2009

User-generated content (UGC) in tourism: benefits and concerns of online consumers

Stephen Burgess  
Victoria University

Carmine Sellitto  
Victoria University

Carmen Cox  
Southern Cross University

Jeremy Buultjens  
Southern Cross University

Publication details
Burgess, S, Sellitto, S, Cox, C & Buultjens, J 2009, 'User-generated content (UGC) in tourism: benefits and concerns of online consumers', in S Newell, E Whitley, N Pouloudi, J Wareham & L Mathiassen (eds), Information systems in a globalising world: challenges, ethics and practices: Proceedings of the 17th European Conference on Information Systems, Verona, Italy, 8-10 June, Department of Business Administration, Faculty of Economics, University of Verona, Verona, Italy.

ePublications@SCU is an electronic repository administered by Southern Cross University Library. Its goal is to capture and preserve the intellectual output of Southern Cross University authors and researchers, and to increase visibility and impact through open access to researchers around the world. For further information please contact epubs@scu.edu.au.
ECIS 2009
The 17th European Conference on Information Systems

INFORMATION SYSTEMS IN A GLOBALISING WORLD: CHALLENGES, ETHICS AND PRACTICES

8-10 June 2009
Verona, Italy
Welcome to ECIS 2009

ECIS 2009 -- Benvenuti a Verona!

Welcome to the 17th European Conference on Information Systems, ECIS 2009, organized jointly by the Department of Management in the Faculty of Economics at the University of Verona, and the Information Systems and Innovation Group of the Department of Management at the London School of Economics and Political Science.

Since its beginning in 1993 ECIS has grown to become the largest and most prestigious Information Systems conference in Europe, attracting scholars from all over the world.

In contrast to past conferences ECIS 2009 does not have specialist tracks, but differentiates only between full research papers and research-in-progress ones. Of 697 submissions, 248 papers will be presented at the conference. In addition, ten excellent panels enrich the programme.

Renewing the traditional structure of the conference required different and more substantial support from the community. We relied on Research Paper Chairs Sue Newell and Edgar Whitley, Research-in Progress-Chairs Nancy Pouloudi and Jonathan Wareham, and Panels Chair Lars Mathiassen, 250 Associate Editors, countless reviewers, and Carla Marisa Bonina for handling the submission process. To all we are very grateful for their excellent work!

We further extend our thanks to the Chairs of the Doctoral Consortium, Gabriele Piccoli and Erica Wagner.

The conference would not be taking place without the enthusiastic support provided by Antonio Cordella and Cecilia Rossignoli, Organisation Chairs, and Lapo Mola who managed the Local Organisation.

A special thank is also for Cinzia Bonomo from Sistema Congressi, Anna Maran and Alessandro Zardini from University of Verona for their support and help in making this event possible.

Finally, we "thank" our delegates, presenters, and session chairs, who will be the real protagonists of the conference.

We are sure the programme and the quality of the participants will make ECIS 2009 in Verona a very memorable event.

Marco de Marco, Claudia Loebbecke, Leslie Willcocks

Conference Co-Chairs
Conference Committee:

Conference Chairs
Professor Marco De Marco (Catholic University, Italy)
Professor Claudia Loebbecke (University of Cologne, Germany)
Professor Leslie Willcocks (London School of Economics, UK)

Research Paper Chairs
Dr. Edgar Whitley (London School of Economics)
Professor Sue Newell (Bentley College, USA)

Research–in–Progress Chairs
Dr. Jonathan Wareham (ESADE, Spain)
Dr. Nancy Pouloudi (Athens University of Economics and Business, Greece)

Panels Chair
Professor Lars Mathiassen (Georgia State University, USA)

Doctoral Consortium Chair
Dr. Gabriele Piccoli (Università di Sassari, Italy)
Dr. Erica Wagner (Cornell University, USA)

Organisation Chairs:
Dr. Antonio Cordella (London School of Economics, UK)
Dr. Cecilia Rossignoli (University of Verona, Italy)

Local Organization
Dr. Lapo Mola (University of Verona, Italy)
Associate Editors, A big thank you!

On behalf of the ECIS 2009 programme chairs and organizing committee we would like to thank you for all your excellent work in acting as an associate editor for ECIS 2009.

This year, we received 680 submissions - up from 566 last year. Nevertheless, because of your excellent work, we were able to make decisions on all of the papers in the normal timescale and notified authors of the decisions on their papers following the programme committee meeting held in Verona in early March.

Around 35% of the papers were accepted and the programme for the conference is now available online at http://www.ecis2009.it/programme.htm.

An important part of the review process is providing an opportunity for the reviewers and associate editors to see how their perspectives, insights and reviews relate to those of the others involved in the process as well as to the final decision on the paper. This allows all involved to learn from the review experience.

The functionality for doing this on Manuscript Central is a little awkward, but if you logon to MC in the usual way, you can now view the full reviews, AE commentary and chair’s decision for all papers you were involved with.

Best wishes

Edgar, Sue, Nancy and Jonathan
Associate Editors:

Mark Aakhus, Rutgers University, USA;
Frederic Adam, University College Cork, Ireland;
Alsayed Alergawy, Magdeburg University, Germany;
Michel Avital, University of Amsterdam, Netherlands;
Karim Axelsson, Linking University, Sweden;
Andrea Back, University St. Gallen, Switzerland;
Richard Baskerville, Georgia State University, Bolivia;
Ronald Batenburg, Utrecht University, Netherlands;
Ulrike Baumoeo, University of Hagen, Germany;
Roman Beck, University of Frankfurt, Germany;
Jörg Becker, European Research Center for Information Systems, Germany;
Frances Bell, University of Salford, UK;
Egon Berghout, University of Groningen, Netherlands;
Nik Bessis, University of Bedfordshire, UK;
Markus Bick, ESCP-EAP European School of Management, Germany;
Paolo Boccardelli, Luiss Guido Carli University, Italy;
Tilo Boehmann, ISS, Germany;
Richard Boland, Case Western Reserve University, United States;
Harry Bouwman, Delft University of Technology, Netherlands;
Sjaak Brinkkemper, Utrecht University, Netherlands;
Laurence Brooks, Brunel University, UK;
Luigi Buglione, Engineering. IT, Italy;
Christopher Bull, Manchester Metropolitan University, UK;
Frada Burstein, Monash University, Australia;
Vincent Buskens, Utrecht University, Netherlands;
Xavier Busquets, ESADE, Spain;
Cinzia Cappiello, Politecnico di Milano, Italy;
Pamela Carter, North Carolina A&T State University, United States;
Andrea Carugati, Aarhus School of Business, Denmark;
Jerry Chang, UNLV, United States;
Michael Chau, The University of Hong Kong, Hong Kong;
Thomas Chesney, Nottingham University Business School, UK;
Mike Chiasson, Lancaster University, UK;
Alina Chircu, Bentley University, United States;
Jyoti Choudrie, University of Hertfordshire, UK;
Torkil Clemmensen, Copenhagen Business School, Denmark;
Sue Conger, & University of Dallas, United States;
Ioanna D. Constantiou, Copenhagen Business School, Denmark;
Stefan Cronholm, & Linkoping University, Sweden;
Mary Daly, University College Cork, Ireland;
G. ‘Hari’ Harindranath, Royal Holloway, University of London, UK;
Khaleed Hassanein, McMaster University, Canada;
Jeremy Hayes, University College Cork, Ireland;
Alessandro D’Atri, CeRSI - Luiss University, Italy;
Linda Dawson, Monash University, Australia;
Giorgio De Michielis, University of Milano - Bicocca, Italy;
Patrick Delfmann, University of Münster, Germany;
Kristine Dery, University of Sydney, Australia;
Gurpreet Dhillon, Virginia Commonwealth University, USA;
Yogesh Dwivedi, Swansea University, United Kingdom;
Amany Elbanna, Loughborough University, UK;
Hakan Enquist, CFA, Sweden;
Owen Eriksson, Dalarna University, Sweden;
Martin Fellenz, Trinity College Dublin, Ireland;
Joseph Feller, University College Cork, Ireland;
Walter Fernandez, Australian National University, Australia;
Elaine Ferneley, University of Salford, UK;
Agata Filipowska, Poznan University of Economics, Poland;
Patrick Finnegan, University of New South Wales, Australia;
Kai Fischbach, University of Cologne, Germany;
Julie Fisher, Monash University, Australia;
Guy Fitzgerald, Brunel University, UK;
Gordon Fletcher, University of Salford, UK;
Chiara Franscalanci, Politecnico di Milano, Italy;
Brent Gallupe, Queen’s University, Canada;
G.R. Gangadharan, Telematica Institute, Netherlands;
Brian Gannon, Birkbeck, University of London, UK;
Janis Gogan, Bentley University, USA;
Gerald Goh, Multimedia University, Malaysia;
Goran Goldkuhl, Linkoping University, Sweden;
Janis Grabis, Riga Technical University, Latvia;
Mary Granger, George Washington University, USA;
Shirley Gregor, Australian National University, Australia;
Catherine Griffiths, London School of Economics, UK;
Marie Griffiths, University of Salford, UK;
Miria Grisot, University of Oslo, Norway;
Norbert Gronau, University of Potsdam, Germany;
Matthew Guah, Erasmus University, Netherlands;
Oliver Guenther, Humboldt-Universitaet zu Berlin, Germany;
Laurence Habib, Oslo University College, Norway;
Felix Hampe, University of South Australia, Australia;
Melisa Handzic, TBA, Bosnia and Herzegovina;
Andrew Hardin, University of Nevada, Las Vegas, USA;
Federico Iannacci, University of Wales, Lampeter, UK;
Giovan Francesco Lanzara, University of Bologna, Italy;
Heejin Lee, Yonsei University, Korea, Republic of;
<table>
<thead>
<tr>
<th>Name</th>
<th>Session</th>
<th>Name</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cathy Urquhart</td>
<td>Session 01 - Gender</td>
<td>Prodromos Tsiavos</td>
<td>Session 43 - Control</td>
</tr>
<tr>
<td>Ping Zhang</td>
<td>Session 02 – E-Gov 1 /</td>
<td>Dianne Wigand</td>
<td>Session 44 – E-Gov 3</td>
</tr>
<tr>
<td>Aurelie Lequerucq</td>
<td>Session 03 - Socio Collaboration</td>
<td>Aurelio Ravarini</td>
<td>Session 45 – Learning And Education 2</td>
</tr>
<tr>
<td>Kalle Lyytinen</td>
<td>Session 04 - Ethics</td>
<td>Hans Van Der Heijden</td>
<td>Session 46 - Agil</td>
</tr>
<tr>
<td>Peter Loos</td>
<td>Session 05 - Enterprise Systems 1</td>
<td>Tom Acton</td>
<td>Session 47 - Adoption</td>
</tr>
<tr>
<td>Sven Carlsson</td>
<td>Session 06 - Strategy</td>
<td>Jan Marco Leimeister</td>
<td>Session 48 - Km And Health</td>
</tr>
<tr>
<td>Milena Head</td>
<td>Session 07 - Soa Methods</td>
<td>Andrea Carugati</td>
<td>Session 49 – Information Systems Development 1</td>
</tr>
<tr>
<td>Nicolas Romano</td>
<td>Session 08 - Economisc 1</td>
<td>Francesco Bolici</td>
<td>Session 50 - Bpr 2</td>
</tr>
<tr>
<td>Bob Galliers</td>
<td>Session 09 – Learning And Education 1</td>
<td>Rolf Wigand</td>
<td>Session 51 – Economics 3</td>
</tr>
<tr>
<td>Paul Alpar</td>
<td>Session 10 - Web 2.0 Cases</td>
<td>Braa Kristin</td>
<td>Session 52 – Project Management</td>
</tr>
<tr>
<td>Nando Pennarola</td>
<td>Session 11 – Governance 1</td>
<td>Frederic Adam</td>
<td>Session 53 – Bpr 3</td>
</tr>
<tr>
<td>Aman Elbanna</td>
<td>Session 12 – Enterprise Systems 2</td>
<td>Richard Baskerville</td>
<td>Session 54 – Km And Learning</td>
</tr>
<tr>
<td>Tina Blegind Jensen</td>
<td>Session 13 - Health 1</td>
<td>Gianmarco Campagnolo</td>
<td>Session 55 – Collaborative Work</td>
</tr>
<tr>
<td>Eugenio Capra</td>
<td>Session 14 - System Use</td>
<td>Andrea Pontigga</td>
<td>Session 56 – Adoption Case</td>
</tr>
<tr>
<td>Michel Avital</td>
<td>Session 15 – Soa Value</td>
<td>Attila Marton</td>
<td>Session 57 Knowledge Sharing</td>
</tr>
<tr>
<td>Tillo Bohmann</td>
<td>Session 16 – Service Innovation</td>
<td>David Sammon</td>
<td>Session 58 – Information Systems Development 2</td>
</tr>
<tr>
<td>Björn Niehaves</td>
<td>Session 17 - Risk / Trust</td>
<td>Paolo De Paoli</td>
<td>Session 59 Is Implementation</td>
</tr>
<tr>
<td>Tony Cornford</td>
<td>Session 18 - New Approaches</td>
<td>Miria Grisot</td>
<td>Session 60 - Health 3</td>
</tr>
<tr>
<td>Hugo Loriet</td>
<td>Session 19 – Web 2.0 1</td>
<td>Annemette Kjaergaard</td>
<td>Session 61 – E-Commerce 1</td>
</tr>
<tr>
<td>Katarina Voutsina</td>
<td>Session 20 – Research And Studies In Is</td>
<td>Jan Ljungberg</td>
<td>Session 62 - Enterprise Systems - Data Warehousing</td>
</tr>
<tr>
<td>Maddalena Sorrentino</td>
<td>Session 21 - E-Gov 2</td>
<td>Aman Elbanna</td>
<td>Session 63 – Adoption Web 2.0 1</td>
</tr>
<tr>
<td>David Wilson</td>
<td>Session 22 – Culture 1</td>
<td>Michael Newman</td>
<td>Session 64 - Is Requirements</td>
</tr>
<tr>
<td>Antonella Ferrari</td>
<td>Session 23 – Is Supply Chain – Ubi_Supplychain</td>
<td>Walter Fernandez</td>
<td>Session 65 - Project</td>
</tr>
<tr>
<td>Francesco Virili</td>
<td>Session 24 - Dec-Agile</td>
<td>Kalinka Kaloyanova</td>
<td>Session 66 – Data Modeling</td>
</tr>
<tr>
<td>Janice Sipior</td>
<td>Session 25 - Social Networks 1</td>
<td>Andrea Resca</td>
<td>Session 67 – Knowledge Management 2</td>
</tr>
<tr>
<td>Hans Van Der Heijden</td>
<td>Session 26 – Enterprise Systems 3</td>
<td>Yingquin Zheng</td>
<td>Session 68 - Design Research</td>
</tr>
<tr>
<td>Maurizio Cavallari</td>
<td>Session 27 - E-Services</td>
<td>Jan Vom Broke</td>
<td>Session 69 - Economics 4</td>
</tr>
<tr>
<td>Jannis Kallikinos</td>
<td>Session 28 - Competences</td>
<td>Rolf Wigand</td>
<td>Session 70 – E-Commerce 2</td>
</tr>
<tr>
<td>Virpi Tuunainen</td>
<td>Session 29 – Mobile Business</td>
<td>Shaikh Maha</td>
<td>Session 71 - Innovation</td>
</tr>
<tr>
<td>Shaikh Maha</td>
<td>Session 30 – Projects And Open Source</td>
<td>Steven Alter</td>
<td>Session 72 - Is Design And Decision Support Systems</td>
</tr>
<tr>
<td>Ioanna Chini</td>
<td>Session 31 - Governance 2</td>
<td>Paolo De Paoli</td>
<td>Session 73 - Green Technology</td>
</tr>
<tr>
<td>Niels Bjørn-Andersen</td>
<td>Session 32 - Knowledgemanagement 1</td>
<td>Jacques Bulchand-Gidumal</td>
<td>Session 74 – Adoption Web 2.0 2</td>
</tr>
<tr>
<td>Carol Saunders</td>
<td>Session 33 - Social Network 2</td>
<td>Giovan Francesco Lanzara</td>
<td>Session 75 - Culture 2</td>
</tr>
<tr>
<td>Pat Finnegam</td>
<td>Session 34 – Open Innovation</td>
<td>Carla Bonina</td>
<td>Session 76 - Is Evaluation</td>
</tr>
<tr>
<td>Guy Fitzgerald</td>
<td>Session 35 - Sme</td>
<td>Chircu, Alina</td>
<td>Session 77 - Mobile</td>
</tr>
<tr>
<td>Ilze Zigurs</td>
<td>Session 36 - Nfc</td>
<td>Maddalena Sorrentino</td>
<td>Session 78 - Sourcing</td>
</tr>
<tr>
<td>Dianne Cyr</td>
<td>Session 37 – Bpr 1</td>
<td>Ole Hanseth</td>
<td>Session 79 - It Governance</td>
</tr>
<tr>
<td>Wolf Ketter</td>
<td>Session 38 – Health 2</td>
<td>Bo Andersson</td>
<td>Session 80 – E-Commerce 3</td>
</tr>
<tr>
<td>Tony Cornford</td>
<td>Session 39 – Inter Organizational Systems</td>
<td>Roy Johnson</td>
<td>Session 81 – Research And Studies In Is 2</td>
</tr>
<tr>
<td>Andrea Resca</td>
<td>Session 40 - Economics 2</td>
<td>Halonen, Raija</td>
<td>Session 82 - Networks</td>
</tr>
<tr>
<td>Ben Eaton</td>
<td>Session 41 – Culture Mobile</td>
<td>Karin Axelsson</td>
<td>Session 83 - E-Service Un The Value Chain</td>
</tr>
<tr>
<td>Aleksi Aaltonen</td>
<td>Session 42 - Web 2.0 2</td>
<td>Hans Ulrich Buhl</td>
<td>Session 84 – Learning</td>
</tr>
</tbody>
</table>
Programme at a glance

SUNDAY, JUNE 7, 2009
09.00 - 17.00 ANCILLARIES:

- OPEN SOURCE AND INNOVATION WORKSHOP
  - Workshop Co-Chairs,
    - Tony Cornford, ISIG, LSE
    - Maha Shaikh, ISIG, LSE

- ESWIS: EUROPEAN STUDENTSWORKSHOP ON INFORMATION SYSTEMS
  - Workshop Co-Chairs
    - Valentina Albano, LUISS Guido Carli, Rome
    - Stefano Za, LUISS Guido Carli

- INVESTIGATING THE POSSIBILITIES OF GREEN IS
  - Workshop Chair
    - Helen Hasan University of Wollongong, Australia

- GLOBAL SOURCING MANAGEMENT WORKSHOP
  - Workshop Chair
    - Leslie Willcocks, ISIG, London School of Economics
    - Eleni Lioliou, ISIG, London School of Economics

18.30 : WELCOME COCKTAIL: GIARDINO GIUSTI

MONDAY, JUNE 8, 2009

<table>
<thead>
<tr>
<th>Time</th>
<th>SESSION 1</th>
<th>SESSION 2</th>
<th>SESSION 3</th>
<th>SESSION 4</th>
<th>SESSION 5</th>
<th>PANEL A</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 AM - 10:30 AM</td>
<td>GENDER</td>
<td>E-GOV 1</td>
<td>SOCIOCOLLABORATION</td>
<td>ETHICS</td>
<td>ENTERPRISE SYSTEMS 1</td>
<td></td>
</tr>
<tr>
<td>10:30 AM - 11:00 AM</td>
<td>STRATEGY</td>
<td>SOA METHODS</td>
<td>ECONOMICS 1</td>
<td>LEARNING AND EDUCATION 1</td>
<td></td>
<td>PANEL B</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>SESSION 6</th>
<th>SESSION 7</th>
<th>SESSION 8</th>
<th>SESSION 9</th>
<th>SESSION 10</th>
<th>PANEL C</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:00 AM - 12:00 PM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WEB 2.0 CASES</td>
<td>PANEL D</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>SESSION 11</th>
<th>SESSION 12</th>
<th>SESSION 13</th>
<th>SESSION 14</th>
<th>SESSION 15</th>
<th>PANEL E</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00 PM - 1:00 PM</td>
<td>GOVERNANCE 2</td>
<td>ENTERPRISE SYSTEMS 2</td>
<td>HEALTH 1</td>
<td>SYSTEM USE</td>
<td>SOA VALUE</td>
<td>PANEL F</td>
</tr>
<tr>
<td>1:00 PM - 2:00 PM</td>
<td>SERVICE INNOVATION</td>
<td>RISK / TRUST</td>
<td>NEW APPROACHES</td>
<td>WEB 2.0 1</td>
<td>RESEARCH AND STUDIES IN IS</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>SESSION 21</th>
<th>SESSION 22</th>
<th>SESSION 23</th>
<th>SESSION 24</th>
<th>SESSION 25</th>
<th>PANEL E</th>
</tr>
</thead>
<tbody>
<tr>
<td>2:00 PM - 3:00 PM</td>
<td>E-GOV 2</td>
<td>CULTURE 1</td>
<td>SUPPLY CHAIN</td>
<td>AGILE</td>
<td>SOCIAL NETWORKS 1</td>
<td></td>
</tr>
<tr>
<td>3:00 PM - 4:00 PM</td>
<td>SERVICES 27</td>
<td>COMPETENCES</td>
<td>MOBILE BUSINESS</td>
<td>PROJECTS AND OPEN SOURCE</td>
<td>ENTERPRISE SYSTEMS 3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>SESSION 29</th>
<th>SESSION 30</th>
<th>SESSION 31</th>
<th>PANEL F</th>
</tr>
</thead>
<tbody>
<tr>
<td>4:00 PM - 5:00 PM</td>
<td>MOBILE BUSINESS</td>
<td>PROJECTS AND OPEN SOURCE</td>
<td>GOVERNANCE 2</td>
<td></td>
</tr>
</tbody>
</table>

Breaking the conventions: a new operating system for the next generation workstation, itsme

KEYNOTE ADDRESS: Prof. Detmar Straub

KEYNOTE: PROF. GIORGIO DE MICHELIS

ECIS2009
**TUESDAY, JUNE 9, 2009**

<table>
<thead>
<tr>
<th>8:00 AM - 4:00 PM</th>
<th>REGISTRATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 AM - 16:30 PM</td>
<td>KEYNOTE ADDRESS: Prof. Dr. Martin Curley Global Director, IT InnovationDirector, Intel Labs Europe</td>
</tr>
<tr>
<td>10:00 AM - 12:30 PM</td>
<td>COFFEE BREAK</td>
</tr>
<tr>
<td>10:30 AM - 11:00 AM</td>
<td>COFFEE BREAK</td>
</tr>
<tr>
<td>11:00 AM - 12:30 PM</td>
<td>LUNCH BREAK</td>
</tr>
<tr>
<td>12:30 PM - 14:00 PM</td>
<td>LUNCH BREAK</td>
</tr>
<tr>
<td>14:00 PM - 15:30 PM</td>
<td>COFFEE BREAK</td>
</tr>
<tr>
<td>15:30 PM - 16:00 PM</td>
<td>COFFEE BREAK</td>
</tr>
<tr>
<td>16:00 PM - 23:00 PM</td>
<td>GALA DINNER: PALAZZO LA GRAN GUARDIA</td>
</tr>
</tbody>
</table>

**WEDNESDAY, JUNE 10, 2009**

<table>
<thead>
<tr>
<th>8:00 AM - 6:00 PM</th>
<th>REGISTRATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 AM - 18:30 PM</td>
<td>KEYNOTE ADDRESS: Prof. Dr. Martin Curley Global Director, IT InnovationDirector, Intel Labs Europe</td>
</tr>
<tr>
<td>10:30 PM - 11:00 PM</td>
<td>COFFEE BREAK</td>
</tr>
<tr>
<td>11:30 PM - 12:30 PM</td>
<td>COFFEE BREAK</td>
</tr>
<tr>
<td>12:30 PM - 13:00 PM</td>
<td>CONCLUSIONS</td>
</tr>
</tbody>
</table>

**ECIS2009**

Proceedings ECIS 2009
DETAILED PROGRAM
### MONDAY PAPERS

**ROOM T4**

#### FROM 11.00 TO 12.30

**SESSION 2**

**E-GOV 1**

**SESSION CHAIR**

PingZhang

**THE CHIASMUS OF DESIGN: PARADOXICAL OUTCOMES IN THE E-GOVERNMENT REFORM OF UK CHILDREN’S SERVICES**

Dave Wastell, Nottingham University Business School, UK; Susan White, Dept. of Applied Social Science, University of Lancaster UK; Karen Broadhurst, Dept. of Applied Social Science, University of Lancaster UK; Chris Hall Centre for Applied Childhood Studies, University of Huddersfield, UK; Sue Peckover Centre for Applied Childhood Studies, University of Huddersfield, UK.

**APPLYING LESSONS LEARNED FROM COUNSELLING: ON NURTURING RELATIONS IN DESIGN PROJECTS**

Diego Calzà, University of Trento, Department of Sociology and Social Research, Italy; Claudia Cattani, eTour, Italy; Gianni Jacucci, University of Trento, Department of Sociology and Social Research, Italy.

**DESIGN CRITERIA FOR PUBLIC E-SERVICES**

Annie Röstlinger Linköping University, Dept of Management and Sweden; Stefan Cronholm Linköping University, Dept of Management and Engineering, Sweden; University of Borås, Business and Informatics Sweden.

#### FROM 14.00 TO 15.30

**SESSION 11**

**GOVERNANCE 2**

**SESSION CHAIR**

Ferdinando Pennarola

**IT-CONTROLLING IN FEDERAL ORGANIZATIONS**

Andreas Schwertsik Technische Universität München, Germany; Petra Wolf, Technische Universität München, Germany; Helmut Krcmar, Technische Universität München, Germany.

**HOW GERMAN HOSPITALS GOVERN IT – AN EMPIRICAL EXPLORATION**

Felix Köbler, Technische Universität München, Germany; Jens Faehling, Universität Kassel, Germany; Jan Marco Leimeister, Technische Universität München, Germany; Helmut Krcmar, Technische Universität München, Germany.

**TOWARDS A SERVICE GOVERNANCE FRAMEWORK FOR THE INTERNET OF SERVICES**

Christian Janiesch, SAP Research CEC Brisbane, SAP Australia Australia; Michael Niemann, Multimedia Communications Lab, Technische Universität Darmstadt, Germany; Nicolas Repp, Multimedia Communications Lab, Technische Universität Darmstadt, Germany.

#### FROM 16.00 TO 17.30

**SESSION 30**

**PROJECTS AND OPEN SOURCE**

**SESSION CHAIR**

MahaShaikh

**INFORMATION TECHNOLOGY AND THE FIRST-LINE MANAGER’S DILEMMA: LESSONS FROM AN ETHNOGRAPHIC STUDY**

Joao Vieira da Cunha, Joao Vieira da Cunha Faculty of Economics, Universidade Nova de Lisboa; Andrea Carugati, Andrea Carugati Århus School of Business, Århus University.

**THE RELATIONSHIP AMONG DEVELOPMENT SKILLS, DESIGN QUALITY, AND CENTRALITY IN OPEN SOURCE PROJECTS**

Donato Barbagallo, Politecnico di Milano, Dipartimento di Elettronica e Informazione, Italy; Chiara Francelancli, Politecnico di Milano, Dipartimento di Elettronica e Informazione, Italy.

**DO BEST PRACTICE FRAMEWORKS FIT OPEN SOURCE SOFTWARE CUSTOMIZATION?**

Steffen Kessler, Institute of Information Systems, Philipps-Universität Marburg, Germany; Paul Alpar, Institute of Information Systems, Philipps-Universität Marburg, Germany.
### Session 3: Socio Collaboration (11.00 - 12.30)

**Session Chair:** Aurelie Lequercq

- **Bitching, Bouncing and Brawling - How Backchannels Brought Colour to Conference Calls**
  - Benjamin Eaton, The London School of Economics and Political Science, UK

- **Development and Validation of a Model for Assessing the Success of Employee Portals**
  - Nils Urbach, European Business School (EBS), Germany; Stefan Smolnik, European Business School (EBS), Germany

- **Exploring the Impact of Real-Time Communication on Media Choice in the Context of Distributed Work**
  - Aleksi Aaltonen, The London School of Economics and Political Science, UK; Benjamin Eaton, The London School of Economics and Political Science, UK

### Session 16: Service Innovation (14.00 - 15.30)

**Session Chair:** Tillo Bohmann

- **The Emergence of a New Form of IS Offshore Enterprise – The Modern Hierarchy**
  - Brian Gannon, University of London, UK; David Wilson, University of London, UK

- **Internal Markets as a Sourcing Option for the Delivery of IS Services: Improving Outsourcing and Insourcing**
  - Jacques Bulchand-Gidumal, University of Las Palmas de Gran Canaria, Facultad de CC.EE. y Empresariales, Spain

- **Sourcing and Automation Decisions in Financial Value Chains**
  - Matthias Henneberger, FIM Research Center Finance & Information Management, University of Augsburg, Germany; Arne Katzmarchik, FIM Research Center Finance & Information Management, University of Augsburg, Germany; Stephan Müller, FIM Research Center Finance & Information Management, University of Augsburg, Germany; Frans-Matthis Pleie, FIM Research Center Finance & Information Management, University of Augsburg, Germany

### Session 25: Social Networks 1 (16.00 - 17.30)

**Session Chair:** Janice Sipior

- **Simmelian Ties, Organizational Justice, and Knowledge Sharing in Virtual Workgroups**
  - Zhi Wei Ho, National University of Singapore, Singapore; Klarissa Chang, National University of Singapore, Singapore

- **Web-Enabled Boundary Spanners and Their Role in the Knowledge Flow Network**
  - Eoin Whelan, University of Limerick, Ireland; Brian Donnellan, NUI Galway, Ireland; Robin Teigland, Stockholm School of Economics, Sweden; Willie Golden, NUI Galway, Ireland

- **Re "Design Networks" Shaped by Their Own Outcomes? Coordination Processes between Actors and Artefacts**
  - Francesco Bolici, OrgLab-DIAM, Cassino University, Italy; Francesco Virili, OrgLab-DIAM, Cassino University, Italy
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Topic</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.00 - 12.30</td>
<td>Session 7</td>
<td>Empirical Comparison of Methods for Information Systems Development According to Soa</td>
<td>Philipp Offermann, Deutsche Telekom Laboratories, Germany; Udo Bub, Deutsche Telekom Laboratories, Germany.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Service-Oriented Architectures: Modeling the Selection of Services and Platforms</td>
<td>Thomas Widjaja, Technische Universität Darmstadt, Germany; Peter BuXmann, Technische Universität Darmstadt, Germany.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Design Factors for Service-Oriented Architecture Applied to Analytical Information Systems: An Explorative Analysis</td>
<td>Barbara Dinter, University of St. Gallen, Institute of Information Management, Switzerland; Florian Stroh, University of St. Gallen, Institute of Information Management, Switzerland</td>
</tr>
<tr>
<td>14.00 - 15.30</td>
<td>Session 12</td>
<td>Teaching Case: Leading the Change - ERP Implementation at Keda</td>
<td>Chi Yeung Fung; Yulin Fang; Huai Qing Wang.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Future Requirements of ERP Software from the Vendors’ Point of View</td>
<td>Norbert Frick, Institute for IS Research, Department of Computer Science, University of Koblenz-Landau, Germany; Petra Schubert, Centre for Applied Information and Communication Technologies - CAICT, Copenhagen Business School, Denmark.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exploring Antecedents of Organisational Adoption of ERP and Their Effect on Performance of Firms</td>
<td>Jiwat Ram, School of Computer and Information Science, University of South Australia, Australia; Malcolm Pattinson School of Computer and Information Science, University of South Australia, Australia.</td>
</tr>
<tr>
<td>16.00 - 17.30</td>
<td>Session 31</td>
<td>Information Infrastructure Governance and Windows of Opportunity</td>
<td>Violeta Sun, University of São Paulo, FEa/UsP - Department of Business Administration, Brasil; Margunn Aanestad, University of Oslo, Department of Informatics, Norway; Espen Skorve, University of Oslo, Department of Informatics, Norway; Gianluca Mischione, Department of Urban and Regional Planning and Geo-Information Management, The Netherlands.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Configuration of Actors and Roles in Establishing ICT</td>
<td>Dongback Seo, Universidad de Chile,, Chile; Ariel La Paz, University of Groningen, The Netherlands.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The Business Intelligence Competence Centre as an Interface Between IT and User Departments in Maintenance and Release Development.</td>
<td>Henning Baars, University of Stuttgart, Germany; Michael Zimmer, University of Stuttgart, Germany; Hans-Georg Kemper, University of Stuttgart, Germany.</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Topic</td>
<td>Authors</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td>FROM 11.00 TO 12.30</td>
<td>6</td>
<td>STRATEGY</td>
<td>Sven Carlsson</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ANTECEDENTS AND DRIVERS OF IT-BUSINESS STRATEGIC ALIGNMENT: EMPirical VALIDATION OF A THEORETICAL MODEL</td>
<td>ALI YAYLA, Binghamton University, USA; QING HU, Florida Atlantic University, USA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAJOR ISSUES IN SISP: INSIGHTS INTO THE MAIN REASON OF SISP FAILURE</td>
<td>ZIJAD PITA, RMIT University, Australia; FRANCE CHEONG, RMIT University, Australia; BRIAN CORBITT, RMIT University, Australia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>THE ROLE OF STRATEGY IN THE EVOLUTION AND INNOVATION OF INFORMATION SYSTEMS: A SIMULATION EXPERIMENT</td>
<td>KEVIN GALLAGHER, Business Informatics Dep., College of Informatics, Northern Kentucky University, USA; BRYAN HOSACK, School of Information Technology, Illinois State University, USA</td>
</tr>
<tr>
<td>FROM 14.00 TO 15.30</td>
<td>18</td>
<td>NEW APPROACHES</td>
<td>Tony Cornford</td>
</tr>
<tr>
<td></td>
<td></td>
<td>THE ICT CONVERGENCE DISCOURSE IN THE INFORMATION SYSTEMS LITERATURE – A SECOND-ORDER OBSERVATION</td>
<td>JAN HERZHOFF, London School of Economics and Political Science, UK.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RESEARCH 2.0: IMPROVING PARTICIPATION IN ONLINE RESEARCH COMMUNITIES</td>
<td>ELAINE FERNELEY, Salford Business School, UK; ALEKSEJ HEINZE, Salford Business School, UK; PAUL CHILD, Virtual Surveys Ltd, Faulkner House, UK.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>THE NOTION OF LIFEWORLD APPLIED TO INFORMATION SYSTEMS RESEARCH</td>
<td>ANDREW BASDEN, RIS, University of Salford, U.K.</td>
</tr>
<tr>
<td>FROM 16.00 TO 17.30</td>
<td>26</td>
<td>ENTERPRISE SYSTEMS 3</td>
<td>Hans vanderHeijden</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A UTILITY-BASED MODEL TO DEFINE THE OPTIMAL DATA QUALITY LEVEL IN  IT SERVICE OFFERINGS</td>
<td>CINZIA CAPIELLO, Politecnico di Milano, Dipartimento di Elettronica e Informazione, Italy; MARCO COMUZZI, City University London, Department of Computing, School of Informatics, UK.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nominated for the best paper award</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNPACKING THE ERP INVESTMENT DECISION: AN EMPirical ASSESSMENT OF  THE BENEFITS AND RISKS</td>
<td>BYRON KEATING, University of Wollongong Australia; TIM COLTMAN, University of Wollongong, Australia; KATINA MICHAEL, University of Wollongong, Australia; VALERIE BAKER, University of Wollongong, Australia.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TOWARDS UNFOLDING CRM IMPLEMENTATION CHALLENGES IN PAKISTAN: A CASE STUDY</td>
<td>HUMA HAMID, Brunel University, UK.</td>
</tr>
</tbody>
</table>
### ROOM T8

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| FROM 11.00 TO 12.30 | SESSION 5   | ENTERPRISE SYSTEMS 1                                                  | **SUCCESS FACTOR VALIDATION FOR GLOBAL ERP PROGRAMMES**
                  |              |                                                                      | GUNTER SEIDEL, Institute of Information Management, University of St. Gallen, Switzerland; ANDREA BACK, Institute of Information Management, University of St. Gallen, Switzerland. |
|               |             |                                                                      | **“AVOIDING MANAGEMENT” OF RESISTANCES DURING IT PRE-IMPLEMENTATION PHASE: A LONGITUDINAL RESEARCH IN A HIGH TECH CORPORATION**
                  |              |                                                                      | RÉGIS MEISSONIER, CEROM, GSCM - Montpellier Business School, France; EMMANUEL HOUZE, CREGOR, IAE, Montpellier II University, France. |
|               |             |                                                                      | **BEHAVIORAL ASPECTS IN THE USE OF ERP SYSTEMS: STUDY OF A GLOBAL ORGANIZATION**
                  |              |                                                                      | SONIA DECOSTER, University of São Paulo, School of Economics, Business Administration and Accountancy, Brazil; RONALDO ZWICKER, University of São Paulo, School of Economics, Business Administration and Accountancy, Brazil. |
| FROM 14.00 TO 15.30 | SESSION 14  | SYSTEM USE                                                          | **EXAMINING THE EFFECT OF USER EXPECTATIONS ON SYSTEM USE ACTIVITY**
                  |              |                                                                      | KEE-YOUNG KWAK, Kookmin University, Graduate School of Business IT, Korea; SONG-WOO OH, Kookmin University, Graduate School of Business IT, Korea. |
|               |             |                                                                      | **THE EFFECTS OF ONLINE COMMENTARY ON USERS’ INFORMATION PROCESSING IN THE CONTEXT OF ONLINE DISCUSSION FORUMS**
                  |              |                                                                      | SUNGHUN CHUNG, Business School, Korea; YEOSUN YOON, Business School, Korea; INGOO HAN, Business School, Korea. |
|               |             |                                                                      | **THE APPLICATION OF A PHENOMENOLOGICAL FRAMEWORK TO ASSESS USER EXPERIENCE WITH MUSEUM TECHNOLOGIES**
                  |              |                                                                      | JESSIE PALLUD, CESAG Research Center, EM Strasbourg Business School, France. |
| FROM 16.00 TO 17.30 | SESSION 29  | MOBILE BUSINESS                                                      | **INTEGRATING VALUE-ADDING MOBILE SERVICES INTO AN EMERGENCY MANAGEMENT SYSTEM FOR TOURIST DESTINATIONS**
                  |              |                                                                      | TOBIAS SCHERER, Goethe-University Frankfurt, Germany; JAN MUNTERMANN, Goethe-University Frankfurt, Germany; HEIKO ROSSNAGEL, Fraunhofer Institute for Industrial Engineering (IAO), Germany. |
|               |             |                                                                      | **TAXONOMY DEVELOPMENT IN INFORMATION SYSTEMS: DEVELOPING A TAXONOMY OF MOBILE APPLICATIONS**
                  |              |                                                                      | ROBERT NICKERSON, San Francisco State University, College of Business, USA; UPKAR VARSHNEY, Georgia State University, USA; JAN MUNTERMANN, Goethe-University Frankfurt, Germany; HENRI ISAAC, Université Paris Dauphine, France. |
|               |             |                                                                      | **EMERGING BUSINESS MODELS AND STRATEGIES FOR MOBILE MIDDLEWARE TECHNOLOGY PROVIDERS: A REFERENCE FRAMEWORK**
<pre><code>              |              |                                                                      | ANTONIO GHEZZI, Politecnico di Milano, Department of Management, Economics and Industrial Engineering, Italy. |
</code></pre>
<table>
<thead>
<tr>
<th>TIME</th>
<th>SESSION</th>
<th>SPEAKER(S)</th>
<th>TITLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FROM 16.00 TO 17.30</td>
<td>SESSION 21</td>
<td>STEFAN HENNINGSSON, HELLE HENRIKSEN</td>
<td>A SAD STORY: THE CASE OF CONSTRAINED INFRASTRUCTURES CAUSED BY IT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STEFAN HENNINGSSON, Center for Applied ICT, Copenhagen Business School, Denmark; HELLE HENRIKSEN, Center for Applied ICT, Copenhagen Business School, Denmark.</td>
</tr>
<tr>
<td></td>
<td>E-GOV 2</td>
<td>ABRAR HAIDER</td>
<td>CONTRIBUTION OF INTERNET TO A DEMOCRATIC SOCIETY</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ABRAR HAIDER, School of Computer and Information Science, University of South Australia, Australia.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ANNE FLEUR VAN VEENSTRA, BRAM KLEIVINK, MARJIN JANSSEN</td>
<td>BARRIERS FOR TRANSFORMATION: IMPEDIMENTS FOR TRANSFORMING THE PUBLIC SECTOR THROUGH E-GOVERNMENT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ANNE FLEUR VAN VEENSTRA, Delft University of Technology, The Netherlands; BRAM KLEIVINK, Delft University of Technology, The Netherlands; MARJIN JANSSEN, Delft University of Technology, The Netherlands.</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Session Chair</td>
<td>Title</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>-----------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>11.00 - 12.30</td>
<td>Session 1</td>
<td>Cathy Urquhart</td>
<td>A REVIEW OF ETHICAL THEORY IN THE 'UPPER ECHELONS' OF INFORMATION SYSTEMS RESEARCH</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GENDER AND VISIBLE MINORITY STATUS: CAREER ADVANCEMENT IN THE CANADIAN INFORMATION AND COMMUNICATIONS TECHNOLOGY SECTOR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MEASURING THE SUCCESS OF INTERVENTION PROGRAMMES DESIGNED TO INCREASE THE PARTICIPATION RATE BY WOMEN IN COMPUTING</td>
</tr>
<tr>
<td>14.00 - 15.30</td>
<td>Session 20</td>
<td>Katarina Voutsina</td>
<td>A SOCIAL NETWORK ANALYSIS OF THE CO-AUTHORSHIP NETWORK OF THE AUSTRALASIAN CONFERENCE OF INFORMATION SYSTEMS FROM 1990 TO 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>USING INTERPRETIVE STRUCTURAL MODELING TO UNCOVER SHARED MENTAL MODELS IN IS RESEARCH</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>THE SIGNIFICANT OTHERS OF SUBJECTIVE NORM SCIENTOMETRIC STUDY OF SUBJECTIVE NORM IN IS TOP-JOURNALS OVER TWO DECADES</td>
</tr>
<tr>
<td>16.00 - 17.30</td>
<td>Session 22</td>
<td>David Wilson</td>
<td>UNPACKING MULTICULTURALISM IN THE ICT WORKPLACE: DIFFERENCES IN RESPONSES TO WORKPLACE SITUATIONS FOR ENGLISH AND NON-ENGLISH SPEAKING BACKGROUNDS.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AN INVESTIGATION OF HOW CULTURE IMPACTS GLOBAL WORK: UNPACKING THE LAYERS OF CULTURE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>COPING WITH CULTURAL AND MATURITY INEQUALITY IN OFFSHORE OUTSOURCING: IS MINIMIZING INTERACTION THE SOLUTION?</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Title</td>
<td>Authors</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>FROM 11.00 TO 12.30</td>
<td>SESSION 8</td>
<td>BID PRICE CONTROL AND DYNAMIC PRICING IN CLOUDS</td>
<td>ARUN ANANDASIVAM, University of Karlsruhe, Germany; MARC PREMM, University of Karlsruhe, Germany.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A MODEL OF PREFERENCE ELICITATION: THE CASE OF DISTRIBUTED RESOURCE ALLOCATION</td>
<td>JOCHEN STÖSSER, Institute of Information Systems &amp; Management, Universität Karlsruhe (TH), Germany; DIRK NEUMANN, Albert-Ludwigs-Universität Germany; TIM PÜSCHEL, Albert-Ludwigs-Universität Germany.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAKING MONEY WITH CLOUDS: REVENUE OPTIMIZATION THROUGH AUTOMATED POLICY DECISIONS</td>
<td>TIM PÜSCHEL, Albert-Ludwigs-Universität Freiburg, Germany; ARUN ANANDASIVAM, Universität Karlsruhe (TH), Germany; STEFAN BUSCHEK, Universität Karlsruhe (TH), Germany; DIRK NEUMANN, Albert-Ludwigs-Universität Freiburg, Germany.</td>
</tr>
<tr>
<td>FROM 14.00 TO 15.30</td>
<td>SESSION 19</td>
<td>NEW LENSES TO INVESTIGATE MEDIA USE: THE LAYERING PROCESS PERSPECTIVE</td>
<td>NABILA BOUKEF CHARKI, Nabila, ESDES Business School, Catholic University of Lyon, France; MICHEL KALIKA, Ecole de Management Strasbourg, Université Robert Schuman, France.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INTERNET SOCIAL NETWORKING – DISTINGUISHING THE PHENOMENON FROM ITS MANIFESTATIONS IN WEB SITES</td>
<td>DANIEL RICHTER, University of Liechtenstein, Liechtenstein; KAI RIEMER, ERCIS, University of Münster, Germany; JAN VOM BROCKE, University of Liechtenstein, Liechtenstein; STEFAN GROSSE BÖCKMANN, University of Liechtenstein, Liechtenstein</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CRIMINALISING FANTASIES: THE REGULATION OF VIRTUAL CHILD PORNOGRAPHY</td>
<td>MARIE ENEMAN, University of Gothenburg, Sweden; ALISDAIR GILLESPIE, De Montfort University, UK; BERND CARSTEN STAHL, De Montfort University, UK.</td>
</tr>
<tr>
<td>FROM 16.00 TO 17.30</td>
<td>SESSION 28</td>
<td>THE EFFECTS OF REGULATORY PRESSURE ON INFORMATION SYSTEM ADOPTION SUCCESS: AN INSTITUTIONAL THEORY PERSPECTIVE</td>
<td>KATHARINA KRELL, The University of Queensland, UQ Business School, Australia; SABINE MATOOK, The University of Queensland, UQ Business School, Australia; FIONA ROHDE, The University of Queensland, UQ Business School, Australia.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A STUDY OF COMPLIANCE MANAGEMENT IN INFORMATION SYSTEMS RESEARCH</td>
<td>SYED NORRISSYED ABDULLAH, The University of Queensland, Australia; MARTAINDULSKA, The University of Queensland, Australia; SHAZIA SADIQ, The University of Queensland, Australia.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ANALYSING THE EFFECT OF SECURITY ON INFORMATION QUALITY DIMENSIONS</td>
<td>MARKUS HELFERT, Dublin City University, Ireland; OWEN FOLEY, Owen, Galway Mayo Institute of Technology, Ireland; MOUZHI GE, Dublin City University, Ireland; CINZIA CAPPIELLO, Politecnico di Milano, DipartimentodiElettronicaeInformazione, Italy.</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Title</td>
<td>Contributors</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>FROM 11.00</td>
<td>4</td>
<td>SESSION 4: MITIGATING RESPONSE DISTORTION IN IS ETHICS RESEARCH</td>
<td>SAMUEL KWAN, HKUS, Hong Kong; KAR TAM, HKUS, Hong Kong.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IDENTITY IN INFORMATION SYSTEMS</td>
<td>URI GAL, Aarhus School of Business, University of Aarhus, Denmark; ANNE METTE KJÆRGAARD, Copenhagen Business School, Denmark.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A SITUATED KNOWLEDGE WORK CONTEXT PERSPECTIVE NON KNOWLEDGE MANAGEMENT SOFTWARE-ORGANISATION MISALIGNMENTS</td>
<td>KHUONG LE-NHUYEN, Royal Holloway College, University of London, UK; ROMANO DYERSON, Royal Holloway College, University of London, UK; G. ‘HARI’ HARINDRANATH, Royal Holloway College, University of London, UK.</td>
</tr>
<tr>
<td>FROM 14.00</td>
<td>17</td>
<td>SESSION 17: TOWARDS OPERATIONAL RISK-AWARE INFORMATION SYSTEMS: A CRITICAL REALIST PERSPECTIVE</td>
<td>KRISTIAN ROTARU, Monash University, Dep. of Accounting &amp; Finance, Australia; CARLA WILKIN, Monash University, Dep. of Accounting &amp; Finance, Australia; ANDRZEJ CEGLOWSKI, Monash University, Dep. of Accounting &amp; Finance, Australia; LEONID CHURILOV, University of Melbourne, Australia.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WHY RISK MANAGEMENT MATTERS IN IT OUTSOURCING – A SYSTEMATIC LITERATURE REVIEW</td>
<td>BENEDIKT MARTENS, University of Osnabrueck, Germany; FRANK TEUTEBERG, University of Osnabrueck, Germany.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXAMINING THE RELATIONSHIP BETWEEN TRUST AND CONTROL IN IT OUTSOURCING RELATIONSHIPS</td>
<td>DANIEL BEIMBORN, Dep. of Information Systems and Services, University of Bamberg, Germany; FRANK SCHLOSSER, Dep. of Information Systems and Services, University of Bamberg, Germany; TIM WEITZEL, Dep. of Information Systems and Services, University of Bamberg, Germany.</td>
</tr>
<tr>
<td>FROM 16.00</td>
<td>27</td>
<td>SESSION 27: PREDOMINANTLY ELECTRONIC OR PERSONAL SERVICE DELIVERY? A CASE IN THE WEALTH MANAGEMENT CONTEXT</td>
<td>ANNE SUNIKKA, Helsinki School of Economics, Department of Business Technology, Finland.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RECONCEPTUALISING THE INFORMATION SYSTEM AS A SERVICE</td>
<td>GUY GABLE, Queensland University of Technology, Australia; ARUN RAI, Georgia State University, USA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MEASUREMENT OF E-SERVICE QUALITY: AN EMPIRICAL STUDY ON ONLINE TRAVEL SERVICE</td>
<td>HONGXIU LI, Turku School of Economics, Turku Center for Computer Science, Finland; YONG LIU, Åbo Akademi University, Turku Center for Computer Science, Finland; REIMA SUOMI, Turku School of Economics, Turku Center for Computer Science, Finland.</td>
</tr>
</tbody>
</table>
# ROOM 1.5

## FROM 11.00 TO 12.30
### SESSION 9
#### LEARNING AND EDUCATION 1
**SESSION CHAIR**
Bob Galliers

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A TEACHING CASE: TOWARDS BRIDGING DISCIPLINARY DIVIDES IN IT EDUCATION</strong></td>
<td>Karen Neville, University College Cork; Ireland; Margaret Healy, University College Cork; Ireland.</td>
</tr>
<tr>
<td><strong>THE ROLE OF TRAINING IN DECREASING ANXIETY AMONG EXPERIENCED COMPUTER USERS</strong></td>
<td>Beretta Sokura, Helsinki School of Economics, Finland; Virpi Tuunainen, Helsinki School of Economics, Finland; Anssi Öörni, Helsinki School of Economics, Finland.</td>
</tr>
<tr>
<td><strong>THE DEVELOPMENT AND TEST OF A RELATIONSHIP MODEL ON SYSTEM USE, JOB LEARNING, AND IMPACT</strong></td>
<td>Jerry Chang, UNLV, Las Vegas, USA; Reza Torkzadeh, UNLV, Las Vegas, USA; Andrew Hardin, UNLV, Las Vegas, USA.</td>
</tr>
</tbody>
</table>

---

## FROM 14.00 TO 15.30
### SESSION 13
#### HEALTH 1
**SESSION CHAIR**
Tina BlegindJense

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>THE &quot;FEAR FACTOR&quot; IN CRITICAL CARE TELE-PEDIATRICS</strong></td>
<td>Janis Gogan, Bentley University, USA; Ryan Baxter, Bentley University, USA; Monica Garfield, Bentley University, USA.</td>
</tr>
<tr>
<td><strong>CONSUMERS’ ACCEPTANCE AND USE OF PERSONAL HEALTH RECORD SYSTEMS: A THEORETICAL MODEL</strong></td>
<td>Vahid Assadi, McMaster University, DeGroote School of, Canada; Khaled Hassanein, McMaster University, DeGroote School of Business, Canada.</td>
</tr>
<tr>
<td><strong>ATTITUDES TO INFORMATION TECHNOLOGY IN HEALTH CARE PROFESSIONS</strong></td>
<td>Ann Svensson, University West, Department of Economics and IT, Sweden; Ulrika Snis, University West, Department of Economics and IT, Sweden; Pia Svanberg, University West, Department of Economics and IT, Sweden; Lars Svensson, University West, Department of Economics and IT, Sweden.</td>
</tr>
<tr>
<td><strong>INNOVATIVE CAPABILITY DEVELOPMENT PROCESS: A SINGAPORE IT HEALTHCARE CASE STUDY</strong></td>
<td>Say-Yen Teoh, RMIT University, Australia; Shun Cai Logistics Institute-Asia Pacific: National University of Singapore, Singapore.</td>
</tr>
</tbody>
</table>

---

## FROM 16.00 TO 17.30
### SESSION 24
#### DEC-AGILE
**SESSION CHAIR**
Francesco Virili

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TOWARDS A REFERENCE MODEL FOR GRASSROOTS ENTERPRISE MASHUP ENVIRONMENTS</strong></td>
<td>Nominated for the best paper award</td>
</tr>
<tr>
<td><strong>HOW FACTORS AFFECTING SELECTION OF IMPLEMENTATION APPROACH INFLUENCE ERP SYSTEM IMPLEMENTATION COSTS.</strong></td>
<td>Björn Johannsson, Copenhagen Business School, Center for Applied ICT, Denmark; Frantisek Sudzina, Copenhagen Business School, Center for Applied ICT, Denmark.</td>
</tr>
<tr>
<td><strong>COMPETITIVE ADVANTAGES OF ELECTRONIC MARKETPLACES IN THE RETAIL AUTOMOTIVE AND MAINTENANCE, REPAIR AND ORDER (MRO) INDUSTRIES.</strong></td>
<td>Carine Dominguez, University Jean Monnet, Saint-Etienne, France.</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
</tr>
<tr>
<td>11.00 - 12.30</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>14.00 - 15.30</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>16.00 - 17.30</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Nominated for the Claudio Ciborra award**
# Tuesday

## Session 37
### BPR1
#### Requirements on IT Business Value Measures for Mobile-Integrated Business Processes
- Kai Habermann, University of Augsburg, Augsburg, Germany
- Key Pousttchi, University of Augsburg, Augsburg, Germany

#### Teaching Case Influences of Six Sigma 'R' Embracement and Abondonment
- Wasana Bandara, Queensland University of Technology, Australia
- Diana Heckl, Frankfurt School of Finance & Management, Frankfurt am Main
- Michael Rosemann, Queensland University of Technology, Brisbane, Australia

**Chair:** Dianne Cyr

## Session 42
### Web 2.0
#### Modelling Use Continuance in Virtual Worlds: The Case of Second Life
- Stuart Barnes, Norwich Business School, UK

#### The Moderating Role of Utilitarian/Hedonic User Motivation on User Behaviour Towards Web 2.0 Applications
- Hsia-Ching Chang, National Taiwan University, Taiwan
- Lunghwa University of Science and Technology, Taiwan
- Chen-Ya Wang, National Taiwan University, Taiwan
- Seng-Cho T. Chou, University at Albany, State University of New York, USA

#### Discussion of Functional Design Options for Online Rating Systems: A State-of-the-Art Analysis
- Axel Winkelmann, University of Koblenz-Landau, Koblenz, Germany
- Sebastian Herwig, University of Münster, ERCIS, Münster, Germany
- Jens Peepelbuess, University of Münster, ERCIS, Münster, Germany
- Daniel Tiebe, University of Münster, ERCIS, Münster, Germany
- Jörg Becker, University of Münster, ERCIS, Münster, Germany

**Chair:** Aleksi Aaltone

## Session 58
### ISD 2
#### Towards a Research Framework for a Human Development-Based "Bottom of the Pyramid" ICT Development Strategy in South Africa
- Walter Brown, Monash University, South Africa
- Irwin Brown, University of Cape Town, Cape Town, South Africa

#### Recognizing Work Priorities and Tasks in Incoming Messages Through Personal Ontologies Supplemented by Lexical Clues
- Roger Tagg, University of South Australia, Adelaide, Australia
- Prashant Gandhi, University of South Australia, Australia
- Raaaj Srinivasan Kumara, University of South Australia, Australia

#### The Emergence of Language Consensus - Intensifying Language Interaction in Information Systems Development
- Marianne Corvera Vargas, Goethe University, Frankfurt, Germany

**Chair:** David Sammon
ROOM T5

FROM 11.00 TO 12.30
SESSION 33
Social Network 2
SESSION CHAIR
Carol Saunders

VALUATION OF ONLINE SOCIAL NETWORKS – AN ECONOMIC MODEL AND ITS APPLICATION USING THE CASE OF XING.COM
MARTIN GNEISER, University of Augsburg, Augsburg, Germany; JULIA HEIDEMANN, University of Augsburg, Augsburg, Germany; MATHIAS KLIER, University of Innsbruck, Innsbruck, Austria; CHRISTIAN WEIÄ, University of Augsburg, Augsburg, Germany.

REVEALING KNOWLEDGE NETWORKS FROM COMPUTER MEDIATED COMMUNICATION IN ORGANIZATIONS
JURRIAAN VAN REIJSEN, Utrecht University, The Netherlands; REMKO HELMS, Utrecht University, The Netherlands

ANALYZING COMMUNITY CONTRIBUTIONS TO THE DEVELOPMENT OF COMMUNITY WIRELESS NETWORKS
ABDELNASSER ABDELAAL, University of Nebraska, Omaha, USA; HESHAM ALI, University of Nebraska, Omaha, USA

FROM 14.00 TO 15.30
SESSION 44
e-GOV 3
SESSION CHAIR
Dianne Wigand

UNDERSTANDING CITIZENS’ BEHAVIOURAL INTENTION IN THE ADOPTION OF E-GOVERNMENT SERVICES IN THE STATE OF QATAR
SHAFI AL-SHAFI, Brunel University, UK; VISHANTH WEERAKKODY, Brunel University, Business School, UK

DOES THE ANSWER LIE IN COLLABORATION? – A CASE STUDY ON E-GOVER
BJÖRN NIEHAVES, University of Münster, (ERCIS), Muenster, Germany; JÖRG BECKER, University of Münster, (ERCIS), Muenster, Germany; KEVIN ORBACH, University of Münster, (ERCIS), Muenster, Germany

THE (MISSING?) VALUE OF IT IN PUBLIC ORGANIZATIONS - THE CASE OF THE SWEDISH RESCUE SERVICES
JAN LJUNGBERG, Gothenburg University, Gothenburg; Sweden; JANE FRISK, Gothenburg University, Gothenburg, Sweden

FROM 16.00 TO 17.30
SESSION 56
ADOPTION CASE
SESSION CHAIR
Andrea Pontiggia

DECISION MODELS AND THE ADOPTION OF WIRELESS TECHNOLOGY
GREGORY GIMPEL, Copenhagen Business School, Denmark.

ACCEPTANCE PROBLEMS OF AMBIENT INTELLIGENCE AND MOBILE TECHNOLOGIES IN HOSPITALS IN INDIA AND GERMANY
TYGE KUMMER, ESCP-EAP Berlin, Germany; MARKUS BICK, ESCP-EAP Berlin, Germany; RAJ GURURAJAN, University of Southern Queensland, Australia

WHY INFORMATION TECHNOLOGY IS NOT BEING USED FOR FINANCIAL ADVISORY
PHILIPP NUISBAUMER, University of Zurich, Zurich, Switzerland; GERHARD SCHWABE, University of Zurich, Zurich, Switzerland
<table>
<thead>
<tr>
<th>TIME</th>
<th>SESSION</th>
<th>CHAIR</th>
<th>TITLE</th>
<th>SPEAKERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FROM 11.00</td>
<td>SESSION 38</td>
<td>Wolf Ketter</td>
<td>HEALTH INFORMATION PORTALS: HOW CAN WE IMPROVE THE USER’S SEARCH EXPERIENCE?</td>
<td>JULIE FISHER, Monash University, Australia; FRADA BURSTEIN, Monash University, Australia; ROSETTA MANASZEWICZ, Monash University, Australia; KATE LAZARENKO, Monash University, Australia</td>
</tr>
<tr>
<td>TO 12.30</td>
<td>HEALTH 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FROM 14.00</td>
<td>SESSION 49</td>
<td>Andrea Carugati</td>
<td>MICROEMANCIPATORY PRACTICES IN INFORMATION SYSTEM DEVELOPMENT</td>
<td>DUBRAVKA CECEZ-KECMANOVIC, Australian School of Business, Sydney, Australia; MARIUS JANSON, University of Missouri-St. Louis, USA</td>
</tr>
<tr>
<td>TO 15.30</td>
<td>INFORMATION SYSTEMS DEVELOPMENT 1</td>
<td></td>
<td>A PRAGMATIC APPROACH TO IS DEVELOPMENT AND SOCIO-TECHNICAL EVALUATION</td>
<td>SABINE MADSEN, Roskilde University, Roskilde, Denmark; RICHARD VIDGEN, University of Bath, Bath, UK</td>
</tr>
<tr>
<td></td>
<td>SESSION 60</td>
<td></td>
<td>TOWARD THE CONCEPT OF POCKETS OF CREATIVITY IN BUSINESS PROCESSES</td>
<td>STEFAN SEIDEL, University of Liechtenstein, Principality of Liechtenstein; FELIX MÜLLER-WIENBERGEN, ERCIS, University of Münster, Münster, Germany; MILAN KAROW, ERCIS, University of Münster, Germany; MICHAEL ROSEMANN, Queensland University of Technology Australia</td>
</tr>
<tr>
<td>FROM 16.00</td>
<td>HEALTH 3</td>
<td>Miria Grisot</td>
<td>COMPONENT-BASED PROCESS MODELLING IN HEALTH CARE</td>
<td>LARS BAACKE, University of St. Gallen, St. Gallen, Switzerland; TOBIAS METTLER, University of St. Gallen, St. Gallen, Switzerland; PETER ROHNER, University of St. Gallen, St. Gallen, Switzerland</td>
</tr>
<tr>
<td>TO 17.30</td>
<td>SESSION 60</td>
<td></td>
<td>AN AGENT-BASED APPROACH TO IMPROVING RESOURCE ALLOCATION IN THE DUTCH YOUTH HEALTH CARE SECTOR</td>
<td>ERIK GIESEN, INITI8, Van Nelleweg 1, Rotterdam, NL; WOLFGANG KETTER, Erasmus University Rotterdam, NLROB ZUIDWIJK, Erasmus University Rotterdam, NL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RFID-ENABLED PROCESS CAPABILITIES AND ITS IMPACTS ON HEALTHCARE PROCESS PERFORMANCE: A MULTI-LEVEL ANALYSIS</td>
<td>MARK LEWIS, Bentley University, Waltham, Massachusetts, USA; BALAJI SANKARANARAYANAN, Bentley University, Waltham, Massachusetts, USA; ARUN RAI, Georgia State University, Atlanta, Georgia, USA</td>
</tr>
</tbody>
</table>
# Room T7

## From 11.00 to 12.30
### Session 34
**Open Innovation**

*Session Chair:* Pat Finnegan

- **Presenting Data for Team-Based Decision-Making in Agile Information Systems Projects**
  - Kieran Conboy, National University of Ireland, Galway, Galway, Ireland; Thomas Acton, National University of Ireland, Galway, Ireland; Raija Halonen, National University of Ireland, Galway, Ireland

- **Enabling Open Innovation: Proposal of a Framework Supporting ICT and KMS Implementation in Web-Based Intermediaries**
  - Sven Carlsson, Lund School of Economics and Management, Lund, Sweden; Vincenzo Corvello, University of Calabria, Rende, Italy; Piero Migliarese, University of Calabria, Rende, Italy

- **Digital Libraries as Information Organizations. The Re-Unfolding of the Memory/Information Paradox**
  - Attila Marton, London School of Economics and Political Science, London, UK.

## From 14.00 to 15.30
### Session 50
**BPR 2**

*Session Chair:* Francesco Bolici

- **The Effect of Information and Communication Technologies, Workplace Re-Organization and Trade on the Demand for Employees Skills: A Comparative Analysis of Greek and Swiss Enterprises**
  - Euripidis Loukis, Federal ETH Zurich, Switzerland; Spyros Arvanitis, University of the Aegean, Island of Samos, Greece

- **Counteracting Forces in Implementation of IS-Enabled Global Business Processes**
  - Møyfrid Sannarnes, University of Agder, Kristiansand, Norway; Bjørn Munkvold, University of Agder, Kristiansand; Kim Andersen, Copenhagen Business School, Howitzvej, Denmark

- **An Approach to Assess the Implementation of Business Process Management in Enterprises**
  - Michael Rohloff, University of Potsdam, Potsdam, Germany

## From 16.00 to 17.30
### Session 61
**E-Commerce 1**

*Session Chair:* Annemette Kjaergaard

- **Would I Use My Personal Blog for Commercial Exchange?**
  - Wee-Kek Tan, National University of Singapore; Chuann-Hoo Tan, City University of Hong Kong, Kowloon, Hong Kong; Hock-Hai Teo, National University of Singapore

- **Consumers’ Perception of Control over Online Information Disclosure. An Electronic Focus Group Study**
  - Mirella Lahteenmäki, Helsinki School of Economics, Finland; Johanna Bragge, Helsinki School of Economics, Finland; Anne Sunikka, Helsinki School of Economics, Finland

- **File-Sharing – A Threat to Intell**
  - Bo Andersson, Lund University, School of Economics & Management, Lund, Sweden; Paul Pierce, Lund University, School of Economics and Management, Lund, Sweden; Markus Lahtinen, Lund University, School of Economics & Management, Lund, Sweden
## ROOM T8

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Title</th>
<th>Chair(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FROM 11.00 TO 12.30</td>
<td>SESSION 39</td>
<td>&quot;WHO IS IN CHARGE AND WHOSE RULES ARE FOLLOWED..?&quot;:POWER IN A INTER-ORGANISATIONAL IS PROJECT</td>
<td>RIITTA HEKKALA, University of Oulu, Finland; CATHY URQUHART, University of Auckland, Auckland, New Zealand; NETTA IIVARI University of Oulu, University of Oulu, Finland</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GENERATIVE MECHANISMS FOR INNOVATION IN INFORMATION INFRASTRUCTURES</td>
<td>BENDIK BYGSTAD, Norwegian School of IT, Oslo, Norway.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A CONCEPTUAL MODELING APPROACH FOR SUPPLY CHAIN EVENT MANAGEMENT (SCEM)</td>
<td>AXEL WINKELMANN, University of Koblenz-Landau, Koblenz, Germany; STEFAN FLEISCHER, University of Münster, (ERCIS), Münster, Germany; SEBASTIAN HERWIG, University of Münster, (ERCIS), Münster, Germany; JÖRG BECKER, University of Münster, (ERCIS), Münster, Germany.</td>
</tr>
<tr>
<td>FROM 14.00 TO 15.30</td>
<td>SESSION 48</td>
<td>DEALING WITH TIGHT COUPLINGS AND MULTIPLE INTERACTIONS IN COMPLEX TECHNOLOGICAL SYSTEMS</td>
<td>MARGUNN AANESTAD, University of Oslo, Oslo, Norway; TINA JENSEN, Aarhus School of Business, Aarhus, Denmark; MIRIA GRISOT, University of Oslo, Oslo, Norway.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>THE MEDIATING ROLE OF IT KNOWLEDGE INTEGRATION CAPABILITY IN THE RELATIONSHIP BETWEEN TEAM PERFORMANCE AND TEAM CLIMATE</td>
<td>LEONARDO CAPORARELLO, Bocconi University Milan, Italy; STEFANO BASAGLIA, Bocconi University Milan, Italy; MASSIMO MAGNI, Bocconi University Milan, Italy.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IT-ENABLED CHANGE INTO THE STRUCTURES OF HEALTH INFORMATION SYSTEMS IN AFRICA: A CASE STUDY IN KENYA</td>
<td>ROBERTA BERNARDI, Warwick Business School, Coventry, UK.</td>
</tr>
<tr>
<td>FROM 16.00 TO 17.30</td>
<td>SESSION 55</td>
<td>A WIKI-BASED APPROACH TO ENTERPRISE ARCHITECTURE DOCUMENTATION AND ANALYSIS</td>
<td>SABINE BUCKL, Technische Universität München, Garching, Germany; FLORIAN MATTHES, Technische Universität München, Garching, Germany; CHRISTIAN NEUBERT, Technische Universität München, Garching, Germany; CHRISTIAN SCHWEDA, Technische Universität München, Garching, Germany.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MOBILE INFORMATION SYSTEMS AND ORGANIZATIONAL CONTROL: A FOUCAULDIAN APPROACH</td>
<td>AURELIE LECLERCQ-VANDELANNOITTE, Aurélie, IESEG School of Management, France.</td>
</tr>
</tbody>
</table>
ROOM 1.1

FROM 16.00 TO 17.30
SESSION 52
PROJECT MANAGEMENT
SESSION CHAIR
Kristin Braa

BIASED PROJECT STATUS REPORTS: A SURVEY OF IS PROFESSIONALS
RON THOMPSON, Wake Forest University, Babcock School of Management, USA; CHARLES IACOVOU, Wake Forest University, Babcock School of Management, USA; H. JEFF SMITH, Miami University, Farmer School of Business, USA.

IS PROJECT EVALUATION METHODOLOGY - SCIENCE OR ART?
FOUAD NAGM, School of Information Systems, Technology and Management, Australian School of Business, Australia; DUBRAVKA CECEZ-KECMANOVIC, School of Information Systems, Technology and Management, Australian School of Business, Australia; MARY ANNE KENNAN, School of Information Systems, Technology and Management, Australian School of Business, Australia.

AN EXPLORATORY EVALUATION OF THREE IS PROJECT PERFORMANCE MEASUREMENT METHODS
CORLANE BARCLAY, University of the West Indies, Departments of Computing & Management Studies, Jamaica; KWEKU-MUATA OSEI-BRYSON, Virginia Commonwealth University, School of Business, USA.
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.00 - 12.30</td>
<td>Session 35</td>
<td>A TRANSACTION COST THEORETICAL ANALYSIS OF SOFTWARE-AS-A-SERVICE (SAAS)-BASED SOURCING IN SMBS AND ENTERPRISES</td>
<td>Alexander Benlian, Ludwig-Maximilians-University of Munich, Institute for Information Systems and New Media, Germany.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONTEMPENT DYNAMICS OF IS ALIGNMENT IN SMES</td>
<td>Marge Levy, Warwick Business School, University of Warwick, UK; Philip Powell, School of Management, University of Bath, UK and University of Groningen, Netherlands; Philip Yetton, Australian School of Business, University of New South Wales, Australia.</td>
</tr>
<tr>
<td>14.00 - 15.30</td>
<td>Session 43</td>
<td>CONTROLLING COMPUTER-BASED MULTITASKING THROUGH PROVISIONING SYSTEMS IN CO-LOCATED LEARNING SETTINGS</td>
<td>Gregory Truman, Information Systems Area, IE Business School, Spain.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MANAGEMENT COLLUSION: KEEPING THE LID ON THE ‘CAN OF WORMS’</td>
<td>Polly Sobepere, Salford Business School, Salford, UK.</td>
</tr>
<tr>
<td>16.00 - 17.30</td>
<td>Session 62</td>
<td>THE ROLE OF BOUNDARY OBJECTS AND BOUNDARY SPANNING IN DATA WAREHOUSING – A RESEARCH-IN-PROGRESS REPORT</td>
<td>Helena Vranesic, Goethe University, Germany; Christoph Rosenkrantz, Goethe University, Germany.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EFFICIENT COMMUNICATION IN FINANCIAL DATA WAREHOUSING PROJECTS – INSIGHTS FROM A MULTIPLE CASE STUDY</td>
<td>Marc Raekers, zeb/information.technology, Germany; Christoph Rosenkrantz, Goethe-University, Germany.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A COMMUNICATION EFFICIENCY MODEL FOR ETL PROJECTS IN FINANCIAL DATA WAREHOUSING</td>
<td>Marc Rakers, zeb/information.technology, Germany.</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Title</td>
<td>Presenters</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>11.00 - 12.30</td>
<td>32</td>
<td>An Ontology-Driven Topic Mapping Approach to Multi-Level Management of E-Learning Resources</td>
<td>Bhavani Sridharan, School of Business Information Technology, RMIT University, Victoria, AU; Hepu Deng, School of Business Information Technology, RMIT University, Victoria, Australia; Brian Corbit, School of Business Information Technology, RMIT University, Victoria, Australia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exploring the Development of Social Alignment Within an Innovation Context</td>
<td>Tadhg Nagle, University College Cork, Ireland; Willie Golden, National University of Ireland, Galway</td>
</tr>
<tr>
<td>14.00 - 15.30</td>
<td>45</td>
<td>A Delphi Study on Collaborative Learning in Distance Education</td>
<td>Susan O'Neill, NUI Galway, Ireland; Murray Scott, NUI Galway, Ireland; Kieran Conboy, NUI Galway, Ireland</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A ‘Uses and Gratifications’ Approach to Understanding the Role of Wiki Technology in Enhancing Teaching and Learning Outcomes</td>
<td>Zixiu Guo, University of New South Wales, Sydney, Australia; Ying Zhang, University of New South Wales, Sydney, Australia; Kenneth Stevens, University of New South Wales, Sydney</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E-Learning Acceptance in Workplace Training: The Case of a Greek Bank</td>
<td>Spiros Borotis, Athens University of Economics and Business, Athens, Greece; Angeliki Poulymenakou, Athens University of Economics and Business, Athens, Greece</td>
</tr>
<tr>
<td>16.00 - 17.30</td>
<td>57</td>
<td>The Impact of Learning Culture and Information Technology Use on Knowledge-Sharing: A Case of KFUPM</td>
<td>Mustafa Eid, KFUPM, Doha, KSA; Nuradden Nuju, KFUPM, Doha, KSA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transfer of Technology and Knowledge - The Story of an Enterprise System Implementation</td>
<td>Abdul Aziz Ahmad, Manchester Business School, University of Manchester, Manchester, UK; Michael Newman, University of Manchester, UK; Norwegian School of Economics &amp; Business Administration, Bergen, Norway</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Digital Differentiation, Software Product Lines, and the Challenge of Isomorphism in Innovation: A Case Study</td>
<td>Lena Andreasonsson, Viktoriainstitute, Gothenburg, Sweden; Ola Henridsson, Viktorialnstitute, Gothenburg, Sweden</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Title</td>
<td>Authors</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>FROM 11.00</td>
<td>40</td>
<td>Economics 2: A Real Options Approach for Valuing Intertemporal</td>
<td>Dennis Diepold, FIM Research Center Finance &amp; Information Management, Germany; Christian Ullrich, FIM Research Center Finance &amp; Information Management, University of Augsburg, Germany; Alexander Wehrmann, Senacor Technologies AG, Germany; Steffen Zimmermann, University of Innsbruck, Austria.</td>
</tr>
<tr>
<td>SESSION</td>
<td>CHAIR</td>
<td>Interdependencies Within a Value-Based IT Portfolio Management - A Risk-Return Perspective</td>
<td>Alessandro Zardini</td>
</tr>
<tr>
<td>FROM 14.00</td>
<td>51</td>
<td>Economics 3: Understanding Suppliers’ Participation in Business-to-</td>
<td>Kitsada Dolpanya, The University of New South Wales, School of Information Systems, Technology and Management, Australia; Lesley Land, The University of New South Wales, School of Information Systems, Technology and Management, Australia; Geoff Dick, The University of New South Wales, School of Information Systems, Technology and Management, Australia.</td>
</tr>
<tr>
<td>Session</td>
<td>CHAIR</td>
<td>Government (B2G) Electronic Auction Markets in the Thai Context</td>
<td>Rolf Wigand</td>
</tr>
<tr>
<td>FROM 16.00</td>
<td>59</td>
<td>IS Implementation: Bridging the Gaps: Packaged Software Implementation</td>
<td>Vuong Nguyen, The National Australian University Canberra, Australia; Walter Fernandez, The National Australian University, Canberra, Australia.</td>
</tr>
<tr>
<td>Time Block</td>
<td>Session Number</td>
<td>Title</td>
<td>Authors</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td>FROM 11.00 TO 12.30</td>
<td>36</td>
<td>NFC BASED SERVICE INNOVATION IN RETAIL: AN EXPLORATIVE STUDY</td>
<td>THOMAS WIECHERT, University of St. Gallen, Switzerland; ANDREAS SCHALLER, Technology Consulting, Germany; FREDERIC THIESSE, University of St. Gallen, Switzerland; ELGAR FLEISCH, University of St. Gallen and D-MTEC, ETH Zurich, Switzerland.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A QUANTITATIVE EVALUATION OF NFC BASED CONTACTLESS PAYMENT SYSTEMS IN RETAIL</td>
<td>THOMAS WIECHERT, ITEM-HSG, University of St. Gallen, Switzerland; FREDERIC THIESSE, ITEM-HSG, University of St. Gallen, Switzerland; ELGAR FLEISCH, University of St. Gallen and D-MTEC, Switzerland.</td>
</tr>
<tr>
<td>FROM 14.00 TO 15.30</td>
<td>47</td>
<td>IN-HOUSE DEVELOPMENT AS AN ALTERNATIVE FOR ERP ADOPTION BY SMES: A CRITICAL CASE STUDY</td>
<td>PLACIDE POBA-NZAOU, Université du Québec à Trois-Rivières, Canada; LOUIS RAYMOND, Université du Québec à Trois-Rivières, Canada.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONSUMER ACCEPTANCE OF BIOMETRICS FOR IDENTITY VERIFICATION IN FINANCIAL TRANSACTIONS</td>
<td>MILENA HEAD, McMaster University, DeGroote School of Business, Canada; MICHAEL BREWARD, McMaster University, DeGroote School of Business, Canada; KHALED HASSANEIN, McMaster University, DeGroote School of Business, Canada.</td>
</tr>
<tr>
<td>FROM 16.00 TO 17.30</td>
<td>53</td>
<td>USABILITY OF IT-SYSTEMS IS MORE THAN INTERACTION QUALITY - THE NEED OF COMMUNICATION AND BUSINESS PROCESS CRITERIA</td>
<td>STEFAN CRONHOLM, Linköping University, Dept of Management and Engineering, Sweden; VINCE BRUNOMIT University, School of Business Information Technology, Australia.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>USING GROUNDED THEORY FOR STUDYING BUSINESS PROCESS MANAGEMENT PHENOMENA</td>
<td>STEFAN SEIDEL, University of Liechtenstein, Liechtenstein; JAN RECKER, Queensland University of Technology, Australia.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONSTRUCTING COMPARABLE BUSINESS PROCESS MODELS WITH DOMAIN SPECIFIC LANGUAGES – AN EMPIRICAL EVALUATION</td>
<td>JÖRG BECKER, University of Münster, Germany; DOMINIC BREUKER, University of Münster, Germany; DANIEL PFEIFFER, University of Münster, Germany; MICHAEL RÄCKERS, University of Münster, Germany.</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Title</td>
<td>Authors</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>----------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>FROM 11.00 TO 12.30</td>
<td>SESSION 41</td>
<td>THE TWO CULTURES AND THE INTERNET REVOLUTION</td>
<td>PAUL LICKER, Oakland University, Rochester, USA; SUSANNAH CAMERON CRICHTON, Cameron Crichton Associates, USA.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OPPORTUNITIES AND CHALLENGES OF MOBILE PERSONALIZATION: AN EXPLORATORY STUDY</td>
<td>SHUK YING HO, The Australian National University, Australia.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>THE IMPACT OF CULTURE ON MOBILE PHONE PURCHASING: A COMPARISON BETWEEN THAI AND BRITISH CONSUMERS</td>
<td>MONTATHIP SRIKES, School of Management, University of Surrey, UK; CATHERINE COLLINS, School of Management, University of Surrey, UK.</td>
</tr>
<tr>
<td>FROM 14.00 TO 15.30</td>
<td>SESSION 46</td>
<td>IT IMPACTS ON OPERATION-LEVEL AGILITY IN SERVICE INDUSTRIES</td>
<td>ONE-KI (DANIEL) LEE, University of Massachusetts Boston, College of Management, USA; PENG XU, University of Massachusetts Boston, College of Management, USA; JEAN-PIERRE KUILBOER, University of Massachusetts Boston, College of Management, USA; NOUSHIN ASHRAFI, University of Massachusetts Boston, College of Management, USA.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNDERSTANDING AGILITY IN SOFTWARE DEVELOPMENT THROUGH A COMPLEX ADAPTIVE SYSTEMS PERSPECTIVE</td>
<td>XIAOFENG WANG, Lero, the Irish Software Engineering Research Centre, Ireland; KIERAN CONBOY, National University of Ireland, Ireland.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A CASE STUDY OF RISK MANAGEMENT IN AGILE SYSTEMS DEVELOPMENT</td>
<td>SHARON COYLE, Centre for Innovation and Structural Change, National University of Ireland, Ireland; KIERAN CONBOY, Centre for Innovation and Structural Change, National University of Ireland, Ireland.</td>
</tr>
<tr>
<td>FROM 16.00 TO 17.30</td>
<td>SESSION 54</td>
<td>KNOWLEDGE CLUSTERS: DEALING WITH A MULTILEVEL PHENOMENON</td>
<td>MAMATA BHANDAR, U21 Global, 5 Shenton Way, Singapore; SHAN PAN, National University of Singapore, Singapore; BERNARD TAN, National University of Singapore, Singapore.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONCEPTUAL CONVERGENCES: POSITIONING INFORMATION SYSTEMS AMONG THE BUSINESS DISCIPLINES</td>
<td>DIRK HOVORKA, Faculty of Business, Technology, and Sustainable Development, Bond University, AU; KAI LARSEN, Leeds School of Business, University of Colorado, USA; DAVID MONARCHI, Information Systems (retired) University of Colorado, USA.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AN EXPLORATION OF THE REAL OR IMAGINED CONSEQUENCES OF INFORMATION SYSTEMS RESEARCH FOR PRACTICE</td>
<td>DAVID AVISON, ESSEC, Cergy-Pontoise, France; SHIRLEY GREGOR, Australian National University, Australia.</td>
</tr>
<tr>
<td>TIME</td>
<td>SESSION</td>
<td>CHAIR</td>
<td>PRESENTERS</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------</td>
<td>--------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>09.00 - 10.30</td>
<td>65</td>
<td>Walter</td>
<td>STEPHEN SMITH, Monash University, Department of Accounting and Finance, Australia; GERAEME SHANKS, The University of Melbourne, Department of Information Systems, Australia; ROBERT JOHNSTON, University College Dublin, School of Business, Ireland; MD MAHBUBUR RAHIM, Monash University, Caulfield School of Information Technology, Australia.</td>
</tr>
<tr>
<td>11.00 - 12.30</td>
<td>76</td>
<td>Carla</td>
<td>OCHINIS MADANAYAKE, Australian National University, Australia; SHIRLEY GREGOR, The Australian National University, Australia; COLLEEN HAYES, The Australian National University, Australia; STEVEN FRASER, Australian National University, Australia.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>JANE FRISK, Applied Information Technology, Gothenburg University, Sweden</td>
</tr>
</tbody>
</table>

**From Business Case to Value Case - Assessing the Organizational Value of IT Investments**

**Measurement, Feedback and Empowerment: Critical Systems Theory as a Basis for Software Process Improvement**

**Objects and Their Participation in the Interdisciplinary Design and Development of Computer Games**

**Service Analysis - A Critical Assessment of the State of the Art**

**Information Systems Evaluation**

**What We Need: Project Managers' Evaluation of Top Management Actions Required for Software Development Projects**
<table>
<thead>
<tr>
<th>TIME</th>
<th>SESSION</th>
<th>ROOM T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>09.00 TO 10.30</td>
<td>63</td>
<td>CONTINUANCE USAGE INTENTION IN MICROBLOGGING SERVICES: THE CASE OF TWITTER</td>
</tr>
<tr>
<td></td>
<td></td>
<td>STUART BARNES, Norwich Business School, UK; MARTIN BÖHRINGER, Technische Universität Chemnitz, Germany.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BEYOND USER ACCEPTANCE: THE DETERMINANTS OF THE INTENTION TO PRODUCE USER CREATED CONTENTS ON THE INTERNET</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HABIN LEE, Brunel University, UK; WOONJUNG KOH, Hanyang University, South Korea; JONG KIM, Hanyang University, South Korea; AHMAD GHONEIM, Brunel University, UK.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INTENTION TO LEARN IN MMOG: EXAMINING THE ROLES OF PEER INTRINSIC AND EXTRINSIC MOTIVATIONS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JOSEPH S L KONG, City University of Hong Kong, Hong Kong; RON C W KWOK, City University of Hong Kong, Hong Kong.</td>
</tr>
<tr>
<td>11.00 TO 12.30</td>
<td>74</td>
<td>BLOGGING TO EXPRESS SELF AND SOCIAL IDENTITIES, ANY ONE?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WEE-KEK TAN, Department of Information Systems, National University of Singapore; HOCK-HAI TEO, Department of Information Systems, National University of Singapore.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNDERSTANDING VIRTUAL WORLD USAGE: A MULTIPURPOSE MODEL AND EMPIRICAL TESTING</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TIBERT VERHAGEN, Knowledge, Information and Networks research group, VU University Amsterdam, The Netherlands; FRANS FELDBERG, Knowledge, Information and Networks research group, VU University Amsterdam, The Netherlands; BART VAN DEN HOOFF, Knowledge, Information and Networks research group, VU University Amsterdam, The Netherlands; SELMAR MEENTS, Knowledge, Information and Networks research group, VU University Amsterdam, The Netherlands.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>USER ACCEPTANCE OF SECOND LIFE: AN EXTENDED TAM INCLUDING HEDONIC CONSUMPTION BEHAVIOURS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NAUMAN SAEED, Faculty of ICT, Swinburne University of Technology, Australia; YUN YANG, Faculty of ICT, Swinburne University of Technology, Australia; SUKU SINNAPPAN, Faculty of Higher Education, Swinburne University of Technology, Australia.</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Title</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>FROM 09.00 TO 10.30</td>
<td>SESSION 64</td>
<td>SERVICE SCENARIOS - A SOCIO-TECHNICAL APPROACH TO BUSINESS SERVICE MODELING</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AN EMPIRICAL EXPLORATION OF REQUIREMENTS ENGINEERING FOR HYBRID PRODUCTS</td>
</tr>
<tr>
<td>FROM 11.00 TO 15.30</td>
<td>SESSION 82</td>
<td>AN EVALUATION OF USER ACCEPTANCE OF A CORPORATE INTRANET</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INTEGRATING VALUE-DRIVEN FEEDBACK AND RECOMMENDATION MECHANISMS INTO BUSINESS INTELLIGENCE SYSTEMS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WHY DO PEOPLE BUY VIRTUAL ITEMS IN VIRTUAL WORLDS? AN EMPIRICAL TEST OF A CONCEPTUAL MODEL</td>
</tr>
</tbody>
</table>
## ROOM T6

### FROM 09.00 TO 10.30
**SESSION 73**  
**GREEN COLLABORATION**  
**SESSION CHAIR**  
Paolo De Paoli

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>GREEN IT: EVERYTHING STARTS FROM THE SOFTWARE</td>
<td>EUGENIO CAPRA Politecnico di Milano, Dipartimento di Elettronica e Informazione, Italy; FRANCESCO MERLO, Politecnico di Milano, Dipartimento di Elettronica e Informazione, Italy</td>
</tr>
<tr>
<td>ENVIRONMENTAL RESPONSIBILITY AND GREEN IT: AN INSTITUTIONAL PERSPECTIVE</td>
<td>MARY DALY; TOM BUTLER</td>
</tr>
<tr>
<td>COLLABORATIVE SOURCING – THE MOTIVATION AND DESIGN OF DEMAND SIDE COMBINATIONS</td>
<td>MARK BORMAN The University of Sydney, Australia</td>
</tr>
</tbody>
</table>

### FROM 11.00 TO 12.30
**SESSION 83**  
**E-SERVICE IN THE VALUE CHAIN**  
**SESSION CHAIR**  
Karin Axelsson

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADOPTION OF CROSS-COMPANY RFID: AN EMPIRICAL ANALYSIS OF PERCEIVED INFLUENCE FACTORS</td>
<td>CHRISTOPH GOEBEL, Humboldt-Universität zu Berlin, Germany; CHRISTOPH TRIBOWSKI, Humboldt-Universität zu Berlin, Germany; OLIVER GUENTHER, Humboldt-Universität zu Berlin, Germany</td>
</tr>
<tr>
<td>AN EVENT-DRIVEN APPROACH TO DYNAMIC SITUATION DETECTION</td>
<td>ANSGER JACOB, Universität Hohenheim, Information Systems 2, Germany; MARCUS MÜLLER, Universität Hohenheim, Information Systems 2, Germany; STEFAN KIRN, Universität Hohenheim, Information Systems 2, Germany</td>
</tr>
<tr>
<td>DEDUCING DEMANDS AT BUSINESS-INTELLIGENCE-SYSTEMS BEYOND UNBUNDLING WITHIN THE EUROPEAN ENERGY MARKETS</td>
<td>JOHANNES BUDER, Technische Universität Bergakademie Freiberg, Germany; CARSTEN FELDEN, Technische Universität Bergakademie Freiberg, Germany</td>
</tr>
<tr>
<td>TIME</td>
<td>SESSION</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
</tr>
<tr>
<td>FROM 09.00</td>
<td>72</td>
</tr>
<tr>
<td>TO 10.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>FROM 11.00</td>
<td>79</td>
</tr>
<tr>
<td>TO 12.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Session Time</td>
<td>Session Name</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------</td>
</tr>
</tbody>
</table>
| FROM 09.00 TO 10.30 | SESSION 71    | Innovation         | NIAMH O RIORDAN, University College Cork, Ireland; FREDERIC ADAM, University College Cork, College Road, Ireland; PHILIP O'REILLY, University College Cork, College Road, Ireland.  
A FRAMEWORK FOR ANALYSING SERVICE ECOSYSTEM CAPABILITIES TO INNOVATE  
CHRISTOPH RIEDL, Technische Universität München, Germany; TILO BÖHMANN, International Business School of Service Management, Germany; JAN MARCO LEIMEISTER, Universität Kassel, Germany; HELMUT KRCMAR, Technische Universität München, Germany.  
SOURCES OF IT DYNAMIC CAPABILITY IN THE CONTEXT OF DATA GENESIS CAPABILITY  
CLAUDIO VITARI, Grenoble Ecole de Management, France. |
| FROM 11.00 TO 12.30 | SESSION 80    | E-Commerce 3        | WEN-HSIEN TSAI, National Central University, Taiwan; JUN-DER LEU, National Central University, Taiwan; WEN-CHIN CHOU, National Central University, Taiwan, Taiwan  
THE DEVELOPMENT OF AN EVALUATION MODEL OF E-COMMERCE WEBSITES FOR THE TAIWANESE AIRLINE INDUSTRY  
HEIKKI LEMPINEN, Helsinki School of Economics, Finland; ESKO PENTTINEN, Helsinki School of Economics, Finland.  
ASSESSING THE BUSINESS VALUE OF ELECTRONIC ORDER-TO-PAYMENT CYCLE  
CLAUDIA HIENERTH, Friedrich-Schiller-Universität Jena, Germany; DAVID MEYER, Wirtschaftsuniversität Wien, Austria.  
A METRICS SYSTEM FOR THE PERFORMANCE MEASUREMENT OF ONLINE DISTRIBUTION CHANNELS OF MULTI-CHANNEL RETAILERS |
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Chair</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>FROM 09.00 TO 10.30</td>
<td>SESSION 66</td>
<td>DATA MODELING</td>
<td>CURRENT TRENDS AND FUTURE DIRECTIONS IN THE PRACTICE OF HIGH-LEVEL DATA MODELING: AN EMPIRICAL STUDY</td>
<td>BARBARA ANGLIM, Faculty of Information Technology, Monash University, Australia; SIMON MILTON, Department of Information Systems, The University of Melbourne, Australia; JAYANTHA RAJAPAKSE, Faculty of Information Technology, Monash University, Australia; RON WEBER, Faculty of Information Technology, Monash University, Australia.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TOWARDS INCREASED COMPARABILITY OF CONCEPTUAL MODELS - ENFORCING NAMING CONVENTIONS THROUGH DOMAIN THESAURI AND LINGUISTIC GRAMMARS</td>
<td>JÖRG BECKER, University of Münster, Germany; PATRICK DELFMANN, University of Münster, Germany; SEBASTIAN HERWIG, University of Münster, Germany; LUKASZ LIS, University of Münster, Germany; ARMIN STEIN, University of Münster, Germany.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A NOVEL DATA QUALITY METRIC FOR TIMELINESS CONSIDERING SUPPLEMENTAL DATA</td>
<td>BERND HEINRICH, Department of Information Systems, University of Innsbruck, Austria; MATHIAS KLIER, Department of Information Systems, University of Innsbruck, Austria.</td>
</tr>
<tr>
<td>FROM 11.00 TO 12.30</td>
<td>SESSION 75</td>
<td>CULTURE 2</td>
<td>GENDER AND WEBSITE DESIGN ACROSS CULTURES</td>
<td>DIANNE CYR, Simon Fraser University, British Columbia, Canada.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CROSS-CONTEXTUAL USE OF INTEGRATED INFORMATION SYSTEMS</td>
<td>GASPARAS JARULAITIS, Norwegian University of Science and Technology, Norway; ERIC MONTEIRO, Norwegian University of Science and Technology, Norway.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DOES TIME MATTER? THE ROLE OF ICT IN SHAPING TEMPORAL ASSUMPTIONS</td>
<td>DANIELA ISARI, Catholic University of Milan, Italy.</td>
</tr>
<tr>
<td>Session</td>
<td>Title</td>
<td>Authors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09.00</td>
<td>STORYTELLING AS A TOOL FOR KNOWLEDGE TRANSFER IN THE IT INDUSTRY</td>
<td>ERIK WENDE, University of Zurich, Department of Informatics, Switzerland; PARISSA HAGHIRIAN, Shohpia University, Japan.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>KNOWLEDGE SHARING IN ONLINE COMMUNITIES</td>
<td>ZHAO LI MENG, Renmin University, School of Information, P. R. China; JIONG GONG University of International Business and Economics, School of International Trade and Economics, Department of Economics, China.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>USE OF BLOGS FOR COLLABORATION IN ORGANIZATIONS</td>
<td>DINESH MIRCHANDANI, University of Missouri – St. Louis, USA; THANAPORN SUNDARAVEJ, University of Arizona, USA; JIESI CHENG, University of Missouri - St. Louis, USA.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.00</td>
<td>MANAGING UNCERTAINTY IN SERVICE PRODUCTION WITH MOBILE SYSTEMS - CASE WASTE MANAGEMENT COMPANY</td>
<td>MIIRA JUNTUMAA; Helsinki School of Economics, Finland; THERESA LAURAEUS-NIINIVAARA, Helsinki School of Economics, Finland; VIRPI TUUNAINEN, Helsinki School of Economics, Finland; ANSSI ÖÖRNI, Helsinki School of Economics, Finland.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A PRELIMINARY LADDERING ANALYSIS ON MOBILE SERVICES USAGE</td>
<td>PATRICIA MCMANUS, Edith Cowan University, Australia; CRAIG STANDING, Edith Cowan University, Australia; RAFFAELE ZANOLI, Università Politecnica delle Marche, Italy.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Topic</td>
<td>Authors</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>09.00 - 10.30</td>
<td>68</td>
<td>Problem Solving Patterns in Design Science Research – Learning From Engineering</td>
<td>Anke Gericke, Institute of Information Management, University of St. Gallen, Switzerland</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Action Research and Design Science Research – Seemingly Similar But Decisively Dissimilar</td>
<td>Nominated for the best paper award Juhani Iivari, University of Oulu, Finland; John Venable, School of Information Systems, Curtin University of Technology, Australia</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Scientific Progress of Design Research Artefacts</td>
<td>Christian Fischer, University of St. Gallen, Inst. of Information Management, Switzerland; Stephan Aier, University of St. Gallen, Inst. of Information Management, Switzerland</td>
<td></td>
</tr>
<tr>
<td>11.00 - 12.30</td>
<td>81</td>
<td>A Design Research Study on Enhancing Creativity – The Case of Developing Product-Service Bundles</td>
<td>Felix Müller-Wienbergen, (ERCIS), University of Münster, Germany; Stefan Seidel, University of Liechtenstein, Liechtenstein; Oliver Müller, (ERCIS), University of Münster, Germany; Ralf Knackstedt, (ERCIS), University of Münster, Germany; Jörg Becker, (ERCIS), University of Münster, Germany.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>From Data Warehouses to Transformation Hubs - A Conceptual Architecture</td>
<td>Hans-Georg Kemper, University of Stuttgart, Germany; Henning Baars, University of Stuttgart, Germany</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process</td>
<td>Jan vom Brocke, Institute of Information Systems, University of Liechtenstein; Alexander Simons, Institute of Information Systems, University of Liechtenstein; Björn Niehaves, ERCIS, University of Münster, Leonardo-Campus 3, Münster, Germany; Kai Riemer, ERCIS, University of Münster, Germany; Ralf Plattfaut, ERCIS, University of Münster, Germany; Anne Cleven, Institute of Information Management, University of St. Gallen, Switzerland</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Topic</td>
<td>Authors</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>09.00 - 10.30</td>
<td>Session 69</td>
<td>ECONOMICS 4</td>
<td>GAME-THEORETIC ANALYSIS OF PAY-AS-BID MECHANISMS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>THOMAS MEINL, Universität Karlsruhe (TH), Germany; Jochen STÖSSER, Universität Karlsruhe (TH), Germany; DIRK NEUMANN, Albert-Ludwigs-Universität Freiburg, Germany.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ALGORITHMIC TRADING ENGINES VERSUS HUMAN TRADERS - DO THEY BEHAVE DIFFERENT IN SECURITIES MARKETS?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MARKUS GSELL, Goethe-University Frankfurt, Germany; Peter GOMBER, Goethe-University Frankfurt, Germany.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>COORDINATING SERVICE COMPOSITION</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BENJAMIN BLAU, Universität Karlsruhe (TH), Germany; TOBIAS CONTE, Research Center for Information Technology (FZI), Germany; THOMAS MEINL, Universität Karlsruhe (TH), Germany.</td>
<td></td>
</tr>
<tr>
<td>11.00 - 12.30</td>
<td>Session 78</td>
<td>SOURCING SESSION</td>
<td>IMPACTS OF INFORMATION TECHNOLOGY (IT) OUTSOURCING ON ORGANIZATIONAL PERFORMANCE: A FIRM-LEVEL EMPIRICAL ANALYSIS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>YASIN OZCELİK, Fairfield University, Dolan School of Business, USA; KEMAL ALTINKEMER, Purdue University, Krannert Graduate School of Management, USA.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ANTECEDENTS OF SUCCESS IN IS OFFSHORING PROJECTS – PROPOSAL FOR AN EMPIRICAL RESEARCH STUDY</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MARKUS WESTNER, Dresden University of Technology, Dresden, Germany.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MANAGING AN IT CARVE OUT AT A MULTI-NATIONAL ENTERPRISE - A TEACHING CASE</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>JENS FAELING, Technische Universität München, Germany; Jan Marco LEIMEISTER, Universität Kassel, Germany; Philip YETTON, University of New South Wales, Australian School of Business, Australia; Helmut KRCMAR, Technische Universität München, Germany.</td>
<td></td>
</tr>
<tr>
<td>FROM 09.00 TO 10.30</td>
<td>SESSION 70</td>
<td>E-COMMERCE 2</td>
<td>SESSION CHAIR</td>
<td>Rolf Wigand</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------</td>
<td>-------------</td>
<td>---------------</td>
<td>-------------</td>
</tr>
<tr>
<td>UNDERSTANDING THE MANIFOLD FORMS OF B2B INTEGRATION - A TRANSACTION COST PERSPECTIVE</td>
<td>CHRISTINE LEGNER. European Business School (EBS), Germany.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-DISPUTES AT THE CROSSROADS: A STAKEHOLDER ANALYSIS OF ON-LINE DISPUTE RESOLUTION MECHANISMS (ODR)</td>
<td>AIKATERINA SIDIROPOULOU, Middlesex University, Law Dep., UK; EVANGELOS MOUSTAKAS, Middlesex University, Marketing and Enterprise Dep., UK.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEVELOPING PUBLIC E-SERVICES FOR SEVERAL STAKEHOLDERS – A MULTIFACETED VIEW OF THE NEEDS FOR AN E-SERVICE</td>
<td>KARIN AXELSSON, Linköping University, Department of Management and Engineering, Information Systems, &amp; Swedish Business School at Örebro University, Sweden; ULF MELIN, Linköping University, Department of Management and Engineering, Sweden; IDA LINDGREN, Linköping University, Department of Management and Engineering, Sweden.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FROM 11.00 TO 12.30</th>
<th>SESSION 84</th>
<th>LEARNING SESSION CHAIR</th>
<th>Andrea Resca</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOCUS GROUPS AND CRITICAL SOCIAL IS RESEARCH: HOW THE CHOICE OF METHOD CAN PROMOTE EMANCIPATION OF RESPONDENTS AND RESEARCHERS</td>
<td>BERND STAHL De Montfort University, UK; MONICA CHIARINI TREMBLAY, Florida International University, Decision Sciences &amp; Information Systems, USA; CYNTHIA LEROUGE, Saint Louis University, Decision Sciences/ Information Technology Management, USA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DESIGNING AND EVALUATING AN INTERACTIVE VIDEO WEBSITE FOR ORGANIZATIONAL LEARNING</td>
<td>STEFAN HRASTINSKI, Computer and Systems Science, Department of Information Science, Uppsala University, Sweden; TERESE MONSTAD, Media and Communication Science, Dep. of Information Science, Uppsala University, Sweden</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MULTI-LEVEL ORGANISATIONAL SENSEMAKING AND LEARNING AS ENABLERS OF INFORMATION SHARING TO MITIGATE IDENTITY AND RELATED CRIMES</td>
<td>ROGER JAMIESON, University of New South Wales, Australia; LESLEY LAND, University of New South Wales, Australia; GREGORY STEPHENS, University of New South Wales, Australia; DONALD WINCHESTER, University of New South Wales, Australia.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MONDAY

PANEL A: Room T.1 - 11.00-12.30
BUSINESS PROCESS MANAGEMENT EDUCATION IN ACADEMIA: ITS STATUS, ITS CHALLENGES AND ITS FUTURE
PANELISTS:
ALINA CHIRCU; WASANA BANDARA; RICHARD WELKE; DIMITRIS KARAGIANNIS; ANTONIE VAN RENSBURG AND JAN RECKER

PANEL B: Room 1.1- 11.00-12.30
AGILE SOFTWARE DEVELOPMENT: TIME TO TAKE IT SERIOUSLY?
PANELISTS:
AMANY ELBANNA; GUY FITZGERALD; KIERAN CONBOY AND DERRICK MURRAY

PANEL C: Room T.1 - 14.00-15.30
IS JOURNALS IN WHICH EUROPEANS SHOULD PUBLISH MORE
PANELISTS:
CAROL SAUNDERS, CAROL BROWN, JANICE SIPIOR, PING ZHANG, ILZE ZIGURS.

PANEL D: Room 1.1 - 14.00-15.30
LABOR IMPORTATION AS THE STEROID FOR NATIONS’ ICT SUCCESS: A DEBATE
PANELISTS:
ERRAN CARMEL; ANTHONY D’COSTA; BILL DELONE AND WILLIE GOLDEN.

PANEL E: Room T.1 - 16.00-17.30
WHY DO WE TOIL? BENEFITING RESEARCH AT THE COST OF PRACTICE OR VICE VERSA?
PANELIST:
STEVEN ALTER; RAIJA HALONEN; BJÖRN NIEHAVES; MAUNG SEIN; CATHY URQUHART AND PING ZHANG
TUESDAY

PANEL F: Room T.1 -11.00-12.30
REGULATION AND GOVERNANCE IN COMMONS-BASED PEER (SOCIAL) PRODUCTION
PANELISTS:
JANNIS KALLINIKOS; GIOVAN FRANCESCO LANZARA; OLE HANSETH; PRODROMOS TSIAVOS.

PANEL G: Room 1.1 -11.00-12.30
IS PHD RESEARCH IN THE 21ST CENTURY – A TALE OF CANDIDATES AND THEIR SUPERVISORS
PANELISTS:
RALPH SPRAGUE; FREDERIC ADAM; SVEN CARLSSON; DAVID SAMMON; GRAEME SHANKS; DANIEL MOODY AND CHRISTINA KELLER.

PANEL H: Room T.1 -14.00-15.30
IS HAS OUTGROWN THE NEED FOR REFERENCE DISCIPLINE THEORIES, OR HAS IT?
PANELISTS:
NIEDERMAN, FRED, CAROL SAUNDERS, KALLE LYTTINEN, RICHARD BASKERVILLE AND ELLEN CHRISTIAANSE

PANEL I: Room 1.1 -14.00-15.30
"ICT-ENABLED GLOBAL WORK – PAST PRESENT AND FUTURE"
PANELISTS:
ADIR EVEN; NAVA PLISKIN; KATHLEEN CURLEY; ERRANCARMEL AND STEFFENZIMMERMANN

PANEL L: Room T.1 16.00-17.30
A CALL FOR ACTION IN TACKLING ENVIRONMENTAL SUSTAINABILITY THROUGH GREEN INFORMATION TECHNOLOGIES AND SYSTEMS
PANELISTS:
JANE WEBSTER; MARIE-CLAUDE BOUDREAU; BRIAN DONNELLAN; STEVE ELLIOT; MARK HUBER; TRACY JENKIN AND CHARLES SHERIDAN.
PANELS

MONDAY 8TH JUNE: 11.00 12.30
ROOM T1
PANEL A
ECIS2009-0766.R1
BUSINESS PROCESS MANAGEMENT EDUCATION IN ACADEMIA: ITS STATUS, ITS CHALLENGES AND ITS FUTURE
PANELISTS
ALINA CHIRCU; WASANA BANDARA; RICHARD WELKE; DIMITRIS KARAGIANNIS; ANTONIE VAN RENSBURG AND JAN RECKER

ROOM 1.1
PANEL B
ECIS2009-0762.R1
AGILE SOFTWARE DEVELOPMENT: TIME TO TAKE IT SERIOUSLY?
PANELISTS:
AMANY ELBANNA; GUY FITZGERALD; KIERAN CONBOY AND DERRICK MURRAY

MONDAY 8TH JUNE: 14.00 15.30
ROOM T1
PANEL C
ECIS2009-0767.R1
IS JOURNALS IN WHICH EUROPEANS SHOULD PUBLISH MORE
PANELISTS:
CAROL SAUNDERS, CAROL BROWN, JANICE SIPIOR, PING ZHANG, ILZE ZIGURS.

ROOM 1.1
PANEL D
ECIS2009-0756.R1
LABOR IMPORTATION AS THE STEROID FOR NATIONS’ ICT SUCCESS: A DEBATE
PANELISTS:
ERRAN CARMEL; ANTHONY D’COSTA; BILL DELONE AND WILLIE GOLDEN.

MONDAY 8TH JUNE: 16.00 17.30
ROOM T1
PANEL E
ECIS2009-0754.R1
WHY DO WE TOIL? BENEFITING RESEARCH AT THE COST OF PRACTICE OR VICE VERSA?
PANELIST:
STEVEN ALTER; RAIJA HALONEN; BJÖRN NIEHAVES; MAUNG SEIN; CATHY URQUHART AND PING ZHANG
TUESDAY 9TH JUNE 11.00 12.30
ROOM T1
PANEL F
ECIS2009-0759.R1
REGULATION AND GOVERNANCE IN COMMONS-BASED PEER (SOCIAL) PRODUCTION
PANELISTS:
JANNIS KALLINIKOS; GIOVAN FRANCESCO LANZARA; OLE HANSETH; PRODROMOS TSIAVOS.

ROOM 1.1
PANEL G
ECIS2009-0764.R1
IS PHD RESEARCH IN THE 21ST CENTURY – A TALE OF CANDIDATES AND THEIR
SUPERVISORS
PANELISTS:
RALPH SPRAGUE; FREDERIC ADAM; SVEN CARLSSON; DAVID SAMMON; GRAEME SHANKS;
DANIEL MOODY AND CHRISTINA KELLER.

TUESDAY 9TH JUNE 14.00 15.30
ROOM T1
PANEL H
ECIS2009-0758.R1
IS HAS OUTGROWN THE NEED FOR REFERENCE DISCIPLINE THEORIES, OR HAS IT?
PANELISTS:
FRED NIEDERMAN, CAROL SAUNDERS, KALLE LYTTINEN, RICHARD BASKERVILLE AND ELLEN
CHRISTIAANSE

ROOM 1.1
PANEL I
ECIS2009-0757.R1
"ICT-ENABLED GLOBAL WORK – PAST PRESENT AND FUTURE"
PANELISTS:
ADIR EVEN; NAVA PLISKIN; KATHLEEN CURLEY; ERRAN CARMEL AND STEFFEN ZIMMERMANN

TUESDAY 9TH JUNE 16.00 17.30
ROOM T1
PANEL L
ECIS2009-0752.R1
A CALL FOR ACTION IN TACKLING ENVIRONMENTAL SUSTAINABILITY THROUGH GREEN
INFORMATION TECHNOLOGIES AND SYSTEMS
PANELISTS:
JANE WEBSTER; MARIE-CLAUDE BOUDREAU; BRIAN DONNELLAN; STEVE ELLIOT; MARK HUBER;
TRACY JENKIN AND CHARLES SHERIDAN.
A Review of Ethical Theory in the ‘Upper Echelons’ of Information Systems Research

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0070.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Ethics, E-commerce (B2B / B2C / B2G / G2C), IS Philosophy, Trust</td>
</tr>
</tbody>
</table>
A Review of Ethical Theory in the ‘Upper Echelons’ of Information Systems Research

Bull, Christopher,
Business Information Technology and Management Science,
Manchester Metropolitan University Business School, Manchester M1 3GH, UK,
c.bull@mmu.ac.uk

Abstract
Despite some notable and rare exceptions and after many years of relatively neglect (particularly in the ‘upper echelons’ of IS research), there appears to be some renewed interest in Information Systems Ethics (ISE). This paper reflects on the development of ISE by assessing the use and development of ethical theory in contemporary IS research with a specific focus on the ‘leading’ IS journals (according to the Association of Information Systems). The focus of this research is to evaluate if previous calls for more theoretically informed work are permeating the ‘upper echelons’ of IS research and if so, how (Walsham 1996; Smith and Hasnas 1999; Bell and Adam 2004). For the purposes of scope, this paper follows on from those previous studies and presents a detailed review of the leading IS publications between 2005 to 2007 inclusive. After several processes, a total of 32 papers are evaluated. This review highlights that whilst ethical topics are becoming increasingly popular in such influential media, most of the research continues to neglect considerations of ethical theory with preferences for a range of alternative approaches. Finally, this research focuses on some of the papers produced and considers how the use of ethical theory could contribute.

Keywords: Information Systems Research, Information Systems Ethics and Ethical Theory

Introduction
Information Systems Ethics (ISE) is an important economic, social, cultural and political topic. This research argues for the development of ISE for the similar reasons that are offered for IS (Davis 2000; Baskerville and Myers 2002; Paul 2007a; Paul 2007b) and because IS phenomena are not always well represented by general ethics, business ethics or the various branches of computer ethics. However, the development of ISE should also appreciate some important work in adjacent fields e.g. computer, information and internet ethics (Wiener 1948; Parker 1968; Weizenbaum 1976; Johnson 1985; Moor 1985; Floridi 1999; Tavani 2007), the rare but notable exceptions within the field of Information Systems (IS) (Walsham 1996; Banerjee, Cronan et al. 1998; Adam 2001; Introna 2002; Wagner and Newell 2004) and capitalise on the renewed interest being shown by the International Conference on Information Systems (ICIS) in 2008 and the European Conference on Information Systems (ECIS), 2008 and 2009. The development of ISE is also important given the relatively poor historic coverage of such work, especially in terms of the quantity of research and of considerations of philosophical ethical theory (Walsham 1996; Smith and Hasnas 1999; Bell and Adam 2004), which is somewhat surprising given the interest in other IS topics e.g. politics and power (Markus 1983; Knights and Murray 1994; Jaspersen, Carte et al. 2002; Howcroft and Light 2006; Avergou and McGrath 2007). Furthermore, the need to develop ISE is also important given the increasing economic, social and political significance of IS. In terms of developing ISE, there are various views on how to proceed. Some have long called for more considered reflections of philosophical ethics in IS (Walsham 1996) and others have argued for ethical theories based on ‘less abstract’ approaches (Smith and Hasnas 1999). In 2004, the response had been fairly disappointing (Bell and Adam 2004).
The focus of this research is to follow such studies and assess if more research is being devoted to ethical topics in IS research beyond the specialist media e.g. Ethics and Information Technology and conference e.g. ETHICOMP, and CEPE etc. This research will evaluate the quantity of work published in some of the ‘upper echelon’ IS journals and how such work is considered. This paper is structured as follows; first, in order to understand the broader context of the research questions a brief overview of what is meant by philosophical ethical theory and some of the main developments in computer, information and IS ethics are presented. Next, the paper presents two views in relation to ISE research (Walsham 1996) and (Smith and Hasnas 1999). This is then followed by a review of ISE work published in the ‘upper echelon’ IS journals (according to the Association of Information Systems) from January 2005 to October 2007 inclusive, in the following journals; Management Information Systems Quarterly, Information Systems Research, Management Science and the Journal of Management Information Systems is considered. Finally, the paper critically reflects on the nature of such research in light of the criteria adopted within this review.

**Ethical Theory**

It is beyond the scope and purpose of this paper to evaluate the full complexities of philosophical ethical theory. Many contemporary sources offer good overviews of the main ideas associated with ethical theory (Johnson 1985; Johnson and Nissenbaum 1995; Walsham 1996; Tavani 2007). However, for the purposes of clarifying this research, ethics relates to the study of morality and the use of such theories in judging issues of right and wrong (Gert 1998; Gert 1999; Tavani 2007). Some claim that there are three moral systems; religious, legal and philosophical ethics (Tavani 2007). Philosophical ethics is an informal system where sanctions are often socially orientated e.g. disapproval or ostracism (Tavani 2007). Within philosophical ethics there are a range of diverse and well established ethical theories. The most popular theories are; deontological (duty-based acts) (Ross 1930; Kant 1988), utilitarianism (consequential-based acts) (Bentham 1948), contract or rights-based acts (Hobbes 1972) and virtue (character-based acts) (MacIntyre 1985).

**The Development of Ethical Theory in IS Research**

Historically, some of the most interesting contributions toward an idea of ISE have been made by those working within adjacent academic fields and by a few notable exceptions from those within IS research. The shaping of ISE from the adjacent academic fields may be explained (to some extent) by the relative infancy of IS and of interests in other subjects. Such work includes; Wiener’s texts on the likely impacts of computers on society (Wiener 1954), Maner’s early work on privacy, crime, and dependency (Maner 1980), both Moor’s and Johnson’s work in relation to clarifying computer ethics (Johnson 1985; Moor 1985), Floridi’s philosophical development of information ethics (Floridi 1999; Floridi and Sanders 2002) and Tavani’s work in cyber ethics (Tavani 2007). Within IS, there have also been a few rare notable contributions; Walsham’s use of philosophical ethics to reflect on the limitations of professional codes of conduct (Walsham 1996), Adam’s contributions on gender issues (Adam 2001), Introna’s work in respect of new media (Introna 2002) and some observations of the ethical nature of packaged software systems (Adam and Light 2004; Wagner and Newell 2004; Adam and Bull 2008). However, whilst many have made some useful contributions to the development of IS (Banville and Landry 1989; Baskerville and Myers 2002; Paul 2007a) and other (broadly similar) topics are much healthier (see comments made in respect of the literatures on politics and power earlier, it is somewhat surprising that previous ‘calls to arms’ (see earlier comments) have struggled to ‘take-off’. So what has happened to ISE since?

In order to assess the condition of ISE, this research conducted an evaluation of ethically related work published in the ‘upper echelon’ IS journals (according to the Association of Information Systems) from January 2005 to October 2007 inclusive. These dates are
appropriate as they follow on from the work of other reviews. This research recognises the debates about journal quality and the nature of such media e.g. would such work be submitted to such publications etc. However, this research is ongoing and for the purposes of scope it begins by evaluating the following publications; Management Information Systems Quarterly (MISQ), Information Systems Research (ISR), Management Science (MS) and the Journal of Management Information Systems (JMIS). This review process used the search engine Business Source Premier / EBSCO in conjunction with some independent evaluation of the data to produce its results. The search began by seeking to find any work that was identified as covering ethics (this may or may not be a Key term). This resulted in finding just 5 papers (see table 1).

Table 1: All Papers That EBSCO Highlights As Covering The Term ‘Ethics’

<table>
<thead>
<tr>
<th>Paper</th>
<th>Journal</th>
<th>IS Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethical Decision Making in Software Piracy: Initial Development and Test of a Four-Component Model, (Moores and Cha-Jan Chang 2006)</td>
<td>MISQ 30 (1) 167-180</td>
<td>Yes</td>
</tr>
<tr>
<td>Recipient Choice Can Address the Efficiency-Equity Trade-off in Kidney Transplantation: A Mechanism Design Model, (Su and Zenios 2006)</td>
<td>MS 52 (11) 1647-1660</td>
<td>No</td>
</tr>
<tr>
<td>Strategic Manipulation of Internet Opinion Forums: Implications for Consumers and Firms, (Dellacros 2006)</td>
<td>MS 52 (10) 1577-1593</td>
<td>Yes</td>
</tr>
<tr>
<td>Project Assignments When Budget Padding Taints Resource Allocation, (Arya and Mittendorf 2006)</td>
<td>MS 52 (9) 1345-1358</td>
<td>No</td>
</tr>
<tr>
<td>The Effect of Payoff Feedback and Information Pooling on Reasoning Errors: Evidence from Experimental Markets, (Budescu and Maciejovsky 2005)</td>
<td>MS 51 (12) 1829-1843</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 1 highlights the neglect in IS of the word ‘ethics’ as a key term. However, this doesn’t necessarily mean that ethical considerations are this scarce. In order to proceed further this review used more specific terms in the study of ethics and referred to the bodies of knowledge in ISE, computer, information and internet ethics in order to be guided on such topics. The ethical topics that appear to generate significant interest are; trust, privacy, piracy and moral hazard. However, whilst the primary focus is based on these four issues, many papers also cover issues such as; IS development, E-Commerce and outsourcing etc. Using these four key terms almost 50 papers appear to be retrieved, this again is slightly misleading as one paper may cover more than one topic (see table 2).

Table 2: The Number of Papers That EBSCO Highlights As Covering The Additional Terms; Trust, Privacy, Piracy and Moral Hazard

<table>
<thead>
<tr>
<th>Journal</th>
<th>Trust</th>
<th>Privacy</th>
<th>Piracy</th>
<th>Moral Hazard</th>
</tr>
</thead>
<tbody>
<tr>
<td>MISQ</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MS</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>JMIS</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ISR</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

After some necessary interpretation of the papers, they were reduced to 32 papers that were deemed to be appropriate to this research (see table 3). This research finds that some human intervention in the use of search engine technologies is necessary for the following reasons. There was some occasional retrieving of inaccurate data e.g. a search for ethics and MISQ also located a paper for the Journal of Computer Information Systems. There was some retrieving of the same paper for a range of topics due to the broad nature of some of the research coverage e.g. covered many themes. Finally, some papers found were not appropriate to the topic e.g. not always IS papers etc. The 32 papers in table 3 are evaluated against the following criteria; investigating if calls for more ISE work have materialised, if so,
what is the nature of the research produced and why and how it is theoretically treated, either support for the use of philosophical ethics (Walsham 1996) or alternative ethical theories (Smith and Hasnas 1999).

**Two Theoretical Approaches in ISE**


Walsham highlights that whilst IS literature sometimes considered ethical issues, there is little in the way of grounding such work within general ethical theory. In addition, he questioned the ‘objectives and value’ of such work. Walsham went on to provide a concise outline of the main theories from normative ethics which he used to re-evaluate the ACM ‘Code of Ethics and Professional Conduct’ of 1992. Whilst Walsham acknowledged the difficulties in creating such codes of ethics, he welcomed such work as helping to guide practitioners and considered much of it to be ‘thoughtful and useful’. However, his research seeks to demonstrate how the use of ethical theory could have produced a more ‘informed and thoughtful reflection and debate’. This is highlighted by examples of some limitations within the Code; unresolved conflicts between duties and goals, problems in what constitutes good (virtuous) IS practice, and little focus in relation to the rights of groups. Walsham concludes by reiterating the value of ethical theory and the role that IS academics can play in making such theories more accessible and explicit in their work. He also highlights that the work of ISE is not restricted to IS but also has important implications for ‘society at large’.


Smith and Hasnas’s also called for more work in ISE and highlighted various problems in such research e.g. the different trajectories of IS and ethical development. They also discuss the problems of such an IS ethical vacuum. However, they seek to focus on the role of IS/IT in business. One of the main claims within the paper is that IS are not an ethically neutral entity, this is supported with an interesting case study of Blockbusters selling customer data to direct marketing companies, described as a well intentioned strategic initiative but conceived in ethical ignorance which resulted in several unforeseen negative responses. In addition, they also call for a more philosophical approach in ISE. However, unlike Walsham, they are somewhat wary of the applicability of philosophical ethics to IS practitioners. Ethical theories are perceived to be ‘too abstract’, not easily accessible to people from a business, engineering, computer science or IS background;

> “Unfortunately, the doctrines of philosophical ethics are highly abstract and are essentially meaningless to one with little or no philosophical training”  
> - Hasnas and Smith (1999), p. 112

They seem to remedy this problem by diverting away from philosophical ethics for a more ‘accessible’ form of ethics e.g. theories of business ethics; stockholder, stakeholder and social contract theory. Although it also acknowledged that such theories are barely used in IS, despite their more accessible / practitioner nature. In summary, despite the recognition of the importance of ISE, we have two fairly similar views on how to proceed, for grounding or considering more research within ethical theory. In the next section, the paper examines the response within the ‘upper echelons’ of contemporary IS research to such previous and high profile calls.


It is beyond the scope of this work to critically evaluate the detailed content of all papers retrieved. However, it is desirable to discuss many of the general themes that emerged. Table
3 outlines the papers studied including the titles (sometimes abbreviated), topics and journal references. In addition, the table outlines the three main evaluation criteria:

- Is ethics or morality acknowledged as a key term (y) or not (n) (column p1)
- Are the terms trust, privacy, piracy and moral hazard acknowledged as a key term (y) or not (n) (column p2)
- Are ethical issues considered by ethical theory (y) or another theoretical concept (n) (column p3)

<table>
<thead>
<tr>
<th>ID</th>
<th>Paper</th>
<th>Topic</th>
<th>Journal</th>
<th>p1</th>
<th>p2</th>
<th>p3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Value of Privacy Assurance. (Hui, Tao et al. 2006)</td>
<td>Privacy</td>
<td>MISQ 31 (1) 19-33</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>Understanding and Mitigating Uncertainty in Online Exchange Relationships: (Pavlou, Huigang et al. 2007)</td>
<td>Trust, Privacy</td>
<td>MISQ 31 (1) 105-136</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>3</td>
<td>E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact, (Bo and Benbasat 2007)</td>
<td>Trust</td>
<td>MISQ 31 (1) 137-209</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>4</td>
<td>The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents, (Komiak and Benbasat 2006)</td>
<td>Trust</td>
<td>MISQ 30 (4) 941-960</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>5</td>
<td>The Impact of Ideology of Effectiveness in Open Source Software Development Teams, (Stewart and Gosain 2006)</td>
<td>Trust</td>
<td>MISQ 30 (2) 291-314</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>6</td>
<td>Ethical Decision Making in Software Piracy, (Moores and Cha-Jan Chang 2006)</td>
<td>Piracy</td>
<td>MISQ 30 (1) 167-180</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>7</td>
<td>The Personalization Privacy Paradox, (Awad and Krishnan 2006)</td>
<td>Privacy</td>
<td>MISQ 30 (1) 13-28</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>9</td>
<td>Perceived Information Quality in Data Exchanges, (Nicolaou and McKnight 2006)</td>
<td>Trust</td>
<td>ISR 17 (4) 332-351</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>10</td>
<td>The Nature and Role of Feedback Text Comments in Online Marketplaces, (Pavlou and Dimoka 2006)</td>
<td>Trust</td>
<td>ISR 17 (4) 392-414</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>11</td>
<td>The Effects of Trust-Assuring Arguments on Consumer Trust in Internet Stores, (Kim and Benbasat 2006)</td>
<td>Trust</td>
<td>ISR 17 (3) 286-300</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>12</td>
<td>How Often Should Reputation Mechanisms Update a Trader’s Reputation Profile, (Dellarocas 2006)</td>
<td>Trust, Moral Hazard</td>
<td>ISR 17 (3) 271-285</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>13</td>
<td>Privacy Protection in Data Mining: A Perturbation Approach for Categorical Data, (Li and Sarkar 2006)</td>
<td>Privacy</td>
<td>ISR 17 (3) 254-270</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>14</td>
<td>An Extended Privacy Calculus Model for E-Commerce Transactions, (Dinev and Hart 2006)</td>
<td>Privacy, Trust</td>
<td>ISR 17 (1) 61-80</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>15</td>
<td>Psychological Contract Violation in Online Marketplaces, (Pavlou and Gefen 2005)</td>
<td>Trust</td>
<td>ISR 16 (4) 372-399</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>17</td>
<td>Maximizing Accuracy of Shared Databases when Concealing Sensitive Patterns, (Menon, Sarkar et al. 2005)</td>
<td>Privacy</td>
<td>ISR 16 (3) 256-270</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>18</td>
<td>Reputation Mechanism Design in Online Trading Environments with Pure Moral Hazard, (Dellarocas 2005)</td>
<td>Trust, Moral Hazard</td>
<td>ISR 16 (2) 209-230</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>
The general findings arising from table 3 in this study are as follows. Whilst there is only one paper that explicitly uses the term ethics (P6), it is clear that many papers (28 out of 32) do explicitly use terms that are popular in ethical research such as; trust, privacy, piracy and moral hazard. It is also clear that whilst it may not always be recognised, that more ethically related work is of interest to IS scholars and is permeating the ‘upper echelon’ journals. Also such research is covering a broad range of diverse topics. However, whilst there may be some significant interest, no papers use ethical theories to consider such emerging IS phenomenon. Subsequently, none will make a philosophical contribution that can challenge any of the established theories used, again somewhat distant from developments in information ethics (Floridi 1999; Floridi and Sanders 2002). Disappointingly, almost all of the authors are moving even beyond Smith and Hasnas’s diversion of philosophical ethics approach (Smith and Hasnas 1999) and by-passing ethical theories completely. This review also unearthed some other findings that also appear to drift beyond previous calls. The majority of the papers prefer to focus on micro-level research, rather than considering many of the broader social issues and there were also some interesting methodological choices and rationale.

In specific terms, in respect of meeting all of the criteria specified in table 3, only one piece of IS research could be said to have succeeded (P6), although there are some theoretical disparities with this work. Paper 18, listed ‘moral hazard’ as a key term, proceeded to highlight a number of potential ethical issues within online reputation mechanisms, but then...
decided to by-pass grounding the work within an established ethical framework. Paper 19, briefly acknowledged the specific ethical and legal implications of its topic, for instance, industry attempts at controlling the illegal downloading of music through peer-to-peer technologies however the work then proceeded to analyse this topic from a market-based perspective.

In terms of the analysis of topics, the issues of trust and privacy are of the most interest. This is perhaps due to the nature of IS developments on the significance of E-Commerce. The issue of piracy is perhaps the most ethically orientated, although the philosophical and empirical ethical issues are relatively neglected. Whilst paper 6, met all of the criteria, there are several other interesting observations, some acknowledged by the authors themselves (Moores and Cha-Jan Chang 2006). Moores and Chang acknowledge some problems with the scope of their research e.g. the testing of an established model within a specific practitioner related context. Thus, the work finds that software piracy is a major economic concern (for those who own such intellectual property), that certain individuals in certain contexts do recognise that pirating is an act of infringing intellectual property and that individuals often don’t judge such acts to be too morally wrong to refrain from such acts. In addition, there are some other interesting observations in relation to the choice of research methods e.g. the use of student sampling within a specific geographical and cultural context. Such methods are somewhat problematic especially in terms of generalisation e.g. society at large. Despite the inevitable problems of representation, student sampling methods were a surprising popular practice in many of other papers (P1, 4, 27, 28 and 30). Some of the problems are somewhat acknowledged in paper 6, but the method is still justified on the basis that students are a ‘prominent source of piracy’? The problems of student sampling are less acknowledged elsewhere (P4, and 27).

Another paper (P19) also specifically acknowledges the ethical dimension arising from the development of digital technologies (peer-to-peer) and the desire by the music industry to prevent illegal sharing. However, whilst this is appreciated, the authors seek alternative approaches to improve our understanding. A similar approach is used in paper 16, where the authors argue that piracy in digitally experienced goods can be positively moderated through the use of realistic marketing and increased sampling and neglect to understand the phenomenon from an ethical perspective. Furthermore, ethics is relegated further and the issues are explored within a marketing strategy solution, a theme that is also developed in other work (P1, 4, 21, 27, and 30).

The research on privacy, moral hazard and trust are also based on micro-level issues and again neglect the use of ethical theory. In terms of privacy, paper 1 and 21 address a professional audience. The papers relating to trust are diversely grounded; e.g. papers 4, 27 and 30 are in reasoned action theory and marketing, paper 10, in psychological contract violation and papers 27, 28, 30 and 32 in trust building theory or marketing. Another interesting piece of research (based on the evaluation criteria for this research) is on trust. Paper 18 (Dellarocas 2005) also uses the term moral hazard as a Key term and the paper sometimes refers to its ethical dimension. The research examines the construction of online reputation mechanisms which deal with seller reputation based on customer feedback. The research is based mostly on the website eBay but also through comparisons to others e.g. Amazon. The paper highlights some broader economic and moral issues of reputation in online environments and explores how effective such reputation mechanisms are designed and managed (or not). The research offers some interesting insights into some of the various flaws within such mechanisms, with some detailed explanations as to why such flaws could be significant. Again like some of the IS research on piracy, this research is developed within a specific context but it attempts to consider several broader social implications.

Finally, this review also highlighted some other generic issues, which may be explained by the popularity of certain authors, some of whom have been very prolific, and their favoured
approaches. In addition to the popularity of student sampling, there was an overwhelming tendency to use quantitative or semi-quantitative methods rather than qualitative methods, including papers 1, 4, 6, 8, 16, 18, 19 and 30, on-line or other survey data was also popular e.g. papers 1, 10, 18, 28 and 32, with noticeable contrasts in terms of the levels of participation and response e.g. papers 1, 10 and 32.

The Use of Ethical Theory in This IS Research

The previous section of this paper largely evaluated the published articles against some fairly interesting but specific criteria. Also the work is evaluated within the authors chosen frameworks. This research now seeks to make a modest contribution by seeking to reflect on how some of the various papers could be with the use of a more explicit ethical theoretical position. Again for the purposes of scope it is impossible to reflect on all of the papers listed, thus what follows is an analysis of the most popular topics covered e.g. piracy and trust. Also the research can only restrict the reflection to a few ethical theories, these being that of deontology (duty-based) and utilitarian (consequentially-based) considerations.

In general, in terms of the issue of piracy (papers 6, 8, 16 and 19), the use of a deontological framework could help to explore why individuals feel that they have a duty not to pirate digital goods. Also do individuals perceive that they have conflicts of interest when conducting such acts and which duties do they subscribe too and why? In terms of utilitarian issues, questions could focus on the individual consequences and if people are concerned about conducting often illegal acts, and if not, why? In terms of the digital music piracy issues covered in paper 19, the authors seem to accept that file sharing is almost inevitable and that many individuals seem to share their files regardless of the multifarious risks involved? Even if such a scenario appeared to be true for many individuals, then such issues may represent such a change in human morality and moral reasoning to question the validity of using such established theories relating to duty and consequence-based ethics, particularly if such ideas are relevant in guiding us in understanding contemporary phenomenon? This methodological position appears to be somewhat similar to the analysis of the technical problems of piracy discussed in paper 16. Thus the approaches prescribed by the authors to the problems of piracy e.g. better marketing and sampling of goods, don’t seem to resolve the root-cause moral issues involved in why people pirate goods and again seems to except such actions as an inevitable feature of contemporary life.

Perhaps the most potentially fruitful paper for a more explicit consideration of ethical theory is Paper 18 on the issue of trust with its focus on reputation mechanisms at e-Bay (and Amazon). Whilst this paper raises several interesting issues and insights into the power of reputation mechanisms in online environments, the work also raises several alternative questions. Some of the neglected questions of interest could be the following. What does such research say about the duties of governance in regulating such new forms of commercial organisations? What accountability issues are raised for online operators in such moral hazard problems? Should online operators devote more resources to predict or manage the moral hazard problems more effectively? Finally, what are the moral, social and legal responsibilities or implications arising from the mass publication of potentially erroneous information within an online reputation mechanism?

Strengths and Weaknesses of this Research Paper

The objective of this research is to make a modest contribution to considerations of the condition of ISE. Some of the weaknesses in this research are as follows. Firstly, the findings are only based on four leading IS journals further research will review other leading journals. Also, although the AIS list was chosen to select the ‘upper echelon’ journals, there are many disagreements about the worth of such ranking systems. Secondly, those engaged in ISE research may actually prefer to target other well respected journals for a variety of reasons e.g. such media may be considered to be more accommodating to such research, although our
research highlights that ethically related topics are now permeating the mainstream, ‘upper echelon’ media. Thirdly, selecting papers that conform to a notion of ISE is also a subjective activity. Others (including the authors of the selected papers themselves) may disagree with the papers identified. Fourthly, because the papers on ISE were purposefully selected, such heightened scrutiny may misleadingly distort the actual strengths or weaknesses of such work. Finally, it is worth remembering that the chosen papers are being judged primarily against a specific set of criteria used for this research and only occasionally against the objectives they sought to develop. It is not the intention to criticise such work but to focus on how a different theoretical treatment may also be considered.

In terms of the strengths of this research, the review does offer a relatively rare contemporary insight into the condition of ISE, particularly in some of the ‘upper echelon’ media. Given the broader significance of IS, it is important to reflect on the condition of ISE research. Although the focus on such journals presents some problems in terms of representation, this specific focus is useful because such work (rightly or wrongly) is likely to be influential to others in IS and other academic fields. Furthermore, the findings from this study challenge those who believe that the appropriate home for ISE research is within specialist media and not necessarily within the ‘upper echelon’ IS media. Such media are increasingly sympathetic to ethical topics, even if this research supports many of the earlier findings (Walsham 1996) (Smith and Hasnas 1999) (Bell and Adam 2004) that authors continue to find an interest in such topics but choose to by-pass interpreting them within a more explicit ethical theoretical framework. Finally, this research also highlights some of the problems associated in conducting such research and some of the problems of an over-reliance on search engine technologies. It also sought to focus beyond reviewing just key term and abstract submissions to also offer some detailed considerations of full paper research submissions. Finally, one of the main strengths of this research is to go beyond highlighting what exists but also try to show what could be? This is demonstrated by the reflections on how such empirical data may be informed (or not) by a reformulation within ethical theory for some of the publications scrutinised.

**Conclusion**

This research has sought to respond to previous calls for a more informed development of ISE research, particularly in relation to considerations of ethical theory in empirical work. Historically, ISE research has been somewhat neglected from mainstream or ‘upper echelon’ media, which is surprising given the support and relative popularity for topics such as: power and politics etc. The historic development of ISE has been influenced by some rare but notable exceptions within IS and by academic work in adjacent fields. Those within IS tend to target their work in either specialist or more sympathetic media.

This contemporary review (2005-2007) of work published in the ‘upper echelon’ IS media highlights that it is something of a myth to believe that such media are inappropriate vehicles for ethically related work. Many ‘leading’ IS journals (and conferences) are increasingly valuing the importance of conducting more ISE related research, perhaps due to the presence of a range of digital technologies that appear to raise an array of interesting ethical challenges. However, whilst such developments are welcome, there appears to be some significant problems in indentifying or appreciating the ethical potential of such research and thus a failure to respond to previous calls to ground such work within ethical theories either in terms of Walsham’s call for more philosophical considerations or Smith and Hasnas’s call to use less ‘abstract’ ethical theories. This research also suggests how the use of such neglected ethical theories may illuminate existing empirical research. Finally, this research is an ongoing and will seek to build further on these initial findings.

**References**


GENDER AND VISIBLE MINORITY STATUS: CAREER ADVANCEMENT IN THE CANADIAN INFORMATION AND COMMUNICATIONS TECHNOLOGY SECTOR

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0204.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Gender, Perceptions, Job satisfaction, Information Worker</td>
</tr>
</tbody>
</table>
GENDER AND VISIBLE MINORITY STATUS: CAREER ADVANCEMENT IN THE CANADIAN INFORMATION AND COMMUNICATIONS TECHNOLOGY SECTOR

Cukier, Wendy, Ryerson University, 350 Victoria Street, Toronto, Ontario, Canada, wcukier@ryerson.ca
Yap, Margaret, Ryerson University, 350 Victoria Street, Toronto, Ontario, Canada, myap@ryerson.ca
Holmes, Mark Robert, Ryerson University, 350 Victoria Street, Toronto, Ontario, Canada, m8holmes@ryerson.ca
Rodrigues, Sara, Ryerson University, 350 Victoria Street, Toronto, Ontario, Canada, srodrigu@ryerson.ca

Abstract

In spite of the economic downturn, demographic factors are expected to continue to produce a labour shortage particularly in the Information and Communications Technology (ICT) sector. In Canada, the sector has signaled that critical skills shortages exist in certain areas although the extent of these shortages is the subject of much debate. While economists have suggested that by the year 2011, all labor force growth in Canada will come from immigration, research has shown that immigrants often face barriers to full employment. Similarly, women are under-represented in ICT and face a range of barriers to employment and advancement. Increasingly, both government and industry have asserted that more effective “management of diversity” could help alleviate the skills shortage. To date, research has focused on barriers to entering the workplace but less has focused career satisfaction even though there is evidence to suggest that these may have significant impact on employee retention and productivity. The role of demographic factors is complex. For example, in Canada, while all immigrants are not visible minorities and all visible minorities are not immigrants; recent immigrants are more likely to be visible minorities. In addition, the intersection of variables, such as gender and visible minority status, complicates the analysis. Our study examines demographic factors affecting career satisfaction among 7110 managers with a minimum of 10 years experience in nine large Canadian companies in the Information Communications and Technology Sector. We found that the gap between visible minorities and white/Caucasian respondents in perceptions of career satisfaction and factors affecting it, to be larger than the gap in perceptions of male and female respondents. More work is needed to address these issues if the ICT Sector is to effectively deal with the skills shortage.

Keywords: Information Systems Professionals, Job Satisfaction, Gender

Proceedings ECIS 2009
1 BACKGROUND

Increasingly government, the private sector and industry organizations are suggesting that effectively managing diversity is, at least in part, the solution to addressing the global labour shortage. Pointing to demographic shifts including the aging baby boomers and declining fertility rates, the Royal Bank of Canada (RBC) has noted that 100% of labour market growth in Canada in 2011 will come from immigration.

From an economic point of view, how well Canada continues to meet the challenges of diversity will determine our future success in attracting talented immigrants as global competition for talent intensifies with the aging of Western societies...in the face of potential labour shortages, employers will miss out on opportunities for growth unless they recognize the potential of all groups in Canadian society (RBC, 2005).

The study also points to the under-employment of women, who are under-represented in certain sectors and occupational groups. Strategies to increase the participation of immigrants, women and other under-represented groups have been advocated to ensure that Canada competes effectively in an increasingly globalized economy.

The problem seems to be exacerbated in the Information and Communications Technology (ICT) sector, where for over a decade industry leaders have been talking about the skills shortage (Boisvert, 2007). There is little doubt that enrolments in computer science and engineering have declined. The Information and Communications Technology Council, for example, estimates that new computer science and engineering graduates which will meet 49%-75% of demand (approximately 7,585 positions) annually (ICTC, 2008). “A new survey shows that a growing shortage of IT skills has put upward pressure on salaries as employers choose experienced workers over graduates....a continuation of what [IT World Canada] calls a “sobering trend” will hurt economic growth and productivity “across all industries” (CBC, 2007). While enrolments of women in other professions — medicine, law and business administration — now exceed males, the percentage of female enrolments in computer science and engineering are declining at a time when their participation in other disciplines is increasing. Female enrolment in computer science is the only discipline where the percentage of women is lower than it was in 1992-93 (ICTC, 2007). Some groups, such as the recently formed, Canadian Coalition for Tomorrow’s ICT Skills, have suggested that increasing the participation of women in engineering and computer science will help address the skills shortage (Ibaraki, 2008). Others, such as the Information Communications Technology Council (ICTC) have emphasized the importance of immigration in filling the gap, with approximately 7,588 immigrants with ICT skills entering Canada annually (ICTC, 2008).

Scholars and industry analysts alike have considered the under-representation of women in the ICT sector from a variety of perspectives. Sue Rossner systematically outlines no fewer than ten different theoretical perspectives to women and technology in the workforce (2006) and stresses the importance of being self-reflective. Liberal feminism, for example, focuses on employment, access and discrimination issues and ways to identify barriers and overcome them. Other approaches, explores ways in which the social shaping of technology has been constructed in ways that exclude women. For example Adam et. al (2006) suggest that Information Technology (IT) skills are gendered requiring women to distance themselves from their own identities. Ramsey & McCorduck (2005) explore issues related to professional identity in the face of systemic stereotyping, dualism, and devaluation. Essentialist approaches assumes all women are united by their biological characteristics and focus, for example, on women’s reproductive role and its implications or on skills and aptitudes considered female. Structural approaches focus on revealing the ways in which patriarchy constrains and focuses on the need for structural and systemic changes. Post-modern feminism questions the very notion of common female experience and posits the notion of multiple identities that are highly context specific. Rossner also describes existentialist, psychoanalytic, racial/ethnic feminism, post-colonial and cyber
feminism. While one might take issue with her categorizations, it is clear that theoretical perspective framing the exploration of this issue shapes the way in which questions are formulated and methods applied. Our study is situated in what is often described as a liberal feminist tradition, based on the assumption that barriers exist to women (and other under-represented groups) which can be identified and addressed to increase participation and advancement in the workplace. We do, however, acknowledge that there are other questions that might be asked.

Currently in Canada, while women are 47% of the workforce, they are only 29% of ICT workers. Females represent only 9% of engineers and 16.7% of programmers compared to 36.8% of analysts and 60.8% of graphic designers and illustrators. (Gunderson et al., 2005). Based on secondary research and interviews with key informants, a study entitled "Diversity – The Competitive Edge: Implications for the ICT Labour Market", (ICTC, 2007) examined barriers to women in the ICT sector. It suggested complex factors affecting participation of women in ICT. Socialization and early education affect the self efficacy and confidence, the development of preferences and choices. Systemic barriers in schools include pedagogical approaches to science and mathematics which are insufficiently applied and the absence of role models, negative perceptions of computing and related work including the “nerd” stereotypes and notions that ICT work is programming. Systemic barriers in post-secondary institutions are similar and have been well documented as creating a “chilly climate” for female engineers and computers scientists (Wasburn and Miller, 2006). In addition, there have been barriers identified to finding employment which include closed recruitment processes and narrow definitions of skills and requirements. Within organizations there are issues related to career advancement, access to training, mentoring, exclusion from informal networks, the absence of role models, stereotypes, communication and negotiation styles and work-life balance issues. Some literature also considers broader socio-political-cultural forces and practices which form the institutional environment of organizations but these issues receive limited attention (For a summary of these see ICTC, 2007).

There is also evidence of under-utilization of immigrants and visible minorities in Canada. While all immigrants are not visible minorities and all visible minorities are not immigrants, recent immigrants are more likely to be visible minorities (Palameta, 2004) and represent a growing proportion of the Canadian population. The data suggest that immigrants tend to be unemployed in spite of their education and technical skill levels. Visible minorities in Toronto face higher unemployment rates than non-visible minorities, although they are more likely to be university educated. There is evidence of under-representation in the ICT sector. Immigrants are 20% of the workforce but only 5% of ICT workers. Visible minorities are 13% of the workforce but only 10% of ICT workers (ICTC, 2007). On average, Visible Minorities experience a 15% wage disadvantage and a 13% earnings disadvantage compared to White/Caucasians (Hum & Simpson, 1999; RBC, 2005). They also face a range of barriers to career advancement (Reitz, 2001).

Similarly, the barriers specific to IEP’s included, “inadequate information about the labour market and processes prior to immigration, settlement and job search challenges, inadequate communications/cultural skills, systemic barriers in employment practices, inadequate support in the workplace”. Based on the barriers relating to lack or inadequate workplace support, this paper attempts to examine the subjective perceptions of employees at work.

To date in Canada, there has been much empirical research on the barriers to immigrants on entering employment but much less on the career progression of visible minorities and immigrants. The evidence suggests that immigrants who are also visible minorities face particular challenges not just in

---

1 In Canada, the term “visible minority” refers to a person who is not an Aboriginal person, who is non-Caucasian in race or who is “non-white” in colour, as defined under the Employment Equity Act. The following population groups comprise the total visible minority in this study: Arab, Black, Chinese, Filipino, Japanese, Korean, Latin American, Middle Eastern, South Asian, South Asian, West Indian, Southeast Asian, West Indian, and multiple visible minority (referring to those respondents who identified as belonging to more than one visible minority group).
terms of obtaining employment but even after they are employed. Similarly, there has been considerable research on women in technology but limited large scale empirical work. As well, we are interested in the inter-relationship of multiple demographic variables, in particular, gender versus visible minority status.

This study draws on previous work on the under-representation of women and immigrants in the ICT sector as well as the extensive body of literature on career success and satisfaction. For example, Cox and Nkomo (1991) reported that women and minorities have lower levels of objective career success than white males. Harter et al. (2002) found that career satisfaction among immigrants and visible minorities is lower and has a significant impact not only on employee performance and retention but also on organization’s productivity and profitability. Greenhaus et. al. (1990) studied 828 manager/supervisor pairs and found that blacks reported having less job discretion and lower feelings of acceptance than whites. They also reported receiving fewer promotions and being less satisfied with their careers and were more likely to report negative organizational experiences than whites. Igbaria and Wormley (1992) who focused on the IT profession found evidence that Blacks receive less career support than whites and tended to have lower levels of met expectations and lower levels of career satisfaction than whites.

Given the focus the concerns in the ICT sector about the talent shortage, a better understanding of the factors affecting career satisfaction and organizational commitment is important to understand how to retain well-qualified, productive employees. A variety of factors affect employee satisfaction (Auster, 2001). Research also suggests that employees who are satisfied with their careers are more engaged and thus be more likely to contribute to achievement of organizations’ success (Harter et.al., 2002).

2 RESEARCH QUESTIONS

Our study explores five inter-related questions including:
1. Are there differences in organizational commitment between visible minorities versus white Caucasian and male versus female employees in Canadian ICT firms? Are there differences in career satisfaction between visible minorities and white Caucasian employees in Canadian ICT firms. Are there differences between males and females?
2. Are there differences in perceptions of relationships with managers and colleagues between visible minority and white Caucasian employees in Canadian ICT firms. Are there differences between males and females?
3. Are there differences in perceptions of career advancement and development between visible minority and white Caucasian employees in Canadian ICT firms. Are there differences between males and females?
4. Are there differences in perceptions of diversity and inclusion practices between visible minority and white Caucasian employees in Canadian ICT firms. Are there differences between males and females?
5. Are there differences in perceptions of educational attainment and credentials between visible minority and white Caucasian employees in Canadian ICT firms. Are there differences between males and females?

3 METHODOLOGY

This paper utilizes survey data collected as part of a larger study that examined career advancement of visible minorities in corporate Canada (Catalyst et. al., 2007). Over 17,000 pre-managers, managers, professionals and executives from 43 organizations responded to an online survey between October 2006 and February 2007 - a response rate of 29%. The survey asked a range of questions pertaining to demographic, human capital, subjective and objective attributes and perceptions. Of these, 6,783 respondents were employed in ICT sector companies and we have analysed these results below. Of the
full-time employees who responded, 54 percent were male and 50 percent possessed some forms of university education. The average tenure with their respective organizations was 11.4 years, at an average age of 41.7 and an average salary of $90,775. There are no discernable differences based on the demographic characteristics between the selected sample and the employees who completed the survey with the exception of salary. About 60 percent of the respondents included in the final sample earned $40,000 or more, compared to only 55 percent of all the respondents to the survey.

4 FINDINGS

All of the survey findings presented below have undergone both t-tests and Chi-square tests, and were found to be significant at p<0.05.

4.1 Organizational Commitment

Most respondents, whether visible minority (87%) or not (90%), indicated that they identify with their organizations core values. Similarly, visible minorities and non-visible minorities both agreed that they were proud to tell others that they are part of their organization (86% and 85% respectively), as well as being willing to put in a great deal of effort beyond what is normally expected to help their organization. As can be seen in Figure 1 below, female (85%) as compared to male (80%) respondents were more inclined to stay with their organization, however visible minority respondents (78%) were less inclined than white/Caucasian respondents (84%) to stay with their organization.

![Figure 1. Respondents perceptual intent to stay with their organization.](image)

4.2 Career Satisfaction

While 72% of male respondents indicated they were satisfied with their overall career progress, 79% of females indicated their satisfaction. In contrast, only 65% of visible minority versus 78% of white/Caucasian respondents were satisfied with their overall career satisfaction. Similarly, 65% of female respondents compared to 57% of male respondents were satisfied with their goals for income, as compared to 64% of white/Caucasian versus 51% of visible minority respondents. The perceptions of achieved goals for development also differed between men (60%) and women (69%) as well as white/Caucasians (67%) and visible minorities (52%). The gap (6%) between female (73%) and male (67%) respondents’ satisfaction with the progress they had made toward meeting their overall goals for the development of new skills was smaller than the gap (9%) observed between the white/Caucasian (72%) and visible minority (63%) respondents. Table 1 below presents the statistics of the key questions pertaining to the satisfaction of females versus males, as well as white/Caucasians versus visible minorities.
### Table 1.

Respondent satisfaction with the progress they have made toward overall career goals, income goals, goals for advancement, and development of new skills.

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>White/Caucasian</th>
<th>Visible Minority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Career Goals</td>
<td>79%</td>
<td>72%</td>
<td>78%</td>
<td>65%</td>
</tr>
<tr>
<td>Goals for Income</td>
<td>65%</td>
<td>57%</td>
<td>64%</td>
<td>51%</td>
</tr>
<tr>
<td>Goals for Advancement</td>
<td>69%</td>
<td>60%</td>
<td>67%</td>
<td>52%</td>
</tr>
<tr>
<td>Development of New Skills</td>
<td>73%</td>
<td>67%</td>
<td>72%</td>
<td>63%</td>
</tr>
</tbody>
</table>

4.3 **Relationship with Managers and Colleagues**

Most respondents indicated that their manager provides them with helpful feedback, although women (73%) were more likely to agree than men (71%), while white/Caucasians (72%) were more likely to agree than visible minorities (70%). There is a significant gap in the perceptions of managerial efforts to learn about their employee’s career goals, between female (78%) compared to men (74%) and white/Caucasians (77%) compared to visible minority (72%) respondents. Similarly, white/Caucasian (89%) respondents compared to visible minority (85%) respondents and female (89%) respondents compared to male (88%) respondents, indicated that they receive the support they need from other co-workers to meet their work objectives. As can be seen in Figure 2 below, the gap (2%) between female (76%) and male (74%) respondent’s indication that their colleagues include them in informal networking is smaller than the gap (11%) between visible minority (66%) and white/Caucasian (77%) respondents. Most white/Caucasian (96%) and visible minority (91%) respondents indicated that their colleagues treat them with respect but, again, white/Caucasians were more likely to indicate this.

4.4 **Career Advancement and Development**

When asked to comment on whether they believed “who you know” was more important than “what you know”, a higher percentage of visible minority respondents (70%) compared to white/Caucasian (60%) respondents agreed, in contrast to 63% of females compared to 62% of male respondents. This reveals a significantly larger gap between white/Caucasian and visible minority respondents (10%), than the gap between men and women respondents (1%). Furthermore, an even greater disparity is revealed between white/Caucasian and visible minority responses to the statement that “people tend to recommend people of their own ethnicity for high visibility assignments”, where just less than 1/3rd of the visible minority respondents agreed with the statement, opposed to no more than 1/10th of the white/Caucasian respondents. A greater number of visible minority respondents (45%) believed that they are “held to higher performance standards than their peers”, compared to white/Caucasian (33%) respondents, while there was no difference between the responses of female (36%) versus male (36%) respondents. As can be seen in Figure 3 below, there is a much larger disparity (17%) between visible

![Figure 2. Respondent perceptions of their colleagues inclusion of them in informal networking.](image-url)

![Figure 3. Career advancement and development.](image-url)
minority (54%) and white/Caucasian (37%) respondents in their perception of how few role models there are for them in their organization, than the gap (3%) between male (40%) versus female (43%) respondents.

![Graph showing respondent perceptions of the amount of role models for them in their organization.]

Figure 3. Respondent perceptions of the amount of role models for them in their organization.

### 4.5 Diversity and Inclusion in the Workplace

Even though there are slight differences in the gaps between male (72%) and female (69%) respondents, and white/Caucasian (72%) versus visible minority (67%) respondents, around two-thirds of all respondents indicated that their organization strives to create a climate supportive of all individuals. Inversely, the gap (11%) between the perceptions of visible minority (18%) and white/Caucasian (7%) respondents that that their organization devotes too little resources to diversity programs, was much larger than the gap (1%) between male (9%) and female (10%) respondents. Visible minority respondents (34%) also agreed to a lesser extent than white/Caucasians (46%) as compared to females (41%) versus males (45%) that their senior management demonstrates a strong commitment to cultural diversity. Of those who indicated whether or not their organization has programs, policies and practices that support the attraction, retention, development, or advancement of visible minorities, 81% of men versus 81% of women (no difference) and 88% of white/Caucasians compared to 59% of visible minorities (29% different) indicated that their organizations had programs, policies and/or practices. Table 2 below presents the statistics of the key questions pertaining to diversity and inclusion in the workplace.

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>White/ Caucasian</th>
<th>Visible Minority</th>
</tr>
</thead>
<tbody>
<tr>
<td>My organization strives to create a climate...</td>
<td>69%</td>
<td>72%</td>
<td>72%</td>
<td>67%</td>
</tr>
<tr>
<td>My organization devotes too little resources to diversity programs.</td>
<td>10%</td>
<td>9%</td>
<td>7%</td>
<td>18%</td>
</tr>
<tr>
<td>My senior management demonstrates a strong...</td>
<td>41%</td>
<td>45%</td>
<td>46%</td>
<td>34%</td>
</tr>
</tbody>
</table>

Table 2. Survey participant responses to diversity and inclusion in the workplace questions.

### 4.6 Attainment and Recognition of Education Credentials

More than half of the visible minority respondents (63%) had a bachelors degree or higher, while less than half of the white/Caucasian respondents (45%) had a bachelors degree or higher. Furthermore, men (53%) are more likely to have bachelor degrees or higher than women (44%). This uncovers a greater education gap between visible minorities and white/Caucasians (18%) than that realised...
between male and female respondents (9%). Table 3 below outlines the breakdown of the levels of education attained by each of the respondent categories under analysis.

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Female</th>
<th>Male</th>
<th>White/Caucasian</th>
<th>Visible Minority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than high school</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>High school</td>
<td>21%</td>
<td>11%</td>
<td>18%</td>
<td>8%</td>
</tr>
<tr>
<td>College certificate/diploma</td>
<td>29%</td>
<td>30%</td>
<td>32%</td>
<td>23%</td>
</tr>
<tr>
<td>Bachelor's degree</td>
<td>35%</td>
<td>39%</td>
<td>35%</td>
<td>43%</td>
</tr>
<tr>
<td>Masters degree</td>
<td>9%</td>
<td>14%</td>
<td>10%</td>
<td>19%</td>
</tr>
<tr>
<td>PhD</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>1%</td>
</tr>
<tr>
<td>Other professional degree/designation</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>6%</td>
</tr>
</tbody>
</table>

Table 3. Educational attainment of survey respondents.

Of the education earned, only 6% of female and 9% of male respondents earned at least some of their education outside of the United States, Canada, Australia, New Zealand, France, Germany or the United Kingdom, compared to only 2% of white/Caucasian versus 23% of visible minority respondents. Also, 44% of visible minorities compared to 25% of white/Caucasians, and 43% of males compared to 33% of females, believe that their education credentials are recognized as being less than their Canadian equivalent.

Recognition of foreign credentials is one of the biggest barriers identified by respondents in the study, as 31% of respondents felt that their education was being underutilized by their organizations. While the results of the survey show that men to women and white/Caucasian to visible minority responses differ the biggest variance is found in comparing visible minorities to men and women. As this is a survey of professionals who have been in the sector for an average of 10 years, survivor bias is an issue. That is, the respondents are women and visible minorities who have persevered in spite of the barriers that they face.

4.7 Overall Findings

While the survey assessed perceptions, it does suggest significant gaps between the satisfaction and perceptions of fairness in the workplace between visible minority and white Caucasians. Specifically, we found:

1. Visible minorities show less organizational commitment than white Caucasian employees in Canadian ICT firms. There is also a gap between male and female respondents but it is smaller and, perhaps unexpectedly, females are more committed than males.

2. Visible minorities show less career satisfaction than white Caucasian employees in Canadian ICT firms. There is also a gap between male and female respondents but it is smaller and women are more satisfied than men.

3. Visible minorities show less favourable perceptions of their relationship with managers and colleagues than white Caucasian employees in Canadian ICT firms. There is also a gap between male and female respondents but it is smaller. Once again females show more favourable perceptions than males.

4. Visible minorities show less favourable perceptions of career advancement and development than white Caucasian employees in Canadian ICT firms believing that its “who you know, not what you know”. There is also a gap between male and female respondents but it is smaller with females more likely to subscribe to this view.

5. Visible minorities show less favourable perceptions of diversity and inclusion practices than white Caucasian employees in Canadian ICT firms. There is also a gap between male and female respondents but it is smaller.
6. Visible minorities perceive that their educational attainment is less valued than white Caucasian employees in Canadian ICT firms. There is also a gap between male and female respondents but it is smaller with females likely to be more satisfied. The fact is that the gaps between visible minority and white/Caucasian satisfaction, is substantial. What was perhaps most unexpected was that in general female respondents were more likely than males to indicate that they were satisfied. This may suggest that differences in expectations or that the barriers to visible minorities are greater than those faced by women. Of particular note is that the expressed commitment to diversity in many of these organizations, particularly at the senior level, contrasts with the perceptions of many visible minority respondents. Not surprisingly, white Caucasian respondents appear to be less aware of these issues.

4.8 Initiatives to Promote Diversity in the Canadian Workforce

While there is a gap between white/Caucasian and visible minority respondents, employers reported a range of initiatives aimed at supporting full employment of immigrants and visible minorities. As noted above, there are also a plethora of initiatives aimed at attracting and retaining under-represented groups in the industry (ICTC, 2007). Companies are also undertaking targeted outreach to immigrants and visible minorities. For example, Nortel Networks has been a sponsoring company of INROADS, a non-profit organization that aids in the training and development of young people of colour for professional careers in business and industry. It offers multi-year summer internships, year-round coaching, and training activities. In addition to sponsorship fees, corporate sponsors are expected to provide a professional internship opportunity each summer for an intern and to consider hiring the intern for full-time employment upon graduation. Toronto Region Immigrant Employment Council (TRIEC) is a multi-stakeholder council which offers services for skilled immigrants matching them with established professionals in the same field. While the control group (without a mentor) got jobs more quickly, their searches were less focused and the employment they secured was in general labour, assembly, security, data entry and telemarketing earning $8-22 per hour. In contrast, the mentored group spent more time researching companies and landed more interviews. They earned an average of $55,000. More than half (55%) of those in a mentoring relationship, found jobs in their field of training through TRIEC’s Career Bridge program (TRIEC, 2006).

Once they are recruited, companies have a range of programs aimed at supporting under-represented groups in the workplace. For example, Nortel includes diversity and employment equity compliance training as part of all its management training. Gennum Corporation, a Burlington, Ontario-based medium-sized company in the semiconductor market, reports that approximately 20% of employees have emigrated from other countries. The company offers mentoring programs, language and communications skills programs on site. Celestica has adopted the Electronic Industry Code of Conduct, a voluntary code that commits participants to ensuring that they and their first-tier suppliers do not engage in “discrimination based on race, colour, age, gender, sexual orientation, ethnicity, disability, pregnancy, religion, political affiliation, union membership or marital status.”

Finally, and most challenging, are initiatives aimed at shaping societal views. Hireanimmigrant.com for example, is a campaign aimed at challenging assumptions and stereotypes through a provocative series of television advertisements. One for example, features two south Asian men discussing an applicant for a job. “But his degree is from Canada”, “His only experience is Canadian”, “And then there is his accent”, “Why are we even looking at him?” It concludes. The advertisement attempts to confront many of the stereotypes and discriminatory views of Canadians. However, the impact of such social marketing efforts is often uneven. While such campaigns may or may not be effective, there is little doubt that systematic attention to the representations of under-represented groups in the media is critical. Large companies can use their spheres of influence to help shape the media representations of under-represented groups and to use their purchasing power.

Barriers to full participation for under-represented groups can be overcome, and that effective programs share certain characteristics and approaches (Matton & Hernandez, 2004). Specifically,
successful workplace diversity initiatives hinge on committed leadership, defined goals/targets or measures of effectiveness, strong diversity professionals, employee involvement and ties to performance evaluation, as well as data to identify, quantify and communicate progress and challenges. At the same time, organizations operate within a social context. There are broad social and cultural factors, including stereotypes, which shape perceptions and expectations, and which must be addressed. Better data and more research on current practices will help advance our understanding of processes for developing strategies which “fit” in particular contexts. “Best practices” which are innovative, make a difference, have a sustainable effect, and can be replicated and applied in other contexts need to be further explored.

4.9 Implications: Towards an Integrated Approach

Discussions of diversity have proposed conceptual models of change. For example, the notion of the “diversity continuum” (Friday and Friday, 2003, Gilbert, et. al., 1999) is based on the assumption that organizations move through a stages on the path to becoming inclusive. Others have proposed that different sectors and companies within those sectors are at different stages along a “diversity curve” (Cukier et. al., 2007). For example, large telecommunications services tend to be higher up the continuum because they are federally regulated, female dominated, consumer-market oriented, knowledge intensive and relatively profitable. Given the complex and inter-related existence of barriers, we suggest that an ‘ecological’ model of change which considers the complex interactions among individual factors, group factors, organizational factors and societal factors which shape choices, create barriers and facilitate opportunities. The model is adapted from public health (McLeroy et. al., 1988) and proposes complex, iterative, interactions between individuals, groups, organizations, institutions and society. The core belief is that no element exists within a vacuum but is influenced and influences other actors. We propose this model in order to address one of the major gaps which has been identified in the literature to date on diversity – that is the relative lack of research exploring the institutional environment of organizations or the larger socio-political-cultural context of organizations. Efforts must address factors at the individual, organizational and societal levels.

At the individual level, focused programs aimed at equipping members of under-represented groups – with the knowledge, skills and attitudes that will increase their chances of success are essential. Within organizations critical issues include:
- A well defined business case and explicit policy on diversity/employment equity
- Senior executive support;
- Targeted recruitment/retention/promotion programs;
- Open and transparent recruitment and promotion policies
- Employee training and communication;
- Diversity performance objectives and targets ideally tied to compensation;
- Maternity/parental benefits and measures to support work-life balance
- Mentoring, networking and support opportunities;
- Implement supportive work practices and culture (eg. work life balance, mentoring, networking)
- Alignment of management of suppliers/procurement policies
- Systems to track diversity data;
- An accountability framework, eg. Public reporting on diversity issues (Jantzi Research, 2007; Catalyst and the Diversity Institute, 2007; ICTC, 2007)

At the societal level, better policies and programs are needed to support immigrants and other under-represented groups. But more importantly perhaps, vigilance is needed to address the stereotypes and hidden barriers that limit both aspirations and opportunities. Our study suggests that while there is strong evidence to suggest that immigrants and other under-represented groups could help address the ICT skills gap, more work is needed to fully utilize their skills.
References


Information and Communications Technology Council (ICTC) (2008). Outlook for the Information and Communications Technology Labour Market, Ottawa, ICTC.


MEASURING THE SUCCESS OF INTERVENTION PROGRAMMES DESIGNED TO INCREASE THE PARTICIPATION RATE BY WOMEN IN COMPUTING

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0328.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Gender, Evaluation, Case Study, Education</td>
</tr>
</tbody>
</table>
MEASURING THE SUCCESS OF INTERVENTION PROGRAMMES DESIGNED TO INCREASE THE PARTICIPATION RATE BY WOMEN IN COMPUTING

Craig, Annemieke, Deakin University, annemieke.craig@deakin.edu.au, Piddons Road, Geelong, Victoria 3217, Australia

Dawson, Linda, Monash University, linda.dawson@infotech.monash.edu.au, Monash University PO Box 197, Caulfield East 3145, Australia

Fisher, Julie, Monash University, Julie.Fisher@infotech.monash.edu.au, Monash University PO Box 197, Caulfield East 3145, Australia

Abstract

Many intervention programmes to encourage greater female participation in computer education and careers have been conducted in the last twenty years. These intervention programmes take considerable time, effort and money to design and implement. If success were to be measured by an increase in the percentage of female students undertaking computing courses then these programmes would have to be considered a failure. This paper describes a research project which examined fourteen intervention programmes in detail. From the perspective of the programme champions each of the intervention programmes was considered successful, even when this success was restricted to specific areas or limited to small groups of individuals. Formal evaluation appeared to have been an afterthought rather than a priority of many of the programme champions. Some programmes appeared to be less effective due to the lack of targeted and clear goals or predetermined evaluation criteria. It is recommended that during the initial planning phase for intervention programmes a clear objective is to consider what a successful programme would look like and what the evaluation criteria would be. Further work is needed to understand how intervention programmes can be better designed and evaluated so that their impact and success can be expanded.

Keywords: Gender, Computing, Recruitment, Intervention programmes, evaluation.

1 INTRODUCTION

Gender imbalance in specific discipline areas, or occupations, becomes a problem when the gender spread has a detrimental effect on some sections of society. For both men and women there are areas of the workforce where either dominates and one of these areas is computing. The early history and development of computing included significant contributions from women but as the discipline has evolved the number of women entering the profession has declined. While the computing profession is still relatively young it employs over a quarter of a million people in Australia yet less than 20% of these workers are women.

The philosophical viewpoint taken in this paper is that the field of computing has been socially constructed as a male domain and therefore while women are capable of taking their place in this field, at present, they are choosing not to. Lazowska (2002, p. 11) suggests that of the many reasons why we should be working to increase the number of women in computing ‘the selfish reason is the most compelling one: the quality of the solutions we achieve is enhanced by the diversity of the individuals contributing to these solutions’. This paper presents a research project which examined fourteen intervention programmes and reports on the achievements of those programmes. The criteria used to
measure success and the characteristics of those programmes that were regarded as successful are presented and discussed.

2 BACKGROUND

If women are not part of the technology solution design then it is unlikely that technology will represent their needs or their family’s needs. As Borg (2001) reflected technology for the future will depend on the designers, builders and controllers as well as those who use it.

... involving women in technology - how it is taught, learned, created and used - benefits everyone. .... It’s about bringing the talents of technical women leaders to bear on developing more competitive products and finding solutions to problems that impact our lives, our nation and our world. (Anita Borg Institute for Women and Technology, n.d.)

Woszcynski, Beise, Myers and Moody (2003, p. 1585) express the concern that without sufficient women in the computing workforce development of ‘technology pursuits may focus more on doing things faster, and less on doing new things that reflect alternative perspectives’. This was also the focus behind the creation of the Virtual Development Center in 1999, a programme of the Anita Borg Institute for Women and Technology. The authors highlight the importance of technology for improving the human condition generally and then pose a number of critical questions:

But what are the problems technology is attempting to solve? Whose priorities are represented? How much of technology truly benefits the world’s peoples? Who are the creators of technology? The creators of most of our current technology, however, represent a narrow stratum of the world’s population – North American males. (Anita Borg Institute for Women and Technology, n.d.)

Wulf (1998) explains that with a gender imbalance in the engineering workforce ‘we limit the set of life experiences that are applied, and as a result, we pay an opportunity cost, a cost in products not built, in designs not considered, in constraints not understood, in processes not invented’. This is equally true of computing and consequently computing is another area where gender imbalance is of concern and needs to be addressed. To ensure that technology is useful for all, a diverse blend of people is needed to design, develop and implement new systems (Camp 2001). Camp (2001) argues that to create better technology more women are needed to be incorporated in development teams as they bring different perspectives, different points of view and different ways of solving problems thus broadening the systems development process. Without women in the computing workforce the different perspective, priorities and operating styles women bring to the design and development of systems is lost (WSET 1995) resulting in systems that may not be useful to the whole population.

The impact of the absence of women in design and development can be best illustrated by examples:

- Margolis and Fisher (2002) report on a voice-recognition system which was calibrated to male voices only and therefore not able to hear women’s voices. A video conference system built around this software, where the camera automatically focused on the person speaking, could not hear the women and therefore the women were also not seen.

- Tests conducted on the first airbags showed that they saved lives. However, once in use, women and children were injured and even killed by inflating airbags. A study reported in the American Journal of Public Health described how the induced injuries from airbags were ‘disproportionately borne by female drivers’ (Segui-Gomez 2000, p. 1575). The original design team had been made up almost exclusively of men (Coonan 2006) and the initial tests had all been conducted with a 50-percentile male crash-test dummy, which was the size of the average American male (Nikkel n.d.). If there had been more women involved in its development, the fatal flaws in the airbag may have been avoided.

There are also examples from science and engineering which ‘show how a product–design group that is not representative of its users can go wrong’ (Margolis and Fisher 2002, p. 3):
• Artificial heart valves were sized exclusively to the male heart by a mostly male design team (Margolis and Fisher 2002).
• The theories about the preventative benefits of taking one aspirin a day were tested on 22,000 men (Kolata 1990) but reported in such a way that they applied to both men and women (Pear 2000).

Yet men and women can respond differently to the same drug (Buring 2000).

Although these examples highlight mainly physical and physiological differences, the creation of technology and computing environments that benefit all of society requires diverse perspectives in developmental and organisational areas also. Brown (2007, p. 334) concludes it ‘is not that men are incapable of representing women’s interests, but that men generally lack the necessary experiences to consistently and competently raise issues that tend to concern women more than men’. Unless more women are employed in computer design and development, these products and services are unlikely to meet the needs and desires of approximately half the population. The current loss of available knowledge from women in IT development can only be considered to be a loss for society as a whole.

Women need to be actively involved in all levels of these new technologies that have such immense potential for social change.

3 THE COMPUTING EDUCATION PIPELINE

Valenduc et al. (2004, p.10) argue that ‘Professional ICT skills are required in the ICT industry and in related jobs in the user industries, in order to create, develop, implement, repair or manage ICT tools (hardware, networks and software)’. Consequently the acquisition of appropriate computing skills and qualification via education is necessary.

Right along the education pipeline, from primary school to secondary school to tertiary institutions there are fewer girls choosing to study computing. In Australia, as in other countries, this issue was first identified as of concern in the early 1980s. The Australian government set a target that, by 1995, 40% of students within tertiary computing courses would be female (DEET 1990). This prompted a range of strategies to encourage the uptake of computing studies by female students. In Australia, as elsewhere where this same imbalance is an issue, numerous intervention programmes aimed at encouraging more female students to consider computing and successfully complete courses have been conducted. Initiatives have ranged from mentor programmes, the production of videos, special classes for female students, curriculum changes to create a more inclusive curriculum, to the running of workshops and computer camps (Greenhill et al. 1997b; Craig et al. 1998; Clayton and Lynch 2002). Intervention programmes have also been created to equip students with the necessary skills needed to work in this field. Other programmes have aimed to provide the necessary contacts to obtain jobs; while yet others have focussed on how to obtain career recognition and for support to stay in the industry.

Statistics however show no improvement in the rate of participation of females in this industry and anecdotal evidence suggests that these intervention programmes have not been as successful or effective as was anticipated. If success is measured, albeit simplistically, by the current percentage of females in tertiary computing courses, then the programmes must be seen as a failure with a decline from 27.2% in the 1990s (Lang 2003) to less than 20% of current tertiary computing students being female (DEST 2008). This view however is too simplistic, as there is no way of knowing what the percentage of females in computing would have been without the intervention programmes. The government’s target level of 40% has, however, certainly never been reached.

This raises questions about the efficacy of the intervention programmes, given the considerable efforts of many to raise the level of involvement of females in computing. Were any of the initiatives successful? How is success defined? How can success be measured?
4 THE RESEARCH METHOD

The aim of this research was to investigate gender-equity intervention programmes, focusing on the enrolment and retention of female students in computing courses in Australia, over the last twenty years. A collective case study of 14 individual intervention programmes was undertaken. Each of these cases was a concentrated inquiry into a particular intervention programme: Each case was investigated individually to try to understand its complexities (Stake 2000).

Research questions included:
- How did, or do, these programmes operate?
- To what extent have these programmes been successful from the initiators perspective?
- To what extent have these programmes been evaluated by those initiating the programmes?
- What criteria were used to measure success?

Data was collected via detailed document and artefact analysis and by in-depth interviews with the instigator/leader of each of the programmes. The analysis of each individual case was followed by a cross-case comparison.

Within the context of qualitative case study research in the computing discipline, the number of case studies (14) is consistent with other research (Orlikowski and Bardoudi 1991). Myers (2002) queries why more than one case is actually necessary but Miles and Huberman (1994) suggest cross-case analysis enhances generalisability as a multiple-case design will deepen the understanding and ability to explain what has occurred. Herriott and Firestone (1983 as cited in Yin 1994, p. 45) suggest that the evidence from multiple cases is more compelling and therefore the study will be more robust than one of single-case design.

Miles and Huberman (1994) suggest that a multiple-case study requires clear choices about which cases to include within the study. The intervention programmes which became the case studies for this research were selected on the following basis:
- One of the programme’s objectives was to increase the number of females who were part of the computing field.
- The programme could be made up of one or more projects however the programme needed to be a sustained activity.
- The principal champion/instigator of the programme was prepared to participate in this research.
- The programme and projects could be completed or be ongoing.
- The programmes were chosen to provide diversity in location and focus.
- How long the programme ran for, it was assumed longer term programmes would be more successful.

Intervention programmes have been conducted by three different types of entities in Australia; educational institutions, government bodies and industry groups. Eight cases were selected from universities, three from government bodies and three from industry groups. A greater number of case studies were chosen from the university sector than the other sectors due to the proliferation of such activities in this sector.

For each case study the programme leader or a major contributor was interviewed. However additional interviews were conducted for five of the case studies either because there were joint programme leaders or another major contributor was available and willing to be interviewed. A total of 19 interviews were conducted. For four of the case studies two people were interviewed and in one instance a third person was interviewed. Interview times ranged from approximately 60 minutes to 100 minutes in duration. For privacy reasons the interview participants are identified by pseudonyms and will be referred to as the programme’s champion in further discussion. Additional data for each case study came from the detailed study of documents and artefacts (consisting in total of 40 published and 10 unpublished papers, 17 reports, 32 surveys, 6 videos and 12 websites). These
documents/artefacts were investigated for detailed descriptions of the intervention programmes in context, as well as evidence of evaluation or the absence of evaluation.

Easterby-Smith, Thorpe and Lowe (1995) suggest that analysis of in-depth interviews initially requires the researcher to become familiar with the material and then to reflect upon it. Familiarisation of the data was strengthened via thorough re-reading of the transcripts. A set of themes and sub-themes important for understanding what was going on then began to develop. The interview transcripts were subsequently incorporated into NVivo 7 software. While such software can support the analysis of qualitative data, by searching, linking, marking and so on, it can not undertake the analysis and decide what all the text means. As Weitzman and Miles (1995, p. 330) explain; computer ‘software will never ‘do’ theory building for you…. but it can explicitly support your intellectual efforts, making it easier for you to think coherently about the meaning of your data’.

The documents and other artefacts collected also underwent the stages of familiarisation and reflection. Videos were watched, websites were explored, and support materials examined. In NVivo a ‘notes section’ was created, one for each of the case studies, and relevant facts and reflections were recorded as the researcher investigated the artefacts thereby creating a chain of evidence. Next other documents (such as evaluation reports, grant applications and research publications) were carefully re-read and entered into NVivo. This process brought together all the data for analysis, regardless of its source, into one NVivo project file enabling sorting, searching and linking. Having all the data in one project file within NVivo also enabled the creation of a meta-matrix as described by Miles and Huberman (1994) to facilitate an analysis for patterns in responses and opinions. Yin (1994) agrees that this approach is a good method for analysing multiple-case study data.

5 THE RESULTS

5.1 The Programmes

Of the fourteen case studies explored eight studies centred on programmes developed in academia through universities, three were state government initiatives, and three were developed by industry bodies (see Table 1). The geographic locations of these organisations span the entire breadth of Australia. At the conclusion of the research, six of the programmes were still highly active but in eight activities have stopped, or were operating only at a minimal level.

<table>
<thead>
<tr>
<th>Organisation</th>
<th>Number of Cases</th>
<th>Limited / No Activity</th>
<th>Highly Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic</td>
<td>8</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Government</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Industry</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 1: Case Study Activity

While the United Nations Development Report (2004, p. xii) advocates that the ‘critical starting point for achieving gender balance in the ICT sector is tertiary level education’, work by Clayton, Cranston, Crook, Egea, Lynch, Orchard, Robinson and Turner (1993) has provided a broader framework which identifies three stages where it is possible to influence the participation by females in computing courses and the computing profession. The first stage of the framework is the Pre-tertiary stage where the focus is on intervention programmes for primary and secondary girls. The Tertiary stage has a focus on decreasing the attrition rates for tertiary female students. The final stage of Post-tertiary is aimed at women in the workforce or returning to the workforce.

University and industry involvement commenced with a focus on only one of Clayton et al.’s three stages, but with a growing awareness of the magnitude of the issues the programmes evolved to incorporate interventions at more stages. Eight of the case studies conducted projects at all three
stages, five conducted intervention projects at two stages, with only one University’s focus remaining on just one stage (see Table 2). It is interesting to note that all groups conducted activities at the pre-tertiary stage. Some programmes concentrated their efforts equally across all stages whereas some focused their activities more at a particular stage. The stage at which the majority of the activities were focused is regarded as the major focus of the programme. A minor focus refers to less activities being directed at that target audience.

<table>
<thead>
<tr>
<th>Pre</th>
<th>Uni1</th>
<th>Uni2</th>
<th>Uni3</th>
<th>Uni4</th>
<th>Uni5</th>
<th>Uni6</th>
<th>Uni7</th>
<th>Uni8</th>
<th>Gov1</th>
<th>Gov2</th>
<th>Gov3</th>
<th>Ind1</th>
<th>Ind2</th>
<th>Ind3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tertiary</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Post</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 2: Focus of Intervention Programmes

![Table 2: Focus of Intervention Programmes](image)

5.2 The Programme Achievements

Some of the intervention programmes were still relatively new (3 years old), and some had been operating for a lengthy period of time (up to 20 years). Examples of the extent and variety of the achievements by these 14 case study entities are described in the following section. An indication of the number of participants at various events is also provided.

In total more than 40 separate Girls in Computing Days or Role Model events were conducted with 14,000 girls, 500 educators, 200 parents and 200 role models participating in these events. More than 200 girls from nine locations throughout one state participated in one multimedia workshops. At least four separate computer camps were conducted for secondary girls; and more than 150 computer clubs now run in schools including one which is being conducted wholly online.

Outcomes of many intervention programmes included the creation of resources such as; a ‘How to’ kit for Girls and ICTs role models events; A ‘Practical Ideas’ booklet for Teachers regarding Girls and ICTs which was distributed to all schools within one state; and ‘Mentoring as a Career Guidance Activity’ resource kit. Other resources produced included six videos, 12 separate websites including one which contained more than 70 pages and had received more then 34,700 hits over a two year period.

For industry women there were five national women in computing conferences conducted as well as two state-based women in computing workshops. More than 20 professional development (PD) sessions were delivered to more than 400 teachers.

Scholarship/awards/bursary schemes now exist where they did not before.

Media exposure and publicity has been generated with ‘Girls and ICT’ articles appearing in numerous school and parent magazines, teacher journals as well as in local and national newspapers. Further, an awareness-raising campaign directly reached 23,044 students in 370 schools across one state, and indirectly reached thousands more.

There were additional projects and achievements, which are too numerous to list, that were more local and specific. Some examples include four girls taking up or intending to take up ICT careers as a result of participating in one particular event; all eleven Year 10 students at one school who participated in intervention events indicated that they intended to take at least one science and/or IT subject in Year 11; 30 students and seven teachers from six secondary schools attended a breakfast with ten female ICT mentors from a local Council.

There are also achievements that are harder to quantify. For example in a number of schools it was noted that there appeared to be an increase in ICT skills and confidence by girls in using ICTs amongst participating girls. Girls were also more willing to take a leadership role in the use of ICTs. In some
schools there was an increase in understanding of the many possibilities and options in ICT study paths and careers. Students also felt they were 'cared about', a sense of 'belonging' assisting with their retention in courses. In universities an improved curriculum and/or an improved learning environment was developed and numerous supportive communities were created enabling networking and mentoring programmes to be conducted. For industry women better networks between entities and therefore better distribution of key information occurred. Workplace strategies have been shared and professional development activities carried out.

As can be seen from these examples, these intervention programmes resulted in considerable achievements with many activities conducted and large numbers of participants involved. Can these programmes therefore be considered successful?

5.3 Criteria used to measure success

Credible evidence, be it qualitative or quantitative or both, should be gathered to evaluate the success of a programme. While formal evaluation was not undertaken of many of the activities within the programmes the evidence upon which the champions evaluated the programmes is summarised in Table 3. Similar activities have been grouped together and duplicate criteria have been removed.

<table>
<thead>
<tr>
<th>Award Programme / Scholarships / Bursaries / Competitions / Grant Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>▪ Number of entries/applicants received and geographical spread of the entries/applications</td>
</tr>
<tr>
<td>▪ Percentage of awards going to women and growth of scholarship/award programme</td>
</tr>
<tr>
<td>▪ Willingness of organisations to sponsor awards</td>
</tr>
<tr>
<td>▪ Feedback (written and verbal) from recipients; Coverage received by the media</td>
</tr>
<tr>
<td>▪ Follow up of outcomes for winners of each category and status reports/feedback from grant recipients</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bridging Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>▪ Number of applicants; the retention rate of participants; the progression rates of participants</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Creating Teacher and Parental Awareness / Professional Development of Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>▪ Number of articles published in literature aimed at teachers/parents and the circulation of materials amongst teachers at the secondary school</td>
</tr>
<tr>
<td>▪ Number of PD sessions organised; number of teachers attending; interest shown during the sessions</td>
</tr>
<tr>
<td>▪ Number new initiatives being conducted within schools by prevision attendees of PD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Computer Club</th>
</tr>
</thead>
<tbody>
<tr>
<td>▪ Change in participant’s attitude, skills, confidence awareness and enthusiasm through ongoing participation</td>
</tr>
<tr>
<td>▪ Engagement of participants; teacher/facilitators observations.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Girls in Computing Days / Role Model Events/ Profiling Successful Women / Speaking at Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>▪ Numbers of participants (students, educators, parents, role models) at the events and number of registrations well before the event; Student behaviour and concentration during sessions; enjoyment and engagement</td>
</tr>
<tr>
<td>▪ Number of ‘good’ questions asked by students and feedback (written and verbal) from students, educators, parents, facilitators, role models and changes in attitude and awareness amongst participants.</td>
</tr>
<tr>
<td>▪ Quality of the role models; innovative activities, high energy, high impact and the how well secondary students relate to the visiting students/speakers/role models</td>
</tr>
<tr>
<td>▪ How well known ‘the brand’ is within the education sector</td>
</tr>
<tr>
<td>▪ Generating a sustainable model that the event continues to be conducted annually. Illustrated also through enquiries from schools about the next event</td>
</tr>
<tr>
<td>▪ Number of interviews and publicity generated</td>
</tr>
<tr>
<td>▪ Financial support (sponsorship) from industry</td>
</tr>
<tr>
<td>▪ Creating relationships between groups: industry, education and government and forming networks</td>
</tr>
<tr>
<td>▪ Students going on to enrol in specialist ICT subjects in later years of schooling and going on to apply for scholarships</td>
</tr>
<tr>
<td>▪ Inspiring girls to set up their own computer club</td>
</tr>
</tbody>
</table>
### Improving the Curriculum
- Gender analysis of students’ results in relevant subjects; retention rate of girls

### Mentoring / Support Community / Networking
- Number of participants/members in the community; membership growth and participation numbers at events conducted; the level of engagement of participants
- Level of activity by the community: Number of activities; scope and range of activities.
- Anecdotal feedback and informal comments from participants
- The amount of time girls used the extra hardware provided. The number of participants who sought references from participating staff members or other support and who kept in touch after graduation
- Programme/group becomes embedded and participant/student driven
- Visibility of the group
- Number of students/participants who ‘opt out’
- Level of activity on discussion board
- Amount of sponsorship received and organisations taking part
- A champion emerges from the group

### Overall Pre-tertiary Initiatives
- Increase in the percentage of girls amongst enrolling tertiary computing students.

### Public Speaking Programme
- Students encouraging others to participate and perceived change in students’ confidence levels

### Residential Summer School
- Change in awareness of the tasks required by a ‘typical computing professional’ and change in interest in studying ICT and interest in a career in ICT. This leading to students choosing to enrol in tertiary ICT courses; the number who went to Uni X.
- Anecdotal feedback from students, facilitators and staff, during the camp and afterwards.
- Students returning to their school and speaking about the camp
- Peer evaluation from team members and other colleagues
- Extensive coverage in the media

### Web sites/Video/Resources
- Number of web site hits; Growth of the website
- Another department/organisation using the video/resources. Circulation of materials amongst teachers at the secondary school
- Others recommending the resources.

### Women-only workshop / Conferences on Women in Computing
- Answers to ‘What did you like about today? What changes would you suggest for the next event?’
- Number of participants; Engagement of participants during the day
- Whether all sectors (education, industry, government) were represented

### Tertiary Orientation Camp
- Student comments during and after the camp; Interaction of students when back at University
- Participants reported gaining a sense of ‘belonging’; retention rates in the courses.

### Other
- Number of referred papers
- Change in the culture of the organisation including a commitment by the organisation demonstrated by a budget line entry and support within the organisation
- Uptake of recommendations from gender audit
- Becoming known as the university of choice for girls wanting to do a computer course
- Equity awards received

---

**Table 3: Evaluation Criteria**
5.4 Defining a Successful Programme

When is an intervention programme a success? How can we measure this? For the staff involved with the intervention programmes at Uni5 success for their intervention programme was defined as passing the subject Computing 101 for the girls who participated in the activities (Uni5_documents). Megan (Uni1) on the other hand suggested that success would need to be broader than one subject and would look like ‘maintenance of numbers in my courses’. For Lesley (Ind2) success required the girls to actively follow up the opportunities that they were made aware of during the role model events:

That they have accessed some of the resources that were indicated from that event, that they have linked up with mentors, that they have done something about joining a computer club or at school the next year, signing up for enrolments in courses, or that they have taken back to their school, this kind of information. That they have encouraged other girls to do things or starting their own club at school, that they have just taken it further and influenced other people in what they have done as well. Because there is a sphere of influence when something can be really significant, not just the 30 girls that are there on the day, it is that whole explosive chart. If 10 of those girls get really affected by the day, and they go and impact other people, it is that classic nuclear explosion outwards. (Lesley_Ind2)

Kerrie (Uni7) suggested that success would be that the girls who attended the summer school would take up ‘careers around technology’. Similarly Jodie (Uni6) suggested that programmes were a success if they changed attitudes:

To see the kids taking the appropriate subjects in senior school to broaden their horizons and then, maybe the icing on the cake, would be to see them into a university or a TAFE college to do IT and then the tip would be to see them here. Jo and I don’t get up there and say Uni6 is good. We don’t need to. That is implicit in the whole programme. (Jodie_Uni6)

For one of the intervention programmes still in its early stages, success focused on the implementation of the programme:

Success will look like having an active web presence, having some sort of mentoring support group within the faculty of ICT that the girls are running, that they are active. Even if it is only four lunches a year, it is something. Having guest speakers, that sort of thing. … Success will be the group running, success will be having organised and built up a relationship with some feeder schools. (Sarah_Uni3)

Gov1 defined success by quantitative results; within a short period of time 30% of state primary schools and 70% of state secondary schools will have participated in Girls’ and ICT initiatives. Additionally 35% of state secondary school students enrolled in ICT subjects would be female (Gov1_documents). For Debbie (Gov3) success revolved around all three stages of Clayton’s framework; participation, retention and the advancement of women in ICT:

For me as a group I am hoping that we achieve a couple of things. One is that we really do get more women into ICT. There are three areas that we are focussed on. I hope that we actually bring more in at an earlier stage into the industry. I hope we can show that we are actually retaining more women and not losing them. Most importantly the reason to retain them is for them to actually progress through the industry so that we actually get more senior women managers, we get more people into positions of leadership that then can make a difference rather than being seen as not in areas of authority or influence. (Debbie_Gov3)

5.5 Were these Programmes Successful?

From the perspective of the programme champion each of the programmes implemented by the case study entities was seen as a success. All the interviewees, when asked whether they considered their programmes a success, said ‘Yes’ though a number of the respondents then qualified their answer;
I think all the events were successful, it depends how you are measuring the success. (Clair_Gov1)

Yes. We tried to measure it quantitatively and we couldn’t. We couldn’t relate anything really to the programme but our qualitative work tells us that women students perceived that the programme was just great and that is enough for me. (Alison_Uni5)

Almost all of the respondents talked about individual instances of having made a difference in one persons’ life and that this was enough to make the programme a success;

Yes. I know that I personally have touched lots of women’s lives from that point of view. That satisfies me but does it really justify the amount of time and effort that I have put into it? I don’t care. Really you cannot always measure things by the direct and tangible benefits so to me if we supported an activity and it made the difference between no girls going into a course and one girl going in I would be quite satisfied with it...(Stacey_Ind3)

While nobody described any of the projects or programmes as failures or unsuccessful, not all responses were completely positive:

There is no doubt that the success of the intervention programme depended largely on the personalities and qualities of the senior women students. Naturally some were better at the task than others. (Uni5_documents)

Partial only. I doubt if I move on that it will last independently. The paradox of the majority gender taking no ownership of the issue persists in this Faculty. We have not been able to affect staff perceptions. (Sarah_Uni3)

It was successful to a certain extent I think but only in a limited way, as far as we know. There was some form of evaluation done but there was so many confounding variables that it really didn’t show a clear influence in decisions made by the girls. (Kerrie_Uni7)

6 DISCUSSION/CONCLUSION

Currently there exist a wide range of programmes that have the potential to increase the participation of females in computing and there is some evidence that these programmes have worked locally. This paper describes a research project which examined fourteen intervention programmes in detail. These programmes were conducted by government departments, industry groups or university departments throughout Australia.

The results indicate that it is clear that one type of programme does not suit all situations and that the characteristics of programmes differ widely. The target audience varied from students, at all levels of education, to women in the workforce or returning to the workforce. Different levels of resources and support were accessible for each of the programmes. While the study showed that programmes differed in various ways some common issues and characteristics emerged:

- Difficulties in measuring intangible and qualitative benefits such as increased confidence
- A focus on implementing programmes
- A lack of focus on planning and evaluation of the programmes
- The dependence on the programme champions including personalities, time available, etc
- The need for management support and buy-in from others in the host organisation was seen as critical for programmes.

There were some identifiable explicit measures of success expressed by programme leaders. These include:

- Participants actively following up the opportunities that they were made aware of during programme activities.
• Participants taking up a career in computing. This included concepts of retaining women in computing careers and getting more women into leadership roles in the computing industry.

• The degree to which management and staff of organisations hosting intervention programmes provide mentoring and support both within the programmes and afterwards.

Most of these measures had not been quantified by the programme champions. This appeared to be due to a lack of initial planning of goals and evaluation criteria but resourcing issues also seemed to play a part. A programme was considered successful by all of the programme leaders if it had made a difference, no matter how small, and even if it was only in specific areas or for small groups of individuals.

The study indicates that intervention programmes currently have some identified measures of success depending on the type of programme and the participants. Although these measures were largely self-reported, they are key to understanding the success of the programmes.

The analysis of the findings indicates that for intervention programmes to be successful they need to be structured, targeted and well planned. They also need to have clearly defined goals and evaluation criteria that can be used to assess the programme at various milestones. The evaluation criteria should be considered in the early stages of the planning of the programme and not just after the programme has been completed. Further work is required to understand how intervention programmes can be better designed and evaluated so that they can be successful and have greater impact.

Ultimately, for intervention programmes to be considered successful, given that the purpose of the programmes is to increase participation, then the participation rate of women in computing education and the computing industry needs to rise. For the knowledge-based society of the future this is necessary in order to reflect the diversity of perspectives across the whole of society and to ensure that the development of computing and information systems is based on the broadest range of solutions available. The current lack of multiple viewpoints in computing and IT development and the loss of available knowledge from women, and potential women, in the computing industry can only be considered to be a loss for society as a whole.

References


THE CHIASMUS OF DESIGN: PARADOXICAL OUTCOMES IN THE E-GOVERNMENT REFORM OF UK CHILDREN’S SERVICES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0041.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-government, Ethnography, Organisational Change, Design research</td>
</tr>
</tbody>
</table>
THE CHIASMUS OF DESIGN: PARADOXICAL OUTCOMES IN THE E-GOVERNMENT REFORM OF UK CHILDREN’S SERVICES

Wastell, David, Nottingham University Business School, Wollaton Road, Nottingham, UK, david.wastell@nottingham.ac.uk
White, Sue, Dept. of Applied Social Science, University of Lancaster, Lancaster, UK, s.j.white@lancaster.ac.uk
Broadhurst, Karen, Dept. of Applied Social Science, University of Lancaster, Lancaster, UK, k.broadhurst@lancaster.ac.uk
Hall, Chris, Centre for Applied Childhood Studies, University of Huddersfield, UK, c.hall@hud.ac.uk
Peckover, Sue, Centre for Applied Childhood Studies, University of Huddersfield, UK, s.peckover@hud.ac.uk

Abstract

This paper presents a detailed ethnographic study of the design problems of a major national IT system in UK child protection and welfare services. The implementation of the Integrated Children’s System (ICS) has disrupted social work practice and engendered growing professional resistance, prompting a fundamental review of its design. Marshall McLuhan’s concept of chiasmus is a central feature of the analysis presented here of the tribulations of the ICS. Chiasmus refers to the tendency of any system, when pushed too far, to produce unintended contradictory effects, and is an intrinsic feature of the behaviour of complex, socio-technical systems. The dysfunctions of the ICS provide a pertinent, large-scale example. The ICS constitutes an attempt, via technological means, to re-organize child welfare services in the UK. Whilst aimed at improving child safety, the ICS has had the opposite effect of increasing the potential for error. This chiasmus has been exposed through the multi-site ethnography reported here, which shows how rigidly designed processes, enforced by IT systems, force social work professionals into unsafe investigative and recording practices which put children at greater risk. The paper ends by proposing an alternative approach to design, based on proven socio-technical precepts, emphasizing the principles of minimum critical specification, user-centeredness and local autonomy.

Keywords: Integrated Children’s System, socio-technical systems, child protection, child welfare, user resistance, chiasmus, user-centred design, performance management, audit.
1 INTRODUCTION

Every process pushed far enough tends to reverse or flip suddenly. Chiasmus – the reversal of process caused by increasing its speed, scope or size (Marshall McLuhan, The Book of Probes, p222)

Chiasmus is an integral property of designing, especially so for the design of organisations. In medicine, the general principle is familiar enough: too much of the “active ingredient” harms, even kills, the patient; indeed, the Greek for remedy (pharmakon) also means poison, as Derrida (2004) archly notes. Chiasmus emphasizes the indeterminacy of design, especially in the context of complex systems, where the presence of multiple, interacting variables means that outcomes are inherently non-linear and unpredictable; it thus furnishes a much-needed antidote to simplistic cause-effect thinking. When designing organisations, structure is the active ingredient, and achieving the optimum dose, the golden mean, should be the pre-eminently concerned of the designer: too little structure risks anarchy, too much can bring on bureaucratic sclerosis. As Argyris puts it ‘All organisations are designs… managers specify ahead of time the jobs and roles of the players as completely as possible without the specifications being so complex that they immobilise performance’. The socio-technical theorists propound the same wisdom: “details of a work system should not be overspecified in advance. Sociotechnical analysis proceeds by specifying… only those things that must be defined: the minimum critical specification” (Pava, 1983). Weick (1987) makes much the same argument, expressing it in terms of the right balance between centralisation and decentralisation.

In this paper, we present a cautionary tale where an excess of structure produced a chiasmus, i.e. a reversal of the desired organizational outcome. We will examine the vicissitudes of a major national IT project, the Integrated Children’s System (ICS)\(^1\), in UK child protection and welfare services. The failings of the ICS were thrown into sharp relief by the tragic and brutal death of a 17 month old child (“Baby P”) in the London Borough of Haringey in August 2007. The trial of the child’s mother, boyfriend and another man, which concluded in November 2008, was widely reported in the British press and caused considerable moral outrage. Reports drew attention to the deficiencies of the ICS and the role it had played. An article in the Guardian newspaper of 19th November (p. 7) is typical of the coverage. Entitled “Child protection stifled by £30m computer system”, the article highlighted the vast amount of time taken filling out forms on ICS and the pressures created by the system’s deadlines, together acting to “restrict the time available for family visits”. The urgent need to review the design of ICS has been designated as a key priority for the Social Work Force\(^2\), a body set up by the Government to review all aspects of social work training and practice in response to the Baby P tragedy. Further indications ICS’s deep-seated troubles are reflected in recent reports of industrial unrest, such as the following from the Unison trade union to which many social workers belong:

UNISON wishes to draw attention to the seriousness of the problems being experienced by social work staff with the Integrated Children’s System. The problems appear to be fundamental, widespread and consistent enough to call into question whether the ICS is fit for purpose…. we have reports of a number of industrial disputes or collective grievances brewing or underway and in many more cases staff are voting with their feet and not using the system when they can get away with it (Unison, 2008).

\(^1\) ICS is not a standard software package as such. Rather it is a standard specification, comprising a set of data requirements, a process model, and a reference set of a data collection forms (exemplars), against which suppliers can develop “compliant” software implementations (Cleaver et al., 2008). A number of ICS software products exist purveyed by a range of vendors, either as adaptations of extant systems, or as de novo packages. ICS was originally conceived in 2000. Its development evolved over a number of years, before “roll out” on a national basis from early 2007 onwards. Although there is an espoused emphasis through its development on piloting and incrementalism, our research has found that, in practice, its design has been highly centralised, led by a small group of senior civil servants and expert advisers. Input from social work practitioners has been very limited and feedback from the field has had minimal impact on its design; oft-heard are tales of requests to simplify the exemplars, for instance, going routinely ignored. Adverse reactions from practitioners have been dismissed as “teething problems” rather than symptoms of fundamental design flaws (see Cleaver and Walker, 2004).

\(^2\) As a result of the research reported here, one of the authors (SW) is a member the Task Force
In this article, we shall tease out the fatal flaws in the design of ICS, flaws which were responsible for the chiasmus of a system designed to enhance child safety seemingly having the contrary effect. These paradoxical outcomes are subtle though, only to be brought to light by careful, ethnographic analysis of the local adaptations of practice arising from the user/technology interaction. It is important to be clear that these failings are not, at root, failings of the IT per se, but of the rigid workflow regime and the excessively complex assessment forms which the technology embodies. Having said this, there is no doubt that the all-pervasive nature of the electronic medium has enabled these arrangements to be more efficiently and comprehensively enacted than would have been the case with a paper-based system, and with less flexibility and “wiggle room” for workarounds and other improvisations.

As a major national initiative directly aimed at transforming social work practice, ICS is worthy of investigation in its own right. We also believe that the problems inherent in ICS are by no means unique, and that there are lessons to be learned for the design of systems in general. These lessons are especially relevant in the context of public sector reform where technology features so centrally as the instrument of “modernization”, with e-Government now installed as a global phenomenon (Wastell, 2006). In child welfare, an important focus for such reform has been the need to increase child safety, through the formalisation of organizational procedures and their enactment by technology.

Our main interest has centred on the interface (the “front door”) between social services departments and other agencies (e.g. police, health services), and with the general public. We draw attention to the short-cuts that the ICS appears to necessitate, given the procrastinuous timescales and excessive audit requirements it imposes. Imperatives to safeguard children and support families appeared at odds with the new modes of e-governance and associated performance targets. Equally, the standardised but tortuous assessment forms appeared to engender a range of problematic recording practices. Whilst new tools and methods can clearly play a role in risk management, we will argue that the design of an effective system needs to be based on the needs of users and on a thorough understanding of their working practices. Only this will provide a secure basis for determining the appropriate balance between structure and autonomy.

2 MODERNISING INITIAL ASSESSMENT

In recent years, practices at the front-door of statutory children’s services in the UK have been subject to significant ‘modernisation’ efforts, aimed at improving their safety and efficiency. In 2000, The Framework for the Assessment of Children in Need and their Families was introduced (Cleaver and Walker, 2004) which clearly defined the initial statutory response as a distinct stage in the assessment process (Horwath, 2002). Since its advent, the importance of initial assessment practices has been further reinforced by the public inquiry into the death of Victoria Climbie (Laming, 2003) which heavily criticized the social services department involved (Haringey again) for failings in their referral and assessment processes.

Initial assessment (IA) emphasizes the importance of the first professional response, especially the need to see the child at the earliest point. The Integrated Children’s System (ICS) has accordingly been designed to ensure that workers rigorously follow the various steps specified in a formally defined “model” of the assessment process, leaving behind an indelible, electronic trace for audit purposes (i.e. showing that the correct procedure was followed). A standardized IA form prompts workers to collect information in a systematic way, with the expectation that the data so-garnered will contribute usefully to further assessment. Consistent with the universal application of “Performance Management” throughout the UK public sector (Bevan and Hood, 2006), a burgeoning range of targets and associated timescales have been also stipulated. Within 1 day of a referral being received, social service departments are mandated to make and log a decision about the requisite response. When an

---

3 We use “technology” in its widest sense to encompass formally-defined procedures and methods (administrative technology) as well as physical machinery, such as the ubiquitous personal computer.
initial assessment is deemed necessary, this must be completed within 7 working days, including the requirement to see the child.

Despite these changes, errors continue to occur, not just concerning the risk of serious harm to children but indications have also been found that “children in need” and their families are not receiving appropriate help (Bostock et al., 2005). Our impoverished understanding of these continuing difficulties arises from a key limitation in the evidence-base. There has been a general tendency within the extant social work research literature to bracket off the technological from the human elements of practice. The present study represents an inter-disciplinary collaboration between social work and IS academics. It proceeds from the general principle that to develop a more effective approach to risk management in child welfare, it is important to conceptualise the ICS as a socio-technical system of inter-related human and technological components, and to draw on relevant ideas from systems theory for enhancing safety. We develop this argument in the next section.

3 THE SYSTEMS APPROACH TO ERROR MANAGEMENT

There is a substantive body of work that argues for such a systems approach to error management (Reason, 1997; Munro, 2005a; Bostock et al., 2005). Reason (1997; 2000), perhaps the most well known exponent of this standpoint, argues that errors in human systems have their origins “not so much in the perversity of human nature as in upstream, systemic factors” (p768, 2000). Coining the notion of “latent conditions for error”, he argues that the analysis of errors in organizational settings should focus on general systemic weaknesses, rather than mistakes made by particular individuals on particular occasions. Such latent conditions refer to generalized, immanent characteristics of a designed system (typically non-obvious and unintended) that increase the overall risk of errors occurring in the operational situation. In aviation, for instance, efforts to automate pilot functions were based on the premise that this would improve safety (Norman, 1990). However, by displacing the pilot from the “control loop”, his/her grasp of what was currently going on in relation to the status of airplane was necessarily eroded. Such reduction of “situational awareness” provides an example of a latent condition of error, a direct but paradoxical consequence of infelicitous design. When the pilot was obliged to take over control in exceptional conditions, the likelihood of mistakes could be increased and indeed several serious accidents were attributed to this aetiology (Norman, 1990).

An extensive quotation from Reason (2000, p.769) provides a useful summary of the main features of the concept. Designating latent conditions as the inevitable “resident pathogens” within the system, he goes on:

They arise from decisions made by designers, builders, procedure writers, and top level management… All such strategic decisions have the potential for introducing pathogens into the system. Latent conditions have two kinds of adverse effect: they can translate into error provoking conditions within the local workplace (for example, time pressure, understaffing, inadequate equipment, fatigue, and inexperience) and they can create long-lasting holes or weaknesses in the defences (untrustworthy alarms and indicators, unworkable procedures, design and construction deficiencies, etc). Latent conditions, as the term suggests, may lie dormant within the system for many years before they combine with active failures and local triggers to create an accident opportunity.

From a systems perspective, approaches to error management that focus on individual breaches of procedure or unwanted aberrations of human conduct will be limited in delivering safer worker practices. Individuals are fallible and will always err, the trick is to design safe systems which minimize the likelihood and the consequentiality of such inevitable failures. A systems perspective requires that more attention is focused on minimizing “the number of latent conditions in the system that can contribute to user error” (Lowe, p.2, 2006).
Another core systems concept we shall draw on is the Law of Requisite variety. In his classic paper on the design of “high reliability” systems, Weick (1987) makes the point that failures inevitably occur when “the variety that exists in the system to be managed exceeds the variety in the people who must regulate it”. Pithily the Law proclaims “Only variety absorbs variety”. This, of course, is the exact opposite of the principle of standardisation, which in the limiting case provides for the same response whatever the input. Rules, policies and procedures are all abstractions and intrinsically lack variety; intelligent human agents provide the necessary “variety amplifiers” (Beer, 1994) which enable bureaucratic systems to work effectively.

In the heavily regulated regime of UK children’s services, requisite variety has necessarily been attenuated and the latent conditions for error are all too readily found. Work volumes create severe pressures in most settings. It is estimated that, on average, some 300 referrals must be processed every month by the typical children’s services department. Teams are legally bound to respond to referrals; they receive no extra compensation or flexibility regarding staff sickness levels, rather targets must be met whatever the particulars of local context or individual case. Pressures are further compounded by the widespread problem of recruiting and retaining experienced staff. Where referral rates are high and resources are constrained, trade-offs are inevitable between urgent child-protection work and assessments leading to more general forms of family support. The general aim of this paper is to draw attention to the unsafe practices that the ICS appeared to necessitate in its demand for rapid case disposals dictated by standardised procedures, and exacerbated by the work pressures and conflicting imperatives that apply in the operational situation.

4 THE FIELD-WORK AND FINDINGS

The project comprised a multi-site ethnographic study, based in five local authorities in England and Wales and drawing data from 15 social work “initial assessment” teams. The five local authority areas comprise: a London borough (Metroville); a county council (Shire); a metropolitan borough in the North of England (Westford); a unitary authority (Seaton); a Welsh rural authority (Valleytown). The field-work began in 2007 and is still underway. Informed by appropriate standards for ethical research, the ethnographies have involved various levels of engagement across the sites. Everyday interactions between team leaders and social workers, middle and senior managers have been observed as well as more formal meetings. In total, we estimate that this has amounted to around 240 days of observation and analysis of everyday practice at the time of writing, supplemented by the inspection of key documents and case files. In addition, a total of 10 focus groups and 60 formal interviews have also been conducted. Transcripts and fieldnotes have been uploaded to a dedicated project web-site to allow the research team to share and discuss the data. Regular meetings have been held to examine and validate emergent themes, supported by group email exchange and discussion. Through these means, we have ensured that the pattern of findings reported below provides an accurate depiction of the situation across the five sites.

Each of the five initial assessment teams is tasked to respond to initial contacts and referrals that come by way of telephone calls, faxes, emails, multi-agency assessment forms etc. These various external contacts cover the range from reports of serious injuries to children, more uncertain concerns about children’s welfare, right through to simple requests for information and advice. In all sites, there is increasingly little opportunity for a “customer” to walk in and directly request help, rather all approaches are mediated through some form of “front of house” customer service interface, either

---

4 The Law was coined by the British psychiatrist W. Ross Ashby, one of the founding fathers of the systems movement, in his Introduction to Cybernetics (1956). Imagine a system (e.g. a family) to be “controlled” in the sense that we seek certain desirable outcomes (e.g. children do not suffer harm and achieve their full potential). Variety simply denotes the number of different “states” that the system can be in, which in the case of a human system (e.g. a family) is both large and dynamic. All families are different and the variety of the social care system must therefore possess a comparable degree of variety (repertoire of responses) in order to deal effectively with this variety.
centralized or within the team. Whilst practices varied across our sites, we found a number of distinct commonalities reflecting the influence of the performance management elements of ICS, and the concomitant preoccupation of staff with maintaining “workflow”. Workers consistently claimed that it was easy to lose sight of the primary activities of supporting families and safeguarding children, to the second order activities of targets and audit. In the rest of this Section, we will focus on the latent conditions for error created by this administrative regime embodied within the ICS. We illustrate these risky adaptations in the sub-sections that follow, limiting our report to three discrete aspects of the initial assessment process, (i) accepting a contact/referral (ii) making further enquiries and seeing the child and (iii) the completion of the IA record.

4.1 Accepting a contact/referral

Our assessment teams reported variable referral rates, ranging from 80 to figures significantly higher than 300 per month. In all but one site, far more contacts/referrals were received than could be managed. The requirement for an initial decision within 24 hours necessitated a rapid but not necessarily reliable response, and where workloads are high, the potential for error is clear:

Admin worker: ‘The phone will be ringing continuously, you put the phone down and it rings straight away... one comes and another one comes... and your mind just gets frazzled, I might have written 5 or 6 pages of A4 paper ...and when I come back to reading them, it’s all looking a bit messy... I can’t quite make out what I’ve got down...’

Such pressures created significant anxieties for experienced staff workers, who understood that the pace of work created less than ideal conditions for practice, but there was also evidence of ‘speed-practices’ becoming habituated and normalized, especially in newer staff. For staff higher up the hierarchy, it was critical that only a manageable number of referrals were actually allocated for initial assessment. In the extract below, taken from one of our busiest teams, the team leader makes clear her reasons for clearing contacts by the end of the day, with the mandate for her actions directly invoking the exigencies of managing work-flow dictated by performance timescales.

Team leader: ‘There are 50 contacts in your inbox... you are under pressure because you have to clear them by the end of the day ...and the question of whether you are more likely to close them in these circumstances? Well yeah... so, really we are looking to close cases not open them... that’s why we work to the highest thresholds’

The IT systems maintained the pace of work, typically by providing digital reminders of deadlines and timescales. In one site, we found an ‘e-tracking device’ in the form of traffic lights, which informed workers about how much time was left before the specific episode was deemed out of timescale. In another site, ‘higher management’ were planning to print out weekly graphs of levels of attainment in meeting targets, alongside tables exposing individual failures.

In order to manage the volume of referrals, we consistently observed the presence of well established ‘general deflection strategies’ that included: strategic deferment, i.e. sending the referral back to the referrer to ask for more information; and signposting, deflecting the case to a more ‘appropriate’ agency. Whilst such adaptations are sensible if proportionate, the inherent risks are also clear. Where insufficient time precluded the pursuit of more detailed information from a referrer, other decision-making heuristics came into play. These included the routine categorization of anonymous referrals as malicious (indeed referrals from neighbours and family members were also often treated as suspect). We were told that children aged over 13 were routinely ‘NFA-ed’ (an outcome of no further action was recorded) on the basis that these young people ‘must have lived with these concerns for a long time and be quite resilient’. Similarly we found questionable methods for dealing with domestic violence notifications through automatic letters generated without any attempt at investigation. In general, we found that well-intentioned, but very busy workers, became habituated to such methods of rationing, with little time to reflect on such rationales and the risks they entailed.
In the following extract, the worker indicates how the imperative to prioritize a case already categorized as child protection (S47, i.e. section 47, Children Act 1989) required that she give less priority to an incoming referral, which also sounded malicious. Although she acknowledges her lack of knowledge, she justified her decision as follows:

*Social worker:* ‘I’ve got this S47 and actually this family are in crisis and I want to put support in for them, before I worry about this other family that don’t even know I’m coming, because it’s an anonymous referral from a neighbor... and you think, well OK, I don’t know if there’s a real risk or not, but from reading it, it sounds a bit malicious, well this family, actually are about to fall apart if you don’t put something in’

In order to manage workflow, we also observed some ‘safer’ locally-improvised methods for meeting timescales, that generally amounted to holding a case open for ‘review’, but logging the IA as complete on the system so as to meet the target. In cases where the seven days had not provided sufficient time to establish confidence about the child’s welfare, this ‘review space’ could enable further information to be gathered. However, such workarounds, even when they are constructive, by their very nature can only survive while they remain undetected by the gaze of senior management.

4.2 Making further enquiries and seeing the child

A number of cases will get through the first layer of filtering and be allocated to a social worker for initial assessment, which will include making further enquiries and seeing the child. At this second stage, we also found short-cuts in operation. There was a tendency to abort an assessment whenever the ‘opportunity’ arose. In the case of the referral from a grandmother below, the routine treatment of referrals from family members as potentially suspicious, and that the health visitor had seen the child, together enabled swift disposal of the case.

*Team leader:* ‘Being a bit cheeky...we contacted the health visitor and said when did you last see the child and lucky enough the health visitor had seen the baby recently and it wasn’t as bad as the grandmother had alleged... so we didn’t take it any further, no further action’

The tempo and volume of work, together with the 7-day target for IA completion, were widely reported as making cases at this second stage equally susceptible to partial analysis and rapid disposal.

*Social worker:* ‘If it’s not looking that serious...sometimes you don’t get all the information and the temptation is then to take a short-cut and maybe not contact the school, or because the school are on holidays you say I think I’ve got sufficient information to make a decision- NFA’

Needless to say, school holidays are not factored into the 7 day timescale! Neither are parents and children who are not at home, nor health visitors who are on sick leave and so forth. These factors necessarily interrupt the expeditiousness of the assessment process, but the system offers no accommodating room for manoeuvre. Thus, timescales can create perverse incentives to dispose early on the basis of incomplete information. Whilst in many cases, an “NFA” decision may be quite appropriate, our file analysis of open cases did find a common pattern of repeated initial assessments of escalating severity, before the case eventually found its way through the front-door.

Front-line team managers played a key role in the operation of ICS. In some teams, acutely aware of the possibility of error, managers worked closely alongside new recruits to defuse their inevitable frustrations and induct them into ‘local methods’ which would enable them to get them to the best out of ICS. However, we also found some tension between workers and their managers regarding the primacy of meeting targets. For example, in a particularly pressurized team, managers described their frustrations with workers spending too much time ‘social chatting’ or needing lessons in ‘diary management’; such critical attitudes would seem likely only to exacerbate anxiety in work that was already stressful enough. Invoking ‘safeguarding’ could buy a worker more time, but only in cases where there was a clear moral mandate to set aside the all-important target. The degree of assertiveness required to challenge the performance system could also lead to overt conflict:
Social worker: ‘My manager said to me “why haven’t you finished that yet?”… and I said “well the health visitor hasn’t called me back”… and they said, “well no, if you’ve decided that it’s family support, then the outcome won’t change, whatever they say”. I said “I disagree” and of course that information informs my assessment, I’m not putting my name to that.’

Where workers were juggling between the completion of IAs and serious cases needing to progress for further investigation, ‘NFA’ was described as a welcome relief. Again, we see the latent potential for errors in this expediency; in the busiest of teams and in spite of the good intentions of workers, time precludes, for example, getting back to the referrer to inform him/her of a decision, which closes down any immediate challenge to the categorization. Seeing ‘the child’ is a central and critical part of initial assessment. However, even in relation to this imperative, we found worrying short-cuts, as the following extract illustrates:

Social worker: ‘My new manager…she comes back and says, it [the IA] doesn’t say have you seen the child, it says “has the child been seen?”, you can put “yes” and then make it clear that the teacher has seen the child. I thought hmm… what’s the point me even doing an assessment if I haven’t seen the children’

Children are not easy to ‘see’ under the conditions of initial assessment, for a variety of reasons. First, there is a requirement to see all children irrespective of a ges, but older children can be difficult to track down. Second, ‘seeing’ should involve talking to the child alone to make an assessment of the child’s development and needs, but this can be hard to achieve within 7 days in a single visit. With initial response to telephone calls increasingly mediated by administrative staff and home visits curtailed to a single visit, the space between help-seeker and help-provider is steadily widening. Skilled workers might attempt to reduce this space, but for many, in the absence of knowledge derived from face-to-face work with families, they fall back on readily auditable, bureaucratic justifications, often offered by fellow professionals, which invoke missed health appointments, school attendance problems and the like.

4.3 Completing the record

The standardized initial assessment record invites workers to comment on a range of factors relating to the child, his/her parents or carers and the presenting concerns. A general observation across our sites was the paucity of information recorded on the actual IA document. In one site, we examined 65 records of individual children; the scantiness of the information, compounded by the difficulties of piecing together fragments of narrative scattered across multiple boxes, made it very difficult for the reader to construct a holistic picture of the child and his/her family. The IA record presents as a rather badly designed tool, requiring copious information that is difficult to glean from one home visit and from other professionals; it thus invites workers to discard the majority of its sections as irrelevant.

With not unsurprising consistency, we found an expedient method of ‘front and back-ing’ (or ‘back-to-back-ing’) had spontaneously sprung up across all our sites, wherein middle sections of the document were omitted altogether:

Researcher: ‘So what about the middle of the document, because everyone seems to miss this out?
Social worker: What middle document?
Researcher: You know, practitioners are concerned with the referral and the outcome on the back, but what about all those pages in between about the child?
Social worker (laughing): To me well… yes, there is a page about the child, I would always put in something, depending on what the child is like…I would always put something in, but in IA you wouldn’t…this is initial assessment’

It was clear that workers were trying desperately to make the form fit their work. Whilst most practitioners welcomed the general principle of electronic recording, the IA was not only overly long.

---

5 The initial assessment form is typically 10 pages in length. Apart from the usual administrative fields for structured data, there are well over twenty free text fields addressing the developmental status of the child and relevant environmental factors: including the child’s ‘social presentation’ or ‘self-care skills’, the ‘family’s social integration’ etc.
but the standardized questions and sub-headings were not easily adapted for this or that case (c.f. White et al. 2008), so workers went straight to ‘analysis’ of ‘that dreadful form’, putting ‘nothing-in-between’. Scrolling through the pages of the record ourselves, we found it difficult to distinguish between the material typed by workers and material already on the form, i.e. the numerous sub-headings and explanatory notes provided. Workers have become experts with the copy-and-paste function, as material is regularly and mechanistically repeated. In addition, the principle that for every family, a record of each child was required, tended to encourage practitioners to produce a general homogenizing account that ‘fitted’ all the children.

**Social worker:** ‘If I know that the IA is more than likely going to turn into no further action, and I know that after I’ve had my conversation with the family, then I will massage the information on each of the children and talk in plural “the children presented”’.

Whilst workers were clearly attempting to work-around the excessive audit demands of ICS, to salvage some time to spend with families, even the most perfunctory response to audit left too little time for the real work of face-to-face communication (Peckover et al. 2008; White et al. 2008). The speed with which workers attempted to complete the IA record also meant that errors of recording were common. In the busiest teams, such errors were compounded when ‘students’ (for example) were asked to catch up with recording cases they knew nothing about! Whilst workers were aware of these errors, they also reported that it was difficult to make corrections as material was ‘locked down’ in the system after 24 hours. After that period, a worker would have to seek special permission to undertake corrections. This stricture, entirely rational in the logic of rigorous audit, is however hardly conducive to good practice. Very obvious errors, such as putting a case note in the wrong file might prompt such requests, but simply improving wording was just too much trouble, though vital to the comprehensibility of the record, especially for other professional readers.

## 5 DISCUSSION

The Platonic approach can be described as Utopian engineering, as opposed to … piecemeal engineering. The Utopian approach is the more dangerous… the Utopian engineer will claim that mechanical engineers plan even very complicated machinery as a whole and that their blueprints cover not only a certain kind of machinery, but even the whole factory…. (Karl Popper, The Open Society and its Enemies, 1999)

Since the Laming report (2003) there has been a very significant reconfiguration of children’s statutory services in the UK. Rigorous information recording and performance management are the mantras of this brave new world; however it is important that the tools we provide for our workers are fit for purpose (Munro, 2005b). We would argue, however, that the increased audit demands of ICS, together with on-going resource constraints, have served to increase the burden on front-line workers (Peckover et al., 2008). From our analysis, it is clear that the design of the modernized initial assessment system of children’s services, embodied within the ICS, is not only flawed, but that its dysfunctions provide the latent conditions for error. In response to the intractability of the IA process, the workers in our study had devised a range of artful “work-arounds”. However, errors are inevitable in the context of such expediencies, which have been extemporized to maintain an overly rigid workflow. An excessive zeal for structure and standardization has thus engendered a reversal of the intended outcome. A chiasmus has been produced in which, rather than improving safety, the latent conditions for failure have been exacerbated and risks increased.

The short-cuts that have been fashioned in the “electronic cage” of the ICS typically take the form of early categorizations based on incomplete information, or the fudging of details of a ‘home-visit’, and so on. As illustrated in our ethnography, it is preferable to dispose of a seemingly nebulous referral in the face of the more immediate performance demands of the ICS. But it is often just those kinds of referrals that appear to be irrelevant, or somebody else’s business, that can provide the warning signs of a more serious malaise (Laming, 2003). Scarcity of resources will inevitably mean that giving priority to one part of the system, the most immediate, will result in cuts in another. It is no surprise to
have found that workers pursue opportunities to deflect incoming work, and dispose of cases on the basis of superficial analysis, or fall back on fallible heuristics such as “it’s probably malicious”. In social services, priority needs to be given to reduce the distance between workers, family and community which many studies have cited (by both service users and front-line workers) as central to good practice (Pithouse and Holland 1999; Gray, 2002). We have seen that the performance-driven ICS only detaches the professional further and further from the possibility of meaningful engagement with service users, offering instead a scientistic veneer of codes, risk scores and metrics.

Lipsky’s concept of the “street level bureaucrat” (Lipsky, 1980) emphasizes the importance of professional discretion in front-line practice in the public services. Lipsky saw such discretion as essential in order to get the job done: “the situations they face are too complex to reduce to prescribed responses”. This is simply a reflection of the ineluctable writ of the Law of Requisite variety: discretion is not some incidental feature, it is fundamental to the operation of any “viable system”, and of particular importance in designing reliable systems for risky environments where failure is highly consequential (Weick, 1987). In a different context, Bourdieu (2003) makes the same point in defining the ability to improvise, to adjust responses to local, situated contingencies, as being the hallmark of competent practice. In the IS field, Ciborra (2002) also exalts improvisation in the modern organisation, as the artistry needed “to fill the gaps of planning, cope with unexpected consequences, with events and situations that do not fit the planned hierarchical procedures” (pp. 153-154).

The findings we present are controversial and are presented in a designedly polemical tone. We draw attention to the multiple opportunities for errors on the part of front-line social work professionals that are exacerbated given the current configuration of initial assessment process and its technological embodiment in the ICS. Performance management (the paraphernalia of targets and indicators) is designed to enhance rather than inhibit quality performance, yet our study has found this regime paradoxically worsens the latent conditions for errors. Whilst it is tempting to berate the maverick professional who subverts correct procedure, it is important to remember that there are “good” organizational reasons for such behaviours, i.e. they are an attempt to reconcile the competing elements of the ICS with imperatives to safeguard children and to support families. The latter role, however prominent in current welfare policy, is particularly vulnerable. Perhaps the real tragedy of ICS is that in busy teams, demands to support families will inevitably be subordinated to pressures to maintain workflow.

Although this paper challenges the huge investment in systems of performance management and IT, we are not arguing for a wholesale Luddite abandonment of new modes of governance and new technology. The remedy, we believe, lies elsewhere, in a radically different approach to design. Radical it would seem for present times, in the public services at least, though hardly ground-breaking; rather the reclaiming of an older design wisdom, sundered-off, ignored, or never encountered. Sociotechnical systems design (STSD) embodies much of what we have in mind. Core principles of STSD include: user participation, minimum critical specification and the optimization of local autonomy (Pava, 1983; Mumford, 2003). Above all, it is essential to focus the design of systems on the needs of users, founded on a rigorous understanding of their working practices. The case for user-centred design (UCD) has been cogently made in many design disciplines, including information systems and human-computer interaction (Norman, 1998). UCD is essential in order to gain reliable knowledge for designing new tools and processes. Failure to involve users in the development of systems inevitably engenders alienation, and there were unmistakable signs of practitioner disquiet (complaints of additional workload and excessive “bureaucratization”) in pilot studies of the ICS (Cleaver and Walker, 2004; Cleaver et al., 2008). It is regrettable, unconscionably-so, that such early warning signals apparently went unheeded, written off with oracular reasoning as “implementation issues” rather than more fundamental problems of design dogma. As a result, the strictures of the work regime imposed by ICS have not only produced unsafe practices but, as we have seen, are now provoking overt resistance from an increasingly frustrated and mutinous workforce (Unison, 2008).

We believe that new systems and technologies can be developed which both assist the users in their daily work and achieve desired organizational goals, but without an ethnographically-informed
understanding of human practice (such as this paper provides), this virtuous circle will not be achieved. Ethnographic studies have shown time and again that even work which seems highly routine is a skilled accomplishment (Gasser, 1986); its orderliness is a product of the artful worker, not determined by the imposition of a formal rule-base. It is noteworthy that some design methods explicitly call for ethnographic engagement in order to develop a valid evidence base for design, e.g. the SPRINT methodology which has been specifically developed for the public sector (Wastell et al., 2007). There is no final guarantee against chiasmus, but the deployment of such user-centred approaches certainly offers a less hubristic way forward.

We claimed in the Introduction that our findings have relevance beyond the domain of children’s services. Certainly, there appear to be parallel examples of similar problems in other large scale IT-enabled modernization projects, such as the gargantuan National Programme for Information Technology (NPfIT) in the UK health service. Eason (2007, p. 258) argues that NPfIT has generally followed a “push strategy, thrusting new technology into the healthcare practices of the NHS”, leaving little room for local design. Eason finds strikingly similar local adaptations (workarounds etc.) to those unmasked here, as well equally concerning symptoms of stress and “mis-use”. He goes on to argue a similar case for a flexible socio-technical approach fostering local diversity, based squarely on user needs. The command-and-control, performance management regime (combining “targets and terror”!) which we have seen in children’s services is pervasive across all UK public services, producing a common pattern of dysfunctional effects (Bevan and Hood, 2006). It is surely time to move away from such crude and self-defeating managerialism. Rather than pressing ever more urgently the cause of top-down centralization, salvation surely lies in reversing the direction of travel, reclaiming an older mode of “control” founded on common values, decision premises and assumptions, i.e. the culture deriving from the shared professional task. Public service agencies (schools, welfare, health etc.) are loosely-coupled organizations; systems such as the current incarnation of ICS, which attempt to over-screw this coupling, inevitably impair performance (Perrow, 1984). Only by loosening the electronic yoke and amplifying discretion will requisite variety, and therefore reliability, be augmented. On this point, it is fitting to end with another sagacious utterance from Karl Weick (1987, p.124):

[A] system in which both centralization and decentralization occur simultaneously is difficult to design. And this is where culture comes in. Either culture or standard operating procedures can impose order,. but only culture also adds in latitude for interpretation, improvisation, and unique action.

Acknowledgments

This research was funded by the ESRC Public Services Programme: Quality, Performance and Delivery, grant number Res-166-25-0048.

References


Lowe, C. M. (2006) Accidents waiting to happen: the contribution of latent conditions to patient safety, Quality and Safety in Health Care, 15 (supplement 1), 172-175.


APPLYING LESSONS LEARNED FROM COUNSELLING:
ON NURTURING RELATIONS IN DESIGN PROJECTS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0310.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>User participation, Trust, Organisational Change, Hermeneutics</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
APPLYING LESSONS LEARNED FROM COUNSELLING:

ON NURTURING RELATIONS IN DESIGN PROJECTS

Cattani, Claudia, University of Trento, Department of Sociology and Social Research, piazza Venezia 41, 38100 Trento, Italy, cattani.claudia@gmail.com
Calzà, Diego, eTour, via Grazioli 63, 38100 Trento, Italy, diego.calza@etour.tn.it
Jacucci, Gianni, University of Trento, Department of Sociology and Social Research, piazza Venezia 41, 38100 Trento, Italy, gianni.jacucci@unitn.it

Abstract

This paper elaborates on the personal relation between the facilitator and the participants in Social Practice Design. It is argued that such processes can not just be managed, but have to fostered in relatively free way, so that results can transcend expectations and more closely approach the actual possibilities. This is explained by aspects of Rogers' theory on therapy.

The paper aims to be an interesting and strong example of the critical need for a good relationship in facilitating design. By itself, such a conclusion would not be surprising, but some of its constituent aspects are detailed: the paper elaborates on the relevance of deploying focus and effort on personal relation, in interventions for organisational innovation.

Supporting the establishment of sense making and trust with Social Practice Design (SPD) approaches is found to be of primary importance in an e-Government development project. Here regional employees user-design a computer-based aid for public tender editing – a tender configurator - with the support of facilitators. We address the structural problem with infra-structural measures including open conversations to promote shared understanding, and user design laboratories to promote concept emergence and learning, while practicing relation and trust building all along. Our constructivist approach renounces from the start to solve the governance problem within a narrow managerial perspective.

The paper offers a demonstration of the mission critical relevance of the relational component in SPD, intertwined with the customary functional component, in resuming governance towards project success. This experience is far from a complete experiment. But a wealth of indications and partial results have been harvested on needs, opportunities, and practices, for promoting shared understanding and trust in the project, and letting emerge idiosyncratic solutions.

We judge the quality of the SPD approach by three requirements (Baskerville and Myers 2004): a contribution to practice (the action), a contribution to research (the theory), the criteria by which to judge the research, and we show explicitly how the research in the case meets these criteria.

Keywords: participatory design, organizational change, social practice design, trust, sense making
1 INTRODUCTION

Governance in design projects of social informatics services is still problematic. User participation is essential for Participatory Design (PD) of sustainable IT use (Bodker, Kensing, and Simonsen 2004). However, often participation alone is not enough. Trying to define better requirements within a given conceptual and pragmatic sense-making frame appears not to be feasible. In these cases, design cannot be confined between requirements and solutions: even the given frame must be abandoned, and new rules of the game and goals co-produced by all stakeholders. This happens in particular when there are no shared concepts, motivations, goals, and even language, among the various actors: the establishment of relation and trust in the project is hindered by ambiguity. To overcome the stall, one has to abandon the given frame, trying to recover sufficient sense making and shared understanding (Jacucci and Martin 2008). If facilitated by management through appropriate infrastructural measures, like in co-production Social Practice Design (SPD) approaches, as soon as trust relations are re-established, a new frame emerges, co-constructed by relevant actors, along with brand new outcomes, solving the structural problem (Jacucci 2007; Jacucci, Tellioglu, and Wagner 2007, 2008; Cattani and Jacucci 2007). Also changing the meaning of the whole process, however, and heavily reshaping its governance beyond what intended, and suggesting new practices and a new deontology in facilitating participative design.

2 THEORY

2.1 Managing ambiguity, taking a second step back

Watzlawick (et al. 1967), in studies of the pragmatics of human communication, identifies content and relation as equally relevant. Bion (1961) in studies of the psychology of groups distinguishes between the task to be executed, and the equally important establishing of relation and trust in the group. Is there a dichotomy in our interventions for innovation in technical systems and organisational processes, between attempts to establish functional rationality through analysis of content and task, on the one hand (reducing uncertainty), and attempts to establish culture change through communication, relation and trust, facilitation and learning, on the other (managing ambiguity)? Posing this question corresponds in the words of Bourdieu (1992) to be taking a second step back. We have discovered in the last decade the need to re-balance focus and effort in current approaches between function and communication. (Cattani and Jacucci 2007; Jacucci and Martin 2008).

2.2 Nurturing trust with Carl Rogers’ qualities

Carl Rogers (1951), in his person centred approach studies of therapy, advocates the respect of three criteria, three qualities of the counsellor (non-judgemental unconditional acceptance, empathy, congruence), for establishing working communication and trust relation with the client. Trust is a pre-requisite for establishing a web of shared understanding. In alignment with criteria introduced by Carl Rogers in personal therapy, and extended already by himself to learning and company consulting (Rogers, 1969), non-judgemental respect, empathy, congruence, are needed here, to allow the establishment of trust. This calls for, on the other hand, introspection, self-awareness, clear and clean intentions, and control on the part of facilitators, as well as are their awareness and intentionality in conceiving, proposing, co-constructing with clients the appropriate path in that context towards the desired development; a path that is more important, if possible, that the very content and task objective.

2.3 Facilitating with SPD

Social Practice Design has recently emerged by urgency of praxis, as a ‘sand-box’ type approach, seeking to construct winning paths to ensure that the potential benefits of envisioned novel technologies can be realised. We consider SPD a methodological extension of PD to the
implementation phase of information systems, often intertwined with design, however. It is not to be regarded as rigorous method, or an apodictic truth, rather as the concocting by bricolage of useful responses to stringent needs. SPD entails aspects of research intervention and facilitation, in which facilitators elaborate with organisation personnel visions of solution, i.e., how to attain desirable change goals by leveraging on organisational assets and strong points, and/or how to cope with problem issues, possibly emerged through ethnographic observations and recognised and accepted in previous phases of the SPD work. In SPD, facilitators thus have a crucial counselling-type job, for the sake of which they must follow a holistic approach, performing an accompanying task from start to finish of the innovation path, and for the success of which the establishment of a good communication and of a good relation of trust is all important. Actions, communication, behaviour of the facilitator establish trust with, and support, the client – at all levels: individual, group, organisation – towards change, for promoting the conscious and proactive care taking and hospitality (Ciborra 2002) to the introduction of socio-technical solutions: a person centred approach, an intentional way to proceed, designed in all its passages.

2.4 Recursive intervention and introspection

This paper puts emphasis on the relevance in SPD of Rogers’ counselling-derived perspective, oriented towards trust gaining, learning, and the reduction of ambiguity. A recursive structure then characterizes SPD: it is a continuous coming and going between client and facilitator in a dialogic process in which the facilitator is capable of observing, performing ethnography, elaborating on the observations, constructing visions, but paradoxically also of forgetting them when coming back to the client (“being without memory and desire”, says the group theoretician (Bion 1961)), to reconstruct afresh visions with them, putting to practice in that instant her competences, especially in the communicative, relational domain (ability of managing a flexible and open communication: Schein (1987; 1999), Rogers (1980)), as well as competences on group dynamics. We underline the process of continuous interaction with the client, and the ability of the facilitator to work substantially on self, in order to acquire the necessary ability and ‘purity’ of approach that really renders her capable of helping the client find themselves the solutions to their own problems (Rogers 1951).

3 INTRODUCING THE CASE BY EXCERPTS

This paper elaborates on experiences of practicing SPD approaches in an e-Government development project, the ‘Online call-for-bids’, at the Regional Administration offices of ‘Celtia’, one of the regions in Northern Italy. Here regional employees are brought to user-design a computer-based aid for public tender editing – a tender Configurator -, with the support of facilitators employing Interactive Use Case (IUC) as a PD tool (Calzà, D’Andrea, Jacucci, and Baskin 2004). IUC is a pictorial and interactive tool positioned in between Use Cases and Mock-up. We call the IUC tool “Interactive Use Case”, underlining the importance of the interaction with users, where it can provide a common place for a dialogue between system developers and system users. The research intervention lasts about half-year. ‘Clients’ of the intervention are the manager, and the employees of the administration, as well as technology designers of the regional administration owned software company: Celtia Informatica, and employees of the RTI (Regional Training Institute).

3.1 Browsing case lab-notes

a) The manager, expert about the tender Configurator issue and capable himself, had built by technicians on his own instructions a non participatory-design prototype, to which personnel charged with tender design responded without enthusiasm, with very few comments. In a subsequent user workshop, proposed by us, it had become evident that the Configurator software employees a philosophy, a procedure for the design of the public tender, which is completely out of tune with respect to the real work practice: the fact always emerges, that software designers abstract, hence depart from reality.

b) A proper user design laboratory is then organised on our proposal, for best results and user appropriation. Although uncertain at the beginning about direct user participation to design, the
manager is struck that we now in-act a true, participatory design laboratory, not just consulting, not just training; a true design laboratory, requiring the construction of a real work team.
c) The manager, kept out of the user design laboratory – our request to avoid hierarchical self-censorship by personnel -, is furthermore struck by our live rendering of that experience, not a power point, not a word document, but live video reproduction of the work session, with access to people discourses and action, with results recorded, and directly usable, via the Interactive Use Case tool.
d) In the end, he is so touched by the content of the presentation, and the obvious success of the user design session, that his doubts vanish, and he now asks that users be invited to redesign the tool, entirely from scratch, without influence from his own prototype; technicians too, present at this rendering session, shared the manager’s inclination to take risk and opportunity and ask the group to redesign the Configurator from scratch.

3.2 Anticipating inferences we shall deal with in the case

i) SPD building trust with manager – Manager: Here we see the high level of trust reached by the manager towards facilitators, his staff, and his own leadership. Trust relation with the manager has been constructed along the way with attentions like: timely and detailed reports, video recording, documents made available online, periodic update meetings, communicative attentions, and so on. So that he has perceived to be taken into good account, both professionally and personally.

ii) SPD empowering users – Participants: Participants enjoyed working in a group, found themselves at ease, as tasks demanded of them were simple, not stressful, albeit serious. They liked the method of hands on, of group work, also of written individual reflections. The video camera disappears, fundamental for documenting the work, no more invasive. People at last talk among themselves, share their work practices, never happened before. All this is more important that the goal: the path as all important.

iii) SPD designing for user design – All: the magic of when somebody with practical knowledge of her own work, builds herself the technology instruments to support her work, and as she goes on building, also performs continuous checks that all knowledge is preserved, embedded in the artefact. If the knowledge in question is tacit, it cannot be transferred to the designer. In turn, when designers build a software application for own use, things invariably go well.

iv) SPD inheriting from counselling - We can see the importance of the facilitator and of her direction of the performance: importance of the events she calls clients to participate in. Importance of how the facilitator IS and BEHAVES in situation (see Rogers: work of facilitator on self, re-elaborating how he/she lived through the group work, must observe how things went – video -, must build the passages each time).

v) SPD facilitating sense making by care taking - Importance that participants work and share viewpoints among colleagues, the facilitator here is fundamental, as must accompany this awareness process on their own work and on the work practices it consists of. It is important to make people work not on theory and abstract concepts, but on what they know, and to make them do, use their hands. It is also important working with them at the meta-level of reasoning, as this promotes awareness. And a new, co-constructed solution emerges.

4 ILLUSTRATING CASE INVESTIGATION AND OUTCOME

We present our investigation, and its outcome, enriched by photographs and brief transcripts of recorded discussions. Links to session videos are also provided.

4.1 Case investigation

We carry out our research-intervention with the specific Action Research methodology of observation, interpretation, step taking, outcome evaluation, and learning. The process is iterative as usual, with a succession of interaction events with manager and users. (see: Baskerville and Myers 2004). We describe the process of applying SPD step moves in the building of relations, sense making, and trust. The use of the SPD approach is exhibited by moving through the various stages of the iterative process, and seeing the congruence between the objectives set, and the characteristics and focuses of
facilitation (e.g., the establishment of trust relationship with the manager, and his understanding of the process). We also hint to what the facilitation team does between meetings: strategizing, analysing previous step, planning next step, material preparation; but also analysing observations in videos, and dialogues: emotions, behaviours, state of mind, the things that went well, the things that went badly, SPD attentions, counselling-like activities, etc. Each step is dealt with individually, recounted, discussed, linked to the next step, albeit we warn the reader that some lesson learned in some steps seems not to be confined to that step, but produced by along the WHOLE pathway. Counselling like qualities (acceptance, empathy, congruence) are constantly emphasised along this pathway. In the reflexive reconstruction we show how things worked because these criteria were followed. Evaluation criteria for AR results will be the degree of satisfaction of company manager, personnel, and facilitators/researchers, respectively, for change produced and final results, for theory and practice.

4.2 Report of selected research-intervention steps (steps are listed in bold)

More specifically, we define a number of interaction steps in our intervention, and give them names. For each step we discuss objectives, observations, lessons learned, and decisions taken for the next step, highlighting attentions paid to counselling-like qualities. For space limitations, we choose to report here the details of selected research-intervention steps only.

1. Needs collection

(Briefing with client for the collection of information regarding expectations and connected problems).

2. Presentation to the client of the project as a laboratory

a. - The objectives of the activity and expected results: Make the manager understand the potential of Participatory Design of computerised systems, and that their problem could be dealt with using this approach. Explain the potential of laboratory-type activity, which involves the group in real tasks.

b. - SPD attention points: Consolidate the bases of trust, collaboration and the personal relationship between the facilitator – Diego - and the manager.

c. - How things went and results achieved: During this meeting Diego presented the PD approach, which is very different from what the client expected. It was presented with great emphasis on the relationship among expectations, complexity and method to deal with the problem. During the presentation the client also stated the results expected more specifically: analysis of the current Configurator, connections with work practices, vademecum on using the instrument, training for all the contract announcement designers in the Celtia Region. These requests the Diego assessed to be coherent and achievable with the method proposed.

d. - Learning: It was important that the presentation of the method be brief, but focused on showing that change can be planned/achieved with a participatory format, and that the new work instruments, and the change that they induce in work practices, must be defined with users. It was important to “really” listen to the client and his statement that his aim was to promote change in his organisation. The client was already aware that realisation of the technological instrument would be the means of change but that the true end was change itself. This legitimated our approach.

more b. - SPD attention points: This phase moment was fundamental for the relationship between Diego and client. The relational/communicative style used conditioned the entire path. It is therefore necessary to conceive and plan how to implement it in a congruent, empathetic and conscious manner.

3. Laboratory: sub-groups tasked with defining and describing the work practices used to design contract announcements (18 March)

a. - The objectives of the activity and expected results: Sharing of work practices among the tender editors, creating a fertile humus for awareness. Get them to examine their own practices, measure the differences between what they do to design tenders and what they are asked to do with the technology.

b. - SPD attention points: Make persons communicate in a protected context (that of exercises/training), enabling them to exit from the usual work setting, with its rules, restrictions, roles (assigned, perceived), consolidated (also in their “pathologies”) rituals and relational communicative
procedures. Unconditional positive acceptance; sense-making, artefacts to facilitate sharing (posters in sub-groups, conceptual maps on stimulus questions). Attention to individual empowerment.

c. - How things went and results achieved: Division into workgroups. First activity with use of post-it notes at a sub-group level on a stimulus question: what does working on a contract announcement involve? Initial brain storming within each group, simple but powerful instrument (post-it notes can be easily moved around on a poster for later categorisation and reorganisation). Afterwards, Diego held a plenary session for sharing results (summary of post-it notes) through co-building of a map. Diego used projector and the Freemind application to represent a summary of the work involving all present.

PHOTO 1: Post-it notes sub-group activity
PHOTO 2: Freemind map, sum of design activities

continues c. - How things went and results achieved: Explore the Freemind map, describing the work scenarios: who does what, how, when, why etc. Work again started from a stimulus question: what do you when you design a call for tender? Each sub-group worked alone. Complicated things: sense making on the concept of scenario (what we expected from them regarding the concept of scenario). No wide-ranging scenarios emerged, contrary to expectations. The final product did not seem very satisfactory to facilitators but it was fundamental the process: birth of a functioning workgroup.

more b. - SPD attention points: Diego demonstrated positive unconditional acceptance towards all the participants and intervened only to help them perform the work requested, responding only to questions about the method to follow; he was also an example of open behaviour, empathetic, welcoming; he supported the sense-making process enacted by the groups on their work practices.

d. - Learning: Ask answerable questions and model them with artefacts. It was this work that gave rise to the idea of organising a laboratory presenting the Configurator in situated form, that is, producing a call for tender from beginning to end. The manager asked for advice on how to present the Configurator and we in fact said: present it with a contextual example.

4. Workshop on the contract announcement Configurator (19 March)

(The manager presented his way, his practice, of creating a call for tender supported by Configurator and those present listened and gave feedback (they were immediately able to see differences, the distance). We are aware of the importance of this encounter between the manager, inventor of the Configurator and the users. The manager immediately realised that the participants had numerous observations, many modifications, and with competence and commitment. The next step – he group analysis of Configurator and its functionalities – was obligatory and natural even for the manager).

5. Laboratory: (a) sub-group analysis of the Configurator, and (b) sharing, managed by Diego who handed over to the various groups, which performed a screen-by-screen analysis. (7 May)

a. - The objectives of the activity and expected results: The stated objectives of the work session were: analyse work practices, measure the distance between the work practices of designing a call for tender in reality and the new practices imposed by the Configurator; construct an artefact.

b. - SPD attention points: Consolidate group’s work practice and sharing process by involving the group in practical tasks through the production of artefacts. Reinforce the participants in their
commitment and participation, creating further opportunities for them to demonstrate their competence, professionalism and commitment/motivation. Continue along the workgroup’s learning path of the laboratory; continue along the path of awareness towards ever greater intentionality.

c. - How things went and results achieved: The detailed analysis of the Configurator in sub-groups takes place in the following conditions: a suitable setting, the use of three artefacts: conceptual map of the first meeting (on ‘what you do when you write a contract announcement?’), printouts of the most important Configurator screens, given to each group, and a poster visible to all participants which depicts the map of the entire Configurator. The task assigned: for each screen define what you consider important, what is wrong, what you would add, what you don’t understand (interface accountability). For the “don’t understands” the designers were deliberately kept away so that they did not respond in place of the users. The facilitator was always present for the sub-group and he provided support on method if he saw that a group was in difficulties. The participants were very active, free to express themselves in-group and in the plenary session. Without censorship. Everybody participated.

PHOTO 3: Group analyses Configurator screens PHOTO 4: The group orienting with screens map
VIDEO 1: Screen analysis activity (video with faces: http://www.etour.tn.it/mcis2008/video1.html)

d. - Learning: The participants immersed themselves in the role of analysts of the user interface. Diego did not expect this and it came as a revelation to find that in this group there were also computer specialists, something that was not clear at first. It is therefore essential to verify, understand and make use of the competences present in the group and to enhance them through involvement in a task. This observation conditioned the next project.

more b. - SPD attention points: All of this is positive and unconditional acceptance. The route was marked out by participants, not by Diego. Empathy: continuous listening, no forcing, use of the participants’ point of view. The participants felt enriched in their professionality (and this will have future consequences, also in their mode of participating, of being involved and present). Importance of selecting the group so that it is truly a “special” group.

Stretches of dialogue, phrases, situations: The group already reveals the differences between the work methods envisaged by the software and the need for data comprised in real work practices.
Second group: We don’t understand why the document folder is called that, because we think it is the menu…. As for the label, we would call it “denomination” not “description”. The item “personal data” is no good in our opinion, because it brings to mind something different from what it is in reality. The software was produced by computer specialists, using the framework of the regional computer system, a framework which works on the concept of document (which is why there is an inappropriate/different use of the word “document”).

6. Presentation of the laboratory results to the client

a. - The objectives of the activity and expected results: To present the results of the laboratory and agree with the client on how to continue (the client did not participate, on request by Diego, so that the participants could work more freely; it is therefore important to provide very rich feedback).
b. - SPD attention points: Ensure that the client was aware that the restitution was transparent and accurate, without interpretation or manipulation by the facilitator. Keeping him linked to each step of the work, consolidating the relationship with him and his faith in the project, giving proof that his hierarchal role and functions have always been considered, as well as his states of mind (anxiety, fear, resistance).

c. - How things went and results achieved: Present at the meeting were Diego, client and several members of the work group especially chosen (assertive, authoritative persons highly motivated to change, with important roles within the agency for change itself), because these persons would be guaranteed a true and not interpreted restitution. Various video clips taken during the laboratory were also used. Diego presented the results of the laboratory using a wide range of multimedia documentation with photographs and videos on comments by the sub-groups on each interface.

VIDEO 2: A clip from the multimedia presentation used to present the results of the work (http://www.etour.tn.it/mcis2008/video2.html)

continues c. - How things went and results achieved: The comments are made by the representatives of the workgroup, with important statements about the differences between actual work practices in designing a contract announcement and how this work could be supported by a computerised system.

Stretches of dialogue, phrases, situations: Statements of the type: “We saw that the Configurator interface contains elements that are not very clear; furthermore the interface does not help those using it to understand how it is used. In some cases the Configurator imposes a work logic that is very different from reality. Analysing the displays it was not easy to understand the operational context of each screen, the type of function performed by the screen”.

more c. - How things went and results achieved: The client, having seen the work carried out by the group and being convinced of the potential of the group, requested the facilitator to organise another laboratory to redesign a new Configurator. For the first time we were given the task of designers.

d. - Learning: The surprising thing that happened was that the participants showed great willingness to involve themselves, to be the protagonists of change (something generally not common among civil servants). Communicating to the manager/client what was happening in the laboratories proved to be crucial. From the moment when the client no longer came to the classroom, Diego always sent him the minutes of meetings within 48 hours and put the videos online so that the client could download them. The client was very impressed by this method of communicating, so rich in materials gathered in the field. This legitimated the method that we had selected. The client “trusted” the group, the method and the facilitator, so much so that he proposed something that was absolutely revolutionary and delicate/risky for the organisation. The client’s words “redesign it” expressed his acceptance of a new challenge, which would lead him to agree with IT management to discard the work done thus far (months of work) and rebuild everything using a new logic.

PHOTO 5: Paper user interface with buttons, labels, navigation menu

PHOTO 6: Image shown during work presentation

Proceedings ECIS 2009
7. Laboratory: construction of the Configurator interface in sub-groups

a. - The objectives of the activity and expected results: Construct the Configurator user interface using an agile approach (paper models). The objective is to construct with the group a new work scenario with work practices respectful of the past but able to exploit the potential of the new instrument.

b. - SPD attention points: The map of work practices constructed March 18 was used, from the vision of the participants (positive and unconditional acceptance), to incorporate these practices in a new artefact (Configurator paper interface) which this group built by themselves. Consolidate within the group the conviction that they are truly taken into consideration, beyond simple demagogy.

c. - How things went and results achieved: Work in sub-groups on construction of the Configurator with paper models, passing through three stages. The first stage led to the construction of posters using post-it notes to imagine what the new system should give to and ask of its users, in terms of data.

VIDEO 3: Brief sequence of the design activity with post-its
(http://www.etour.tn.it/mcis2008/video3.html)

continues c. - How things went and results achieved: Second stage concentrated on the sequence in which interaction with the system takes place (that is, what comes first and what comes later). In the third phase they worked to transform these into user interface elements (buttons, labels, menus etc.).

VIDEO 4: Brief sequence of sharing by the sub-group work in an assembly
(http://www.etour.tn.it/mcis2008/video4.html)

more c. - How things went and results achieved: After the sharing they agreed to create a summary of the work by building a new artefact (a map with the stages of work and results achieved, each node of the map a video or an artefact with notes on what was learned). One of the richest summary phases, here the work practices of the first day were re-discussed and redefined within the new instrument.

more b. - SPD attention points: A great deal of sharing and mutual listening, respect for the interventions of others. They discussed work process and sense making of the technical.

more c. - How things went and results achieved: Each person made a great effort to accept the contributions of the others. At the end of the session the group produced a large poster identifying all the interface elements positioned in the order that they wanted. In a somewhat unstructured debate they mixed aspects of what and how needed to be done. The facilitator suggested shaping the discussion through the shared construction of a poster depicting the system with its inputs and outputs.

PHOTO 7: Poster with the final interface design
PHOTO 8: Nodes on map connects to videos
d. - Learning: Again the key role of artefact construction. The crucial role of a facilitator able to accompany the group understanding its needs (for autonomy, but also being directive when the work proceeded in a disorderly and time-wasting manner).

8. Presentation of the results (4 August) to the client and RTI (Regional Training Institute)

(Presentation of all the work carried out in the laboratory, the methodology adopted and the results achieved. Obtain the go-ahead to continue the relationship with new projects even of a different kind. There is a risk that after having activated participation, involvement, change, it may disintegrate and be lost, backfiring both against the client who gave his trust and against the facilitator).

6. LESSONS LEARNED/REFLECTIONS ABOUT THE WHOLE PATH

Section 5 shows how our method made it possible for the client group to go beyond the initial project conceived by the manager, and even, later, by the facilitator. ‘Go beyond’ in the sense that not only were the users of the system able to enrich the instrument with new functionalities derived from a more detailed analysis of the work to do in its various dimensions, levels, functions, but they were also able to find the right language to voice their need for an instrument with a different philosophy which then emerged from their work, and which therefore went beyond what the designers had thought, and brought to light new potential and logic for the instrument, which neither the designers nor manager had thought of: an entirely new goal and outcome, an experience of which they became fully aware, and that they also later engaged to divulgate and preach to others.

What we learn above all is that, to ensure that making the client fully aware of what PD can do, is an objective to be pursued during the whole path; it is a goal that requires experimentation of the method. The client must be enabled to create prudent conditions so that the method can be practised, and these prudent conditions create space for the pursuit of smaller, more pragmatic objectives without forcing the client’s hand, or pushing him to do things prematurely (respect Kairos, the appropriate time).

We should remember that the objectives stated at the outset were only to analyse the Configurator already designed by designers, obtain feedback from its future users, construct a vademecum, and plan a training course for a larger number of users. These objectives of course fall short of a fully ‘blossomed’ Participatory Design approach. In reality the possibility to undertake genuine participation – not an easy task (Bodker et al. 2002) - in the design of the Configurator, and with it to redefine the work practices, was understood and granted by the client and the workgroup during the ongoing work, on the basis of results achieved, and after continuous interaction/exchange of information and thoughts between the s, client, and group (a path of reciprocal learning, of awareness). The group legitimised itself by doing: it was not evident at the start that this would occur.

Let’s list relevant practices towards this important result:

• To GIVE SPACE, HEED, AND ROLES to all involved, not only the manager, but also users, technicians, other managers, etc., with their different/opposing points of view, needs and interests.
• Of LISTENING, OBSERVATION, AWARENESS, REFLECTION, ELICITATION by the facilitator of the needs of the client and the different actors involved. Needs not decided upon once and for all but continually reviewed, re-discussed, in a process of learning/awareness making the client and group ever more involved, increasingly active.
• Of the client’s FAITH in the facilitator, because this allows the client to trust the facilitator, to feel accepted and respected, as well as protected.
• Of the ‘s CONSISTENCY in his relationship with the client (be authentic, coherent between the verbal and the non verbal, be convincing, assertive).
• Of ’s TRUST in the client, this generates positive unconditional acceptance, empathy.

Here is perhaps the key to understanding the entire process according to the logic highlighted in our AR: the stakeholders involved here, all have different perspectives, and interests. Some are interested in the PD approach because of their desire to use it in other cases (the manager of one of the computer systems divisions), some are interested in technology tout court, some in defining new work scenarios.
The employees in our group are not simply application software users: they are stakeholders expert in the design of call for tenders and their redesign started from work practices. On the other hand, the Configurator was previously designed by technicians used to design starting from abstractions, working on the possibility of processing the data according to a computational logic, a logic different from that of routine work. In the last case (technical design): $\rightarrow$ABCDE (from abstraction to work practices). In the second case (redesign participants in laboratories) $\leftarrow$EDCBA (from work practices, with ramifications, to abstraction).

Our SPD approach was successful in managing ambiguity and establishing trust towards a new governance, where users could in fact redesign system and work practices according to their own user-strategies, with the consent of all involved, all recognising the improvement of this state of affairs.

We judge the quality of the SPD approach by three requirements (Baskerville and Myers 2004): a contribution to practice (the action), a contribution to research (the theory), the criteria by which to judge the research, and we show explicitly how the research in the case meets these criteria: Contribution to practice. In the ‘Online call for bids’ project, users not only to enriched the instrument with new functionalities, they were also able to voice their need for, and designed, an instrument with a different philosophy which went beyond what the designers had thought.

Contribution to theory. Confirmation of the usefulness of:
- recursive dialogic process for sense making and the involvement of company personnel
- person centred counselling-like facilitation qualities for trust building and empowerment

Evaluation criteria for AR results. Hi degree of satisfaction of company manager, personnel, and facilitators/researchers, respectively, for change produced and final results.

5 CONCLUSION

In conclusion, we have experimented with user design laboratories based on the Interactive Use Case technique of PD, and on open conversations and learning, as ways of practicing SPD facilitation towards co-construction in a regional administration setting. These activities clearly enabled and supported participation in engaging, in sense making and awareness, and in the acquisition of perspectives of others. People generally felt comfortable and not at risk of being judged. Participants expressed how important the experience of working creatively on solving “real problems” had been for them. We can understand this also as a result of the in-depth observation and intervention character of our SPD engagement with people in the project, which provided us with good knowledge about their work practices, potentials and problems on the one hand, and allowed trust building on the other.

In closing, let is recall the basic attention points anticipated all along, to close the circle.

5.1 The group of facilitators prepares with care the activities in the field, in order to practice SPD indications – especially those from Rogers - in a real, ethic, coherent, manner.

Before each meeting there is always a consultancy workgroup session during which the team analyses the materials collected in the previous meeting, the observations made, the exchanges and the dialogues, content aspects, relational aspects, cross referencing all with first-off impressions, the notes taken by the s in a “log book” during and after the meeting. The primary aim of these sessions is to analyse the work performed “on stage”, reprocessing it through analysis of the observation materials and discussion among the s (both those who went directly into the field and those who remained back stage), to assess what happened in the previous step, the situations experienced, the content and relational aspects, the initial impressions/emotions, the purpose being to design and define the next step in light of the path taken thus far, the results achieved and the lessons emerging from the analysis. Planning means: thinking flexibly about what will be done later, who will do it, with what aims, strategies, communicative/relational concerns etc. For each of these consultancy team meetings a log-book was kept, an internal workgroup document which was used to keep accurate track of these steps. It was from these notes taken in the field and during the consultancy team meetings that the points developed for each step were taken.
5.2 They closely attend SPD attention points during intervention steps, opening the way to success

Facilitators pay attention to the relation of the facilitator with all concerned (managers, employees).

- **Incipit**: at the beginning of the relation, dedicating time to gathering the needs of all concerned through SPD attentions, i.e., facilitating the development in of awareness in individuals and groups
- **Ruit**: in all crucial passages of the project, activating and mixing individual and group meetings
- **Exit**: in approaching conclusion, activating occasions for visible, tangible results, as feasible with that group, with those people, in that context, giving sense to and justifying the entire path.

This means communication attentions (language, non verbal aspects, setting), observation, listening, both in managing meeting with participants, and facilitating internal relations in the organisation.

This means consolidating the trust relation with the manager through open communication and choice of appropriate language, unconditional positive acceptance, empathy.

This means, in laboratory activities, attention to: whom to involve in the work activities (caring for roles, people, competences, skills) → key persons; whom should be, or not be, present w.r.t. hierarchy (e.g., manager causing self-censorship on others); subgroup composition; setting and climate (a true laboratory, protected, non judgemental); role of the facilitator (observation and coordinator); communication management in the group; choice of job assigned (starting from own work practice, describing and comparing them among peers: this helps involve and motivate people, reassures them on their abilities, makes them learn new things in a protected environment: a job, in sum, of designing something useful for real, close to participants needs).

Lab activities strengthened the group, favoured contributions by individuals, and favoured training of all. In this environment, SPD behaviour of facilitators were winning bets: empathy, acceptance, congruence. The group felt guided, protected, non pushed, respected in its time and way (kayros). Its work and commitment was valued, yielded outcome. No demagogy or exploitation, real involvement, true trust in the group and in its potential. All this consolidated trust in facilitators, and in the project, increasing motivation.

References


Bion, W.R. (1961) Experiences in Groups and other Papers, Tavistock Publications Ltd.


Rogers C.R. (1951) Client-Centered Therapy, Houghton Mifflin Company, Boston


DESIGN CRITERIA FOR PUBLIC E-SERVICES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0659.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-government, E-Service, Human computer interaction (HCI), Usability</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
DESIGN CRITERIA FOR PUBLIC E-SERVICES

Annie Röstlinger, Linköping University, Dept of Management and Engineering, 581 83 Linköping, Sweden, annie.rostlinger@liu.se
Stefan Cronholm, Linköping University, Dept of Management and Engineering, 581 83 Linköping, Sweden & University of Borås, Business and Informatics, 501 90 Borås, Sweden, stefan.cronholm@liu.se

Abstract

This paper focuses on design criteria for public e-services. The paper challenges traditional lists of usability criteria supporting design and evaluation of IT-systems. Traditional criteria within the field of human-computer interaction are perceived as too limited and an expanded view of the communication and business process is proposed. Within this expanded view, existing usability criteria are considered as important but not sufficient. Based on this claim, criteria also supporting communication and business actions are proposed. The proposed criteria are generated from an empirical study concerning a development of an e-service. The aim of this e-service is to support the communication between companies and the municipality when companies are going to get permission for running business. Analysing the communication need in this context has been of particular interest in order to generate complementary criteria for design and evaluation. Examples of criteria generated are: good conditions for users to become knowledgeable (support for perception, comprehension and information), differentiate between informative and performative parts, provide meta information (purpose and content).

Keywords: Usability, human computer interaction, e-government, e-service
INTRODUCTION

In the e-government field there is a rapid growth in the development of e-services (Anca\-rini, 2005; Buckley 2003). Such an e-service is a public service mediated electronically through a user interface that is generally available. Public e-services can be used by citizens or company. These users can be inexperienced and infrequent users of the public e-service. If e-services are to be used, they must be easy to use and be beneficial to the user. Important questions are: How to create usable e-services and what properties are desirable for such IT-systems?

There exist several lists of usability design criteria. Examples of criteria lists are: 10 usability heuristics (Nielsen, 1993), Eight Golden Rules (Shneiderman, 1998), design guidelines for small screen devices (Kärkkäinen & Laarni, 2002), context-aware mobile applications (Häkkilä & Mäntyjärvi, 2006) and Participatory Heuristic Evaluation (Muller et al, 1998). These lists are explicitly not web oriented; however, Nielsen (1999) has presented criteria for designing web usability.

In our literature research we could not find any list belonging to the context of public e-services. The importance of identifying different contexts is acknowledged by Henninger et al. (1995). They claim that “If the potential of usability guidelines as an interface design technique is to be fully realized, they need to be augmented with context-specific guidelines and examples that synthesize isolated guidelines into domain-specific solutions to design problems”.

A reasonable question to ask is “why are not existing criteria good enough?”. Our answer to this question is that criteria are always derived from a specific theory or perspective. That, which means that the usage of criteria lists support a specific focus of the phenomenon studied. To set a focus on something also means that other aspects are de-emphasized. The identified criteria lists are constructed in the field of human-computer interaction and their aim is to focus on the aspect of the interaction between a human and a computer. Our hypothesis is that the existing usability criteria can be too limited for supporting design and evaluation of the context of e-government. The reason for this claim is that we view e-government as a communication intensive context. In this context there are citizens or companies that need to communicate with e.g. representatives of municipalities or other authorities. When citizens communicate with municipalities they are often instructed to use artefacts such as answering machines or the internet. It is not an exaggeration to say that communication via traditional face-to-face situations is successively replaced or complemented by communication via information technology (e-services). Based on this we claim that there is a need for a complementary perspective (and design criteria) focusing on the e-service mediated communication between humans instead of a more limited human-computer interaction (see section 3). Our hypothesis is supported by Hartson et al (2001) who claim that “Usability is seated in the interaction design”. Our hypothesis is also supported by findings in Cronholm & Bruno (2008). A result from this study is that “most of the existing usability criteria are predominantly supporting a human-computer interaction perspective”. Furthermore, Gould and Lewis (1985) claim that existing design guidelines are limited since they are not detailed enough. We are not saying that the criteria lists referred to above are unusable; on the contrary they seem to provide good support for analyzing human-computer interaction, but they are not mainly supporting design of IT-mediated communication between humans.

Many of the existing criteria lists seem to be constructed for usage in any situation (e.g. Nielsen, 1993; Shneiderman, 1998). The risk of using too general criteria is that these do not provide enough contextual support that is needed in a specific domain. They can be perceived as too superficial and thereby not usable enough.

The aim of this paper is to propose a set of design criteria primarily to be used to support professional designers and evaluators within an e-government context. In this paper we restrict the discussion of criteria to design criteria for public e-services. This is the context of our case study described below. From this context we have found the traditional design criteria not enough supportive for our design decisions concerning communication between the municipality and the companies using the e-service.
The proposal of criteria is based both on existing criteria lists (see section 4) and on new experiences from the case study (see section 5). The introductory section is followed by a description of the research method (section 2). Section 3 describes the theoretical basis used and in section 5 we present the design criteria used in the case study. Finally, in section 6 we present some conclusions.

2 RESEARCH APPROACH

This study is part of a research project for improving the conditions of commercial and industrial life in municipalities. The project is a joint project between four municipalities and some researchers. In order to develop design criteria for e-services we have used three different knowledge or data sources: a case study, existing usability criteria lists and theories of communication and business processes. In the case study we acted as both researchers and designers of an e-service. The research process can therefore be characterized as action research since we intervened in the process (e.g. Lewin, 1946; Avison et al, 2001; Baskerville, 2001).

The case study has functioned as a data source for inducing design criteria. The overarching questions asked were “How can companies and municipalities benefit from an e-service?”, “How can companies and municipalities communicate with each other through an e-service?” and “How can companies interact with an e-service?”. The study started with a broad process analysis discovering problems and goals concerning the interaction between the companies and the municipalities. Based on this investigation a decision was taken to start developing e-services and redesigning processes. A critical analysis of forms was conducted. It was important to critically examine existing forms, e.g. content, structure and terminology. This was done in order to avoid transferring bad existing paper forms into e-forms.

We have used existing usability criteria as a source of inspiration and have reused or refined relevant criteria. The design criteria generated were also explicit compared to three existing usability criteria lists. The reason for choosing two of them was that they are well known. These lists are: 10 heuristics (Nielsen, 1993) and Eight Golden Rules (Shneiderman, 1998). The third list is called Actability Principles (Cronholm & Goldkuhl 2002). This list was chosen since it is based on a communicative perspective. It is more comprehensive but is not as commercially used as the other two lists. Behind the idea of comparing the developed design criteria with the list of Actability Principles was to understand if and how criteria generated from a communication perspective could contribute to design criteria for the domain of e-services. Finally, the communicative and business perspective has also had an impact on the development of the design criteria. We have continuously refined the proposals of design criteria and the e-service, which means that there has been a shift between action and reflection/evaluation. Through this the design criteria have emerged in an iterative way.

3 THEORETICAL BASIS: A COMMUNICATIVE PERSPECTIVE

In this paper we apply communicative and business perspective in order to develop design criteria for e-service. The theoretical sources for viewing IT-system as communication and business systems are social action theory (e.g. Weber, 1978) and language action theory (e.g. Goldkuhl & Lyytinen, 1982; Habermas, 1985; Searle, 1969; Winograd & Flores, 1986). The main message in Weber’s (1978) theory of social action is that communication is intentional. Using a social action perspective means that it is not acceptable to view IT-systems as a black box with some social and organizational consequences (Dietz, 2001). IT-systems are instruments for business action.

Language action theory discusses communication as one type of action, and IT-systems are not considered as “containers of facts” or “instruments for information transmission” (Goldkuhl and Ågerfalk 2002). The language action perspective emphasizes that communication is to establish interpersonal relationships between the sender and the receiver (Searle 1969). The communicative perspective emphasizes a user communicating with other users through a user interface. It is a human-
via-computer-to-human communication; it is not a limited human-computer interaction. Thereby, the use of IT-system is viewed as a social process consisting of technology mediated business communication.

In order to present our view on interaction we use the Elementary InterAction Loop (EIAL). EIAL was originally introduced in Ågerfalk et al (1999) and revised in Goldkuhl et al (2004). EIAL consists of an interactive situation divided into four phases: informing, execution, IT-system reaction and interpretation (see figure 1). The informing phase means that a business actor interprets the action repertoire offered and possible messages from other business actors in order to reach a decision about what to do. The execution phase describes that the user is performing the action chosen. The IT-system reaction phase describes the IT-system’s response to the business actor’s action. Finally, the interpretation phase describes how the business actor interprets the result of the IT-system’s reaction.

In the middle of the EIAL a screen document is placed. A screen document is any designed interface with which the user interacts (e.g. a form on a web page). The screen document plays different roles in the different phases. Therefore the screen document is multifunctional. In the informing phase the screen document is used when the user is reading the screen to figure out what to do. It contains information about the action possibilities and other action conditions. In the next phase the screen document is used for execution. In this sense, the screen document functions as an action medium. For example, the user enters some data in a field and clicks on a button on the screen in order to perform an action. Finally, the phase of the IT-system reaction should be understood as a response to the user execution. The IT-system’s reaction can result in changes of the screen document (as a feedback to the user). In this sense, the screen document consists of action results and functions as a basis for interpretation.

"I did" (Execution)
"What can I do?" (Informing)
"What has happened?" (Interpretation)

Figure 1. Interaction: The Elementary InterAction Loop (based on Goldkuhl et al., 2004)

The central concepts of interaction, communication and business process discussed above are depicted as three sets or three levels (see figure 2). As pictured, an interaction is the interplay between a user and an IT-System. The communication takes place between two users, mediated by an IT-system. The arrows in the figure above symbolize the business; a business that produces something (a product or a service) for a client. The communication level is viewed as a subset of the business process level and the interaction level is viewed as a subset of the communication level. This means that if the IT-system is considered as providing good support for the business level it also provides good support for the lower communication level and interaction level. On the other hand, if an IT-system is evaluated at the interaction level no predictions can be made of the higher communication level and business process level.

The aim of figure 2 is to bring forward three important and related levels. We are not saying that criteria formulated on the interaction level are useless; ideally we would like to see criteria residing on all the three different levels and that they are coherent and complementary. The model is chosen since e-services are supposed to support a type of business that is highly communication intensive (see section 5.1).
In order to motivate our proposal of design criteria for e-services we have analyzed two familiar and popular usability criteria lists; the Ten Heuristics proposed by Nielsen (1993) and the Eight Golden Rules proposed by Shneiderman (1998). We have also analyzed a criteria list that is not equally well known, this list is called Actability Principles (Cronholm & Goldkuhl, 2002, 2005).

An overarching result from studying these criteria lists is that we have found them important to consider while designing IT-systems; but they are not sufficient. Analyzed from a communicative perspective they are viewed as too limited, i.e. they are supporting the interaction that is going on between a user and a computer. Our claim is that they do not fully support communication between humans and they are not formulated to support business processes. Examples of criteria representing this limited view are “Visibility of system Status”, (Nielsen, 1993), “Enable frequent users to use short cuts” (Shneiderman, 1999) and “Clear feedback” (Cronholm & Goldkuhl, 2002, 2005).

Another result from our analysis is that we find several criteria hard to understand. The problem of comprehensibility is often to do with that: 1) the criteria are not categorized 2) the label of the criterion does not correspond to the adjacent explanatory text and 3) the label and the description of one proposed criterion often consist of several criteria. The criteria proposed by Nielsen (1993), Shneiderman (1999) and Cronholm & Goldkuhl (2002, 2005) are presented as flat uncategorized lists; i.e. the criteria are considered as belonging to the same abstraction level. The advantage and role of a multilevel abstraction hierarchy is discussed in Rasmussen et al (1994). They compare a multilevel abstraction hierarchy with a means-end hierarchy and claims that a multilevel abstraction hierarchy is often used in practical problem-solving processes.

The second problem, the correspondence between the label and the adjacent explanatory text, is best illustrated by an example chosen from Shneiderman (1998). The label of the criteria reads: “Offer informative feedback”. The explanatory text reads: “For every operator action, there should be some system feedback. For frequent and minor actions, the response can be modest, while for infrequent and major actions, the response should be more substantial”. The keyword in the formulation of the label is “informative feedback”. The explanatory text is discussing when and how much feedback that should be provided. There is no explanation or recommendation of how to provide informative feedback.

Another example of the lack of correspondence between the label and the explanatory text can be found in the criterion “Flexibility and efficiency of use”, proposed by Nielsen (1993). In the explanatory text Nielsen (1993) mainly discusses accelerators. We conceive it to be a too large abstraction gap between “flexibility and efficiency of use” and “accelerators”. We did not expect to find a discussion about accelerators in order to define flexibility and efficiency of use.

The third problem, that the label and the description of one proposed criterion actually consists of several criteria, contributes to the confusion of what is going to be evaluated. Several formulations of labels refer to more than one criterion. To clarify our criticism we use the same example again:
5 PROPOSAL OF NEW DESIGN CRITERIA

5.1 Prerequisites for the case study

A company may need different permissions from the local authority to set up and run a business. Permissions are often necessary e.g. when selling food, use chemicals or building a factory. To get permission easily and quickly is important for the company since permission is a prerequisite for running a business in a legal way. We have identified five main phases in the permission process: 1) Search info about application [company] 2) Formulate and send application [company] 3) Assess application [municipality] 4) Decision [municipality] and 5) Send decision [municipality]. This process is easy to conceptualise but different problems may occur when companies are going to apply, e.g.:

- Companies do not have sufficient knowledge about: 1. The legal rules; e.g. what are we allowed to do and not to do, is permission needed or not for different kinds of businesses? 2. The process of application, assessment and decision organised by the municipality; e.g. who is the executing officer, is the application supported by a form, what aspects are assessed, what is the fee, when do we get a decision? 3. How to get answers to questions; who is the right person to contact, by telephone, e-mail or by visit, how to find understandable information about regulation and permission process and application forms?

- Executing officers at the municipality are overloaded with rather simple questions, they have to handle incomplete and difficult-to-read applications, and they need to contact the companies to get completed applications.

The companies regarded the municipality as too bureaucratic and not service-minded enough while the municipality regarded the companies as ignorant and unwilling to follow legal rules. The identified problems are all located in the communication process between the municipality and the companies and also have implications for the conditions in the business process. The consequences are both inefficient municipality processes and inefficient and delayed business processes.

The four studied municipalities all had existing web sites but these were not informative enough. The web sites did not seem primarily to be designed to support the companies’ application processes. Some municipality web sites had reading functionality and application forms in pdf-files but none had functionality for sending information, e.g. questions and applications. Many of the application forms were designed to support the case handlers’ assessment processes rather than the companies’ process to produce a complete application. There were no proper e-services within this domain.

To improve the communication between the case handlers and the companies we started a pilot study to design e-services for firms active in food business. The first step in the study was to support the communication in the first two phases of the permission process: companies searching information about food permission and then sending their application to the municipality.

5.2 Designing the e-service

The pilot study has resulted in a web-based IT prototype. In the design process we have actively used a prototyping approach (Vonk, 1990). The e-service prototype holds information relevant for the application process and companies going to apply, e.g. information about legal rules, municipality principles for handling applications and decision making, information about the subsequent municipality control of the company handling food in a proper way. The e-service also includes
different supporting functions for applying: 1) an application form to fill out by companies 2) guides with proper questions depending on type of food business 3) completion control of forms 4) function for sending an application. Some of the food case handlers participated in the design process together with the researchers. The aim of the e-service was to support both companies and case handlers. Besides making it easier for companies, the aim was to ensure that case handlers achieved correct and necessary information.

Knowledge from the workpractice was an important basis when we created the e-services prototype but also practical theory including usability criteria and actability principles was important. The aim was to create an e-service easy to use for experienced and inexperienced users and frequent and infrequent users, i.e. all companies trading in food. The prototype was refined in an iterative design process until we got a satisfied e-services prototype. In this refinement process the researchers and the case handlers reviewed the prototype based on their different views on “what is meant by useful”.

When we demonstrated the prototype, also case handlers from other permission areas were interested. We decided to transfer properties from the food prototype to a new e-service prototype for building permission. In order to detect characteristic and important properties we examined the prototype ex-post and also reconstructed our design decisions. We found several characteristics of importance for success when using the e-service in the permission process. Then, when we compared the detected attributes to the existing design criteria we found it difficult to relate each of our attributes to any particular criterion from the criteria lists. We had achieved an e-service which met the case handlers’ and researchers’ aims and expectations, but it was not because of the existing criteria. As mentioned above, we found the existing criteria important and valid but limited in terms of support to the designer.

5.3 Different aspects of an e-service

A public e-service is about two actors communicating. A company needs information from the municipality in order to make decisions about the business and the application, i.e. the company needs to be informed by the municipality and the municipality needs to inform the company. The company needs to send an application to the municipality to get a decision from the municipality; i.e. the company needs to inform the municipality and the municipality needs to be informed about the application from the company. The company is both a sender and a receiver and these are also the roles of the municipality in the communicating process.

A public e-service is also about two organisations conducting their respective tasks. One important task for a municipality is to give permission to firms doing business, and for the company it is important to get permission in order to produce results, e.g. selling food in a shop or in a restaurant. This means that the e-service developed is created to support the producing processes in two different types of workpractices, one producing food permissions (a municipality) and the other producing food to customers.

Of course, public e-service is also about a human actor interacting with a computer. Handling the computer in an efficient way is important and a prerequisite for achieving communication between actors via the computer. A good working communication between the sender and the receiver contributes to results in the workpractices. We have identified three levels of e-services (see figure 3) where the IT-system plays an important role. The three levels concern different aspects of the workpractice and also focus on different aspects of the IT-system. Thus, to get an e-service that is useful and supports producing results in workpractices we have to consider the three levels of IT when designing e-services. As mentioned above, we found the existing criteria in the lists mostly related to the interaction level. A public e-service is not only about human-computer interaction. It is also about human to human communication via computer and furthermore about humans (with IT support) producing results in favour of humans.
When designing e-services it is important to recognise all three levels of human and IT-system actions and the e-service must be designed to support these levels. Thus, when designing e-services for public municipalities we also need criteria related to these levels, not only to the interaction level.

![Figure 3. Three levels of e-services in organisations](image)

5.4 Criteria for usable e-services

Every item carrying information intended for the user, i.e. words, pictures and sounds, should also be **perceived** by the user. The user should quickly and easily be able to detect adequate information without risk to miss something. Important features for achieving this goal are e.g. colour, shape, size/volume and location. All these properties may have values which make them highly perceivable or not perceivable at all. The perception differs between people, e.g. different people perceive colours in different ways. It is also important to be careful with hidden and concealed functionality.

Every sign perceived should also be **comprehended** by the user. The user interprets the perceived signs to understand its meaning. The user should easily and quickly be able to understand the information without risk of misunderstanding. The key features are e.g. simplicity, recognition and consistency. - Simple but informative expressions and symbols. - Signs related to the task and the users’ knowledge and understanding. Important aspects are e.g. the language and vocabulary used, signs related to standards, local standard in the IT-system and global standard in other IT-systems, other well known artefacts and other objects. - Avoid homonyms and be careful with synonyms. To perceive information is a prerequisite for achieving the goal of comprehension. The different properties for perception (colour, shape, size/volume and location) also have functions with importance for achieving simplicity, recognition and consistency.

The user should not only be able to perceive and understand; the user should also be **informed** by the signs. The information perceived and comprehended should make a difference; the user should become **knowledgeable** and prepared for action. That the receiver is informed by the sender is an aim of communication via e-service. Important features contributing to knowledge when the user perceive and comprehend the information are: information related to the user and to the task, the action which the user is going to perform or has performed, the process in which the information is a part of. In this context, properties as structure, thematizing, level of detail (detail, aggregation and summary) and level of knowledge (elementary, advanced) are important.

Perception, comprehension and informing are all basic aspects of importance for the IT-system and its use. These aspects must be taken into account in the designing process, whether the information is about a subject like food permission, navigation in the IT-system, feedback information or error.
messages. In this way these basic aspects can be related to all three levels of e-services in organisations (in figure 3),

What is a user going to do with the acquired knowledge? When the user becomes knowledgeable and prepared, the action taken can be: a) within the IT-system, b) via the IT-system or c) outside the IT-system. To navigate to the next page in the web site is an action within the IT-system (a) on the interaction level. To read information initiated by the case handler about e.g. legal rules and permission is instead an action via the IT-system (b) on the communication level. To send a question to a case handler is also an action via the IT-system (b) on the communication level. It is important to recognize that when a user communicates with a case handler through e-service the user is at the same time interacting with the computer. Thus, this action is multifunctional. Action multifunctionality builds on and extends the concept of pragmatic duality (Sjöström & Goldkuhl, 2004). To send a complete application to the case handler is an action via the IT-system (b) but on the business process level. An action via an IT-system on the business process level always also includes both the interaction and the communication levels. To have obtained food permission through the e-service and then running a restaurant is also an action on the business process level but outside the IT-system (c). Actions outside the IT-system are separated from the IT-system and these actions are restricted to the business process level only. The three different levels (interaction, communication, business process) are related in a way that all the lower levels are prerequisites for all the higher levels (see figure 3). Lack of quality on a lower level affects the quality on higher levels.

Informing and performing are different e-service functions. In a user interface every object is information. But the information objects can have different functions: to be used to inform or to be used to perform. Through the informing function the user is being informed about something. Every object has some informing function. But some objects also have a performing function which allows the user to perform something and that means more than being informed. There is a big difference between the user only getting information about something, e.g. how to pay, and the user performing a payment. In an IT-system it is highly important to clearly differentiate between these two functions. The user has to be aware of what is going on. Only receiving information is often not associated with heavy consequences. However, to conduct by sending information (e.g. a payment, an application, an update or a deletion) can be more serious. In the e-service for food permission we clearly differentiated between informative e-services and performative e-services. By the informative e-services the user can get information about e.g. legal rules and the application process. By the performative e-service the user can perform the application process, i.e. parts of the process with filling out the form and sending the complete application to the municipality. The IT-system for food application permits every user to read information (an informative e-service) but only logged in users are permitted to perform an application (a performative e-service).

**Purpose and content;** i.e. meta information about the main topic. What is the user offered by the e-service? Who benefits from the use of the e-service? It is important that a user as soon as possible is able to get an overview of what the system can give; what information to receive and what information to send? The user should have a basis to decide, “is this a ‘right’ site for me or should I leave the web site or the current page?” The function of this information is also to stimulate interest among prospective users. For example, on top of the first page on the web site for food e-service there is an overview picture and a synoptic description of purpose and content and the expected type of user. The description of purpose and content continues on the top of all other pages, where the user can find a description or a title that indicates what the page is about.

Clearly expressed main topic; the informative and/or performative e-service on the communication and the business process level. This is the content in the communication between a sender and a receiver, and also the purpose of using the e-service. The structure and the context of presenting the main topic are very important. The information should be displayed from the users’ point of view and the purpose of using the e-service (Eliasson & Hedström, 2005). In the example from the food e-service we related to the application process and structured the information in 1) information about legal rules 2) information about the application and 3) information about permission process. This
information is displayed for the user in both an aggregate and a detailed way. The user is able to receive information about the application process but also to conduct the process via the e-service. Filling in the form is structured according to different generic themes (who, what, where, when) related to the business process. The application process begins with a clear starting point and ends with a clear ending point.

**Navigation.** Information about principles and functions for information access and also information that allows performing navigations. Where and how can the user find the information in the IT-system? The design of the navigation has consequences for accessing the informative and performative e-services. The navigation can be open or limited. Open is when the user has the option to choose among pages and functionalities. Limited is when the user does not have any choice but the IT-system chooses the path within the IT-system. A fully open navigation is hard to realise; various mixtures of open and limited are more realistic. In an open navigation the user has flexibility to perform different actions, but in a limited navigation the IT designer may choose a proper way according to the user and/or the task. In the food e-service we have chosen a design to make the user knowledgeable before conducting an application. The user will be guided to read information before performing an application. The user has also the possibility to navigate directly to pages for performing an application (aimed for more advanced users). The pages for filling out the form can be accessed in any order. The informative and performative e-services concerning the main topic are supported by functions as e.g. forward, feedback, redo and rectify.

**Indicate forward.** The user should be able to navigate and act in a confident way within and through the IT-system. The IT-system must contain clear information that indicates what result the user can expect by using the system. In the e-service for food permission all buttons have a descriptive sign, for instance.

**Feedback.** The system should give feedback to the user, information about what the user has done and what the IT-system has done. The IT-system can display the result, e.g. accessed information about rules for food permission, or describe the result, e.g. your application is sent to the municipality.

**Redo.** If possible, the IT-system should permit the user to change actions already done. Information about opportunities to change is imported. Sometimes changes are possible and sometimes not.

**Rectify.** Support for error handling. When and how it is possible to correct an error, (i.e. an undesirable result not changeable by redo).

When designing the e-service it is important to pay special attention to the main topic, the navigation function and the difference between the informing and performing functions. The functions of feedback, indicate forward, redo and rectify are relevant in all actions in the IT-systems. However, the most basic prerequisites when using the e-service is its support to the trajectory of perceiving, comprehending and becoming informed as a basis for action.

## 6 CONCLUSIONS

An IT-system is about information, information to be used for some purpose and with some expected effects. The purpose and effects of the e-service in our case study can be related to the different levels in the three levels model (see figure 3). We used the model as a basic approach when creating the e-service prototype and also when articulating and analysing the criteria. When a criterion was identified at one level, corresponding criteria residing on the other levels were searched for. In this way, the model has governed us to pay attention to formulations of criteria and to search for criteria residing at different levels.

The knowledge contribution of this paper consists of design criteria for e-services. We have generated new criteria, refined existing criteria and reused existing criteria. In this sense our approach can be characterized as a cumulative approach. The most important new criteria that we have identified are to create good conditions for users to become knowledgeable, differentiate between informative and
performative parts and provide meta information (purpose and content). Examples of reused criteria from the HCI-field are feedback, redo and rectify.

The special character of governmental services has influenced the evolution of these e-service criteria. A municipality’s role is both to exercise public authority and to provide service. The municipality needs to inform the companies and handle applications in a legal and proper way. Different legal rules need to be clarified and readily available. The company managers need to be knowledgeable before making an application. The municipality has regulated obligations toward companies but also the companies have regulated obligations towards the municipality. The companies are in some situations forced to get permission and then to turn to the municipality to get that permission. There are no alternative ways for the companies. These conditions have influenced the e-service criteria, e.g. knowledge support, thematized forms and division into informative and performative services.

One message in this paper is that criteria proposed by HCI scholars are important to consider, but they are not sufficient. Our claim is that they are too limited and that they are not formulated to support the communication process or the business processes. They are formulated to support one individual user interacting with one computer, just restricted to human-computer interaction.

Our view is that human-computer interaction is important, but we view IT-systems as media or instruments used to support communication between users and business actions that are performed in the business process (Goldkuhl 2005). Therefore, the use of an IT-system is only a means to achieve a higher goal: to support the communication between a sender and a receiver in order to successfully manage a business process that produces value for the client. This means that we propose a changed perspective on design or evaluation of public e-services: An expansion from a more limited human-computer interaction perspective to a human-to-human communication via computer perspective.

ACKNOWLEDGEMENTS

This research has been financially supported by the Swedish Governmental Agency for Innovation Systems (VINNOVA).

REFERENCES


BITCHING, BOUNCING AND BRAWLING - HOW BACKCHANNELS BROUGHT COLOUR TO CONFERENCE CALLS

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0186.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>Case Study, Computer supported collaborative work (CSCW), Distributed organization / teams, Instant messaging</td>
</tr>
</tbody>
</table>
BITCHING, BOUNCING AND BRAWLING

HOW BACKCHANNELS BROUGHT COLOUR TO
CONFERENCE CALLS

Benjamin D Eaton, Ph.D. Student
b.d.eaton@lse.ac.uk
The London School of Economics and Political Science

Abstract

This paper addresses the research question: How do digital backchannels, mediated by Real-Time Communications, facilitate the progress of conference calls? It is argued that whilst the fields of Distributed Work and Computer Supported Collaborative Work have been studied in depth, research into Real-Time Communications (RTC) and Instant Messaging (IM) is at an early stage. There is a paucity of research into the use of these technologies in distributed settings within organisations, especially from a perspective of generating social cohesion within teams. Furthermore, there is no known research into the use of instant messaging to mediate “hidden” interactions between individuals, through digital backchannels, within conference calls. Qualitative empirical evidence is obtained from two case studies of teams using an established corporate RTC product. Goffman's Interaction Order (1983) is presented a suitable lens through which to interpret this mediated form of interaction. Empirical evidence points to concurrent working and use of digital backchannels during meetings and conference calls. Dramaturgy (Goffman 1959) is used to identify and analyse three interaction styles of “bitching”, “bouncing” and “brawling”. It is found that interactions over digital backchannels fulfil criteria of Interaction Ritual (Goffman 1967) and may contribute to the creation of social cohesion.

Keywords: Digital Backchannels, Real-Time Communication, Instant Messaging, Distributed Teams, Interaction Order, Social Cohesion
INTRODUCTION

In their research report (2008) on a Delphi study of Real-Time Communications (RTC), Frößler et al. described it as an emergent technology, which has recently come to the attention of academics. RTC represents a group of products originating from the convergence of the telecommunications and groupware markets. It consists of two main components, presence, or availability status, and unified communications (UC). Presence provides contextual information regarding the availability status of group members for engaging in communication. By indicating their availability for communication and their preferred means of being reached, individuals are able to signal when they can be engaged in communication. UC unites access to communication tools such as voice, instant messaging (IM) and e-mail in one application. As a result, UC facilitates the management and coordination of a user’s communication. To date there has been little academic research into RTC (Frößler et al. 2007) and its uses.

Companies are also adopting flexible modes of working, where teams are geographically distributed across sites. Many users of RTC form part of distributed teams involved in dispersed communications activities (Olsen et al. 2000) and are impacted by strains that such an environment might impose (Hinds et al. 2005). A common genre of communication and coordination employed by these individuals is the conference call. It has been noted (Olsen et al. 2000) that such forms of interaction are limited in context, interpersonal signals and cues when compared with face to face meetings. This may impact upon interaction and cohesion amongst teams.

RTC tools allow concurrent communication channels to be set up, using IM, during an instance of voice interaction. These secondary, private and unofficial channels of communication between two or more parties are known as Digital Backchannels and are invisible to the wider group taking part in the primary interaction (Cogdill et al. 2001). The use of backchannels may occur, for example, in the scenario of two teams on a conference call. One team may use IM to surreptitiously debate a point made by the other team, such that they can respond in a coordinated and effective manner. It is of interest to understand how conference call participants find backchannels useful, to see how backchannels may add context within calls and how they may create cohesion within teams. This paper reports on the use of backchannels occurring within audio conference calls, within two companies using an established corporate RTC product.

Whilst there has been little research into the use of RTC in the organizational sphere, there have been studies of IM. This research is mainly concerned with increasing user adoption (e.g. Aducci 2008) or on local practices and user adoption (e.g. Cameron et al. 2005). Some papers (e.g. Kellogg et al. 2006) mention the use of IM mediated backchannels within the context of conference calls, but without providing examples or theoretical understanding of what is occurring. No research was found regarding the use of RTC or IM within dispersed teams and the subsequent development or maintenance of cohesion. There are few papers covering IM mediated backchannels in other contexts. Cogdill et al. (2001) cover the use of backchannels within Multi User Domains (MUDs), but do not describe the understanding or theory of use. McCarthy et al. (2005) investigate the experimental use of IM mediated backchannels amongst attendees during conference presentations. The paper categorises use, but does not attempt to describe how the taxonomy was derived. Their use of backchannels varies from those encountered in this research. In their setting, backchannels were visible and open to all, rather than being hidden. The authors conclude that “a number of issues warrant further study, such as the nature of the social bonding that takes place in collocated verses distributed groups”.

Within the fields of distributed work and computer supported collaborative work, there is coverage of the need for cohesion within distributed teams for work to take place effectively. Olsen et al. (2000), describe situations when distributed work can be effective. One factor that they describe is common ground, which is the need for teams to be able to establish a shared identity, level of context and trust before they are able to engage in successful collaboration at a distance. This is reflected by Hinds et al.
(2005), who indicate that shared identity, context and spontaneous communication can have a moderating effect on conflict arising within geographically distributed teams. Cohesion that has been developed or maintained by mediated interaction is covered by a number of scholars. Licoppe (2004) posits that frequent mediated and co-present interaction between groups is interwoven and increases cohesion. Ling (2008) examines how mobile telephony affects interactions and suggests its use strengthens social ties within the close circles of frequently called contacts.

1.1 Aims and Objectives

The objectives of the research, described within this paper are threefold. First, it aims to assess the practicability of the applying Goffman’s theories of Interaction Order to understand interactions between employees mediated by RTC during and within meetings. Second, it assesses whether the research method, is sufficient to answer the research question. Last, it ascertains whether knowledge of how RTC is used within organizations is advanced, and whether this is of practical benefit.

2 METHOD

The empirical data used in this research, and the method through which it was collected, was informed by a number of sub questions: How do employees communicate? What forms of communication do they use? In what working contexts do they communicate? And what ends do they achieve through communication? The unit of analysis that was applied in the research is that of “situated instances of interaction”.

The epistemological orientation of this research is that human interactions can have different meanings for different human subjects, as well as for the researcher investigating them (Lee 1994). There may be an element of subjectivity in individuals’ opinions of meaning. In addition these meanings are inter-subjective as they are socially constructed and agreed amongst individuals within a team or organization, as well as between subjects and researcher. This research is broadly based on social constructivism and interpretive epistemology. An interpretive understanding, at a micro sociological level, is applied as to how interviewees understand their interaction using backchannels.

Qualitative interviews were carried out at two companies to form two case studies. The interview questions concerned individuals’: work style and practices; use of communications; use of presence; process of contacting colleagues; tendency to switch between communications methods. The first case study considers AppsCo, an established provider of business software applications. The team at AppsCo were concerned with product management and marketing of corporate office solutions. The second case study considers TelCo, an established provider of telecommunications services to corporations. The team at TelCo were concerned with propositions management of corporate mobile solutions. Ten interviews were undertaken, six at AppsCo and four at TelCo. The interviews were typically ninety minutes long. Approximately 750 minutes of recordings, covering ten interviews, were transcribed into 280 pages of notes.

A first stage of coding “pre-processed” the corpus of data, to break it down into a set of situations of mediated interactions, according to the unit of analysis. A broad set of coding categories was used in order to do this. Category groupings included: communications behaviours; forms of communication; working contexts within which communication occurred; purposes of instances of communications. A second stage of coding assisted the analysis of interactions and meanings by labelling situations of interactions with concepts from Goffman’s theories concerning Dramaturgy and then Interaction Ritual. The results of this second stage of analysis are considered in this paper.
3 THEORY

When considering theories relevant to the research question, it is useful to refer to Fiske’s definition of communication (1990) as “social interaction through messages”, involving symbols and codes. He assumes that these symbols and codes are transmitted to others and that communication is central to the existence of society. Fiske divides theories of communication into the process and the semiotic school. The semiotic school understands communication as how messages interact with people to produce meaning. Whereas the process school sees social interaction as the process by which one person relates to another, the semiotic school considers social interaction as that which constitutes the individual as a member of society. Similarly, the process school considers a message as that which is transmitted by the communication process, and the semiotic school views messages as the contribution of symbols, which through interaction with the receiver, to produce meanings (Fiske 1990).

The need to understand the meaning of interaction over backchannels, places the research question in the semiotic school of communications theory. The understanding of meaning is highly dependent upon the context of the backchannel interaction, as “The appreciation of the various details that define context is a subtle mental process and provides the key for interpreting a situation and the meaningful elements that pertain to it” (Kallinikos 2001). The interaction that takes places within the “backchannel” is highly contextualized as it takes its meaning from the primary interaction that is taking place, namely from within the conference call or meeting. In a similar way, any discussion or secondary interaction within the “backchannel” may provide context to subsequent primary interaction within the conference call or meeting.

The framework of Interaction Order (Goffman 1983), provides a seam of logic and method through which to understand the meaning of face to face, and possibly, mediated interaction. Goffman argued that people construct meaning for their social world through language and interaction. Interaction socially constructs meaning, and through this he believed that the sense of self was constructed as a product of this social meaning. Furthermore he thought that this meaning can be accessed through the careful analysis of interaction behaviours. Goffman’s focus was on understanding meaning behind individual instances of face to face interaction, and his methodological approach was generally based on ethnographic observation of individuals and groups in their day to day settings. Goffman’s unit of analysis was the situation within which interaction was taking place.

Ling (2008) applies Goffman’s theories, through extensions of Goffman’s Interaction Ritual (1967), to interpret the meanings mobile phone users ascribe to their mediated interaction. In this research, two aspects of Goffman’s framework of Interaction Order (1983) and extensions of these theories (Collins 2005; Ling 2008) were used in order to provide a lens through which to interpret the meaning that subjects give to their interactions through backchannels within the context of the conference calls. In addition, Goffman’s theories give clues as to how these mediated backchannel interactions may contribute to a sense of cohesion amongst distributed teams.

3.1 Dramaturgy and its Application

Dramaturgy, which was put forward in Goffman’s book “Presentation of Self in Everyday Life” (1959), is the first theory that was applied. Dramaturgy is an approach to interpreting how the self manages impressions and what is conveyed based on interactions within groups of people. He used the metaphor of “life as theatre” as a lens through which to analyse interaction within life situations. He uses this lens in order to understand the tactics and techniques by which people use the ritual of theatre to construct the impressions that they give and that individuals take. The main premise in this analysis is that human interactions are dependent on contextual factors such as time, place and audience. The impressions that one human presents to another is based on cultural norms, values and expectations that are contextualized. Goffman uses the theatrical metaphor to provide terminology to describe how individuals go about organizing their performances. Goffman assumes that interactions are social
performances in which individuals are concerned with giving off and maintaining certain desired impressions of themselves to others with the aim of achieving their goals of interaction. The aim of this act of self presentation is to gain acceptance from other people ("the audience") through manipulation. If "the actor" succeeds then audiences perceive "the actor" as she would like.

There are a number of theatrical terms that Goffman introduces to illustrate the performance of impressions management. Goffman introduces the concepts of "front stage", where the self is visible to "the audience" and "back stage", where the self is invisible. When "the actor" is present on the "front stage" she is visible to "the audience" and must manage her impressions or "role" as part of the "performance". Her demeanour and the way she behaves may well be quite different "back stage" to what it is "front stage". When she is "front stage" she is in role and her impressions management need to convey this. As in theatre "boundaries" are required in order to prevent or restrict the movement of "the audience" sitting "front stage" from getting access "back stage", in order that the performance remains credible. The ways through which hidden "back stage" interactions using backchannels, facilitate the progress of meetings can be examined using the language of Dramaturgy.

3.2 Interaction Ritual and its Application

The second theory taken from Goffman's Interaction Order (1983), is that of Interaction Ritual (1967), which analyses the rules of conduct that bring actors together. In Interaction Ritual (1967), Goffman builds upon the concepts of ritual, with origins in Durkheim's "Elementary Forms of Religious Life" (2008). Durkheim was concerned with the dynamics of ritualistic events at a macro level in order to study the societal transitions from one state to another (Ling 2008). The ritualistic events, that he studied, were special occasions, such as weddings, and typically took place in a religious setting rather than everyday life. Durkheim was of the opinion that the context of group interaction in rituals generated solidarity through shared symbolism, and that it is ritual and ritualistic events that bind society together (Ling 2008). Goffman applied Durkheim's concept of ritual to every day interactions, moving the ritual taken within the perspective of the macro to the micro. Goffman believed that there was a sharing of mood and mutual recognition of a situation in mundane rituals (Ling 2008). Building on Durkheim's ideas he posited that interaction ritual provides rules of conduct that provide cohesion and bind actors together (Ling 2008).

Collins (2005) develops Interaction Ritual further in an attempt to connect Goffman and Durkheim in order to link the micro perspective of day to day ritual to the macro perspective of the reinforcement and stability of social structure. Collins (2005, p.48) extends the work of Durkheim and Goffman to define ritual interaction as "A mechanism of mutually focused emotion and attention producing a momentarily shared reality which thereby generates solidarity and symbols of group membership." Through the perspective of Interaction Ritual Chains (2005) individuals focus on the same ritualistic object or action. By (re)enacting these interactions individuals practice deference and demeanour (Goffman 1967) and they develop a common mood. The repetition of micro interactions leads to the significance of artefacts (jargon, symbols or totems) being revitalized and they become charged with emotional intensity (Collins 2005; Ling 2008). Collins claims that the effervescence (Ling 2008) of these events creates social cohesion, as there is a shared sense of energy amongst the participants. Finally, Collins posits that this process allows a sense of solidarity or allegiance to form and that this in turn bolsters institutional stability (Ling 2008).

Ling (2008) develops Collins ideas in order that they can be applied to interaction within communities mediated by the mobile phone. The concern that Ling faces with when attempting to make this connection is that the theories of ritual interaction put forward by Durkheim, Goffman and Collins are based on co-located rather than mediated interaction. Ling circumvents this issue in order that the ritual interaction chain can occur. He claims that once a bond has been formed through co-present means, mediated interaction is as effective in generating the intense emotion that is needed to develop and maintain cohesion. He cites Licoppe (2004) who provides evidence to support these claims.
It is proposed to further extend Collins’ work to the context of IM, using the same logic as Ling. Whilst the individuals interacting by IM may not be co-present they may have bonded as working colleagues on previous occasions. When participants interact over a backchannel they are focusing on the same set of rituals, namely the meeting and the purpose of the backchannel interaction. These ritualistic interactions in turn provide a common mood and a degree of emotional intensity by the fact that they feel the need to use a discrete backchannel. In turn the degree of shared energy caused by this interaction may be little different to that caused in the mobile mediated interactions that Ling observes, so that they may in turn lead to cohesion within in a distributed group. The lens of Dramaturgy is used to analyse the empirical data before applying the lens of Ritual Interaction. The insights Dramaturgy provide into the use of the backchannel bring out the drama in IM mediated interaction, and expose the shared energy which is needed for the generation of cohesion in the context of Ritual Interaction.

4 RESULTS

A large quantity of generalised communications usage data was obtained from interviews. Part of this data concerns the use of digital backchannels in the context of meetings and conference calls, and is considered in this paper. The remaining data concerning other form of communications usage and behaviour is considered elsewhere.

Backchannels in the context of conference calls are of interest as participants are physically remote from each other and the possibility of the use of physical notes, gestures and signals to communicate discretely is impossible. Furthermore interactions over the backchannel may be richer and more sophisticated than what can be achieved using signals and gestures. A selection of examples of backchannel usage that arose during interviews are illustrated in table 1.

<table>
<thead>
<tr>
<th>Table 1 - Example quotations illustrating backchannels mapped to behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. “And the XXXX guy was speaking about something and the XXXX will come on “this guy’s bullshitting” …. it’s like who’s IM-ing who about what. Got to watch what I say …”</td>
</tr>
<tr>
<td>2. “So our group manager, once a year she goes away to (headquarters for a review) … she can’t possibly remember everything. And you know the second that she’s in that meeting and she’s on because all the IMs start popping up … it’ll either be her or a person she sits next to … she obviously gets asked a question, and she’ll say “right you want to know what we’re doing for Product X next year and where we see the biggest opportunities”? Then the answer comes back, apparently the same happens with the big execs on the other end. Our side will say “yes, well we think - we’re only four percent market share”. And someone will go shwe-shwe in the big guy’s ear; you know whispers in his ear. And then they say “actually I know it’s 10 percent”.”</td>
</tr>
<tr>
<td>3. “It’s a great tool for asking questions amongst your little community. Having everyone in context, you can write “I didn’t understand that, did anyone else?, should we press for this, should we do this”. It allows you to then have … a much stronger positioning meeting because you can ask the questions you won’t normally do in a face-to-face like “you ask him, if I’ll ask this”. Get that pincer attack on closing down on the negotiation points.”</td>
</tr>
<tr>
<td>4. “You could argue “is there a moral ethical dilemma about that. No, not really because … you would’ve prepared for that meeting anyway, you would’ve assigned the roles of who’s going to be the negotiator. All you’re doing at this point is cementing and tackling the stuff that comes off the left field that you might not have thought about.”</td>
</tr>
<tr>
<td>5. “we’ve been discussing recently with a supplier about … various conditions that XXX enforces as part of it’s procurement and increasing the cash flow is becoming a major issue for companies. And XXX is seeking … to extend its payment terms too …. in this instance because of our systems we’ve haven’t paid this particular supplier on a number of occasions and we’ve actually made it onto their bad debtors list. The procurement guy wasn’t aware that this was the situation when were trying to renegotiate the contract. So I logged on, quickly got in touch with him and said “this is the situation” and that guy had to change his tack slightly and he didn’t push quite so hard to get it.”</td>
</tr>
</tbody>
</table>

Three categories of backchannel use within the context of Conference Calls were identified. They have been given the descriptive titles “bitching”, “bouncing” and “brawling”. “Bitching” describes
interactions which do not contribute to the progression or success of the meeting. These interactions are typically much more “social” in nature and are typically used to “let off steam” in the form of sharing frustrations and the sharing of jokes. “Bouncing” describes interactions between two or more individuals in a conference call or meeting and two or more individuals outside of the meeting. Unlike other examples of IM mediated concurrent work that may occur during meetings, these IM exchanges are directly related to the progression and maintaining the success of the meeting. The principle use of “bouncing” involves the pulling in of information from outside, into the conference call or meeting, in order to facilitate “its momentum” and to prevent the meeting from “stagnating”. “Brawling” describes interactions which seek to clarify a point or to strategise over an issue occurring in the meeting, so that one or more of the backchannel party can then reengage in the conference call or meeting and tackle the issue effectively. These categories are listed against the examples of backchannel usage in table 1.

These interactions are now analysed in depth, firstly through the lens of dramaturgy and then through the lens of interaction ritual. They are illustrated with example “thick descriptions”.

4.1 Conference Calls as a Performance

In the context of Dramaturgy, conference calls and meetings take place within the context of “front stage” which is the meeting room for face to face meetings, or virtual space for conference calls. Without a backchannel, there is little opportunity for “back stage” interactions, so that all of the actions by the “actors” can be seen or heard by the “audience”. However, with the addition of laptops enabling IM, it becomes possible for the actors to engage in “back stage” interactions over a backchannel, whilst the meeting or “performance” is unfolding “front stage” at the same time. The “back stage” is invisible to the “audience” and “actors” can interact without their behaviour being seen or heard by the “audience”, e.g. “You can have conversations where you couldn’t have physical conversations at the same time”. The descriptions of backchannel usage that were identified from the interviews were analysed by applying dramaturgical concepts. Table 2 illustrates dramaturgical concepts occurring in the example quotations of table 1.

<table>
<thead>
<tr>
<th></th>
<th>Front stage (meeting)</th>
<th>Back stage (backchannel)</th>
<th>Actor(s) (backchannel participants)</th>
<th>Audience (individuals external to backchannel)</th>
<th>Boundary (dividing subjects and objects)</th>
<th>Performance (front presented by actors)</th>
<th>Asides (comments between actors)</th>
<th>Prompting (useful provided over backchannel)</th>
<th>Improvisation (actions not prearranged)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

“Bitching” interactions are typically “social” and joking in nature, e.g. Quotation 1. This comment indicates the necessity to guard “boundaries” between “front stage” and “back stage” and to prevent the “audience” gaining access, otherwise the “performance” will no longer be credible. “Bitching” as a form of “back stage” interaction resembles an “actor” complaining that the “audience” is particularly flat that evening. These “asides” distinguish “Bitching” from the other two categories. The “audience” are not intended to see any of these “back stage” backchannel interactions; otherwise their impression of the “performance” might be ruined.

The principle form of “bouncing” involves the pulling in of information from outside, in order to keep up the momentum of the conference call or meeting, e.g. Quotation 2. The process of “bouncing” is similar to the “prompting” of actors when they forget their “lines” which is occasionally required to keep up the momentum of a “performance”. However, the level of “prompting” must be kept as subtle as possible, as in a play, in order to maintain the “appearance” and “front” of the “actor”, so that their
“role” remains credible and convincing. Ideally the “audience” won’t be aware of “prompting” and the backchannel is relative subtle for this purpose, as it remains “back stage” out of sight. It is “prompting” that makes “bouncing” unique amongst the categories.

There were examples of individuals on conference calls using the backchannel to clarify points within their “team of actors” “back stage” before “brawling” “front stage”. The advantage of being able to have these interactions “back stage” on the backchannel, whilst the “performance” or meeting is continuing “front stage”, is that the “audience” remains oblivious as to these discussions. As a consequence the “actors” are able to coordinate and manage the “performance” without having to interrupt it. The ability for “actors” to be able to have these “back stage” dialogues enables sophisticated strategies to be developed quickly and responsively to achieve the objectives of the meeting, e.g. Quotation 3. There was some debate as to the morality of strategising “back stage”, but it was considered ethical, as teams would typically prepare and “rehearse” before meetings, and the backchannel is used for dealing with the unexpected, e.g., Quotation 4. In this last quotation, the interviewee raises an additional point of interest regarding to be able to react to unforeseen circumstances. The ability to be able to have a dialogue “back stage” is that it enables the actors to “improvise” better. It is “improvisation” that sets the category of “brawling” apart from the others. The ability, that backchannel provides, to allow people to be “front stage” and “back stage” at the same time, without interrupting the “performance” provides an ability to fine tune, or “direct”, events whilst they are unfolding, e.g. Quotation 5.

4.2 Conference Calls as Ritual

There is logic in applying both the concepts of Dramaturgy and Ritual Interaction. First, they both form part of Goffman’s Interaction Order (Goffman 1983). Second, the two perspectives are linked so that themes from Dramaturgy inform Interaction Ritual, in terms of the study of human interaction, its methodological and epistemological approach, and in terms of the concepts and vocabulary used. Last, the initial Dramaturgic analysis helps identify the shared energy that is created in parties interacting, which then feeds the sense of social cohesion resulting from Interaction Ritual.

For ritual interaction to take place, Collins (2005, p.48) states the following conditions must be in place: Two or more physically assembled people; Boundaries to outside; A common focus of attention through which participants are mutually aware of each others’ focus of attention; Sharing of mood (including: Collective engrossment; Sense of emotional energy; Development and markers of shared relationship; Standards of morality (e.g. sense of rightness) associated with being part of the group and willingness to defend against transgressors). Ling (2008) develops Collins’ ideas further in order that they can be applied to interaction mediated by the mobile phone. The first condition becomes: Two or more physically assembled people or people who are connected via mediated communication. Table 3 illustrates how all these conditions are matched within each example of backchannel usage. The paragraphs that follow describe how this comes about for each example.

| Table 3. Chain of Evidence: Ritual Interaction Concepts occurring in Example Quotations |
|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
|                                | Two or more physically assembled people | Boundaries to outside | Common Focus of Attention | Sharing of Mood: Collective Engrossment | Sharing of Mood: Sense of Emotional Energy | Sharing of Mood: Development of Shared Relationship |
| Q1                              |                                |                                |                                |                                |                                |                                |
| Q2                              |                                |                                |                                |                                |                                |                                |
| Q3                              |                                |                                |                                |                                |                                |                                |
| Q4                              |                                |                                |                                |                                |                                |                                |
| Q5                              |                                |                                |                                |                                |                                |                                |
Since “bitching” interactions are typically much more “social” and joking in nature, they may transform a tense energy within a meeting into energy which is more agreeable, e.g. Quotation 1. Applying Collins’ framework to these interactions, it can be seen that “two or more people are connected via mediated communication” by the very fact that they are engaged in IM interaction. There are clearly “boundaries to outside” as there are other people within the primary context of the conference call / meeting, for whom the content of the “back stage” interaction is not intended. There is a “common focus of attention through which participants are mutually aware of each others’ focus of attention”, the act of “bitching” within the primary context of the conference call/meeting is the common focus, and the backchannel is the means through which the “actors” are mutually aware of each others’ focus of attention. With regards to the final aspect of Collin’s framework, “sharing of mood”, there is a sense of “collective engrossment” and “emotional energy” through the fact that the act of “bitching” is transforming one form of emotional energy (e.g. frustration at being bored) to another (e.g. relief that someone else might be feeling the same way). It can be claimed that the conspiratorial nature of sharing a joke about how the meeting may develop or maintain “markers of shared relationship” and may indicate that there are “standards of morality (e.g. sense of rightness) associated with being part of the group and willingness to defend against transgressors”.

The principle form of “bouncing” involves the pulling in of information from outside, into the conference call or meeting, e.g. Quotation 2. Once again, applying Collins’ framework to these interactions, it can be seen that “two or more people are connected via mediated communication” by the very fact that they are engaged in IM interaction. There are clearly “boundaries to outside” as there are other people within the primary context of the conference call / meeting, for whom the “back stage” interaction is not intended. There is a “common focus of attention through which participants are mutually aware of each others’ focus of attention”, the act of “bouncing” within the primary context of the conference call/meeting is the common focus, and the backchannel is the means through which the “actors” are mutually aware of each others’ focus of attention. With regards to the final aspect of Collin’s framework, “sharing of mood”, there is a sense of “collective engrossment” and “emotional energy” through the fact that the act of “bitching” is urgent and emotionally charged in nature. It can be claimed that the hidden nature of sharing information and co-operating together may develop or maintain “markers of shared relationship” and may indicate that there are “standards of morality (e.g. sense of rightness) associated with being part of the group and willingness to defend against transgressors”. The “actors” in these quotations are sharing information with their “audience” to defend their point of view, and there is a sense of “us and them”.

Examples of “brawling” are reported for two observed purposes. Each has its own role in interaction ritual. The first is to enable a group to discuss strategies improvise solutions, “back stage” whilst being “front stage” at the same time, in order to cope with unforeseen circumstances on a conference call, e.g. Quotation 3. The second use of “brawling” is to influence, or “direct”, the “performance” of another “actor” in order to maintain the success of the overall “performance” of the conference call. Again this is done “back stage” whilst being “front stage” at the same time, e.g. Quotation 5. Applying Collins’ framework to these interactions for the last time, it can be seen that “two or more people are connected via mediated communication” by the fact that they are engaged in IM interaction. There are clearly “boundaries to outside” as there are other people within the primary context of the conference call / meeting, for whom the content of the “back stage” interaction is not intended. There is a “common focus of attention through which participants are mutually aware of each others’ focus of attention”, the act of “brawling” within the primary context of the conference call/meeting is the common focus, and the backchannel is the means through which the “actors” are mutually aware of each others’ focus of attention. With regards to the final aspect of Collin’s framework, “sharing of mood”, there is a sense of “collective engrossment” and “emotional energy” through the fact that the act of “brawling” is inclusive of a group and requires their attention. In addition there is likely to be a common sense of charged energy as the “actors” plot to outwit their “audience”. It can be claimed that plotting together may develop or maintain “markers of shared relationship” as those who are involved are very much part of the “in group” in contrast to the “out group” against who they are scheming. The act of brawling, intended to defend the interests of the “actors” against those of the “audience”,
indicates that there are “standards of morality (e.g. sense of rightness) associated with being part of the group and willingness to defend against transgressors”.

5 DISCUSSION

The framework of dramaturgy can be usefully applied to understanding and interpreting interactions over backchannels in the context of conference calls and meetings. It would also appear that the extensions of Interaction Ritual (Collins 2005; Ling 2008) can be used to understand interactions mediated via IM through backchannels, and that meaningful interpretations can be drawn from this analysis. A number of insights emerge from these analyses, which can be assessed, in terms of relevant findings, contributions and failings, against the objectives of this research.

5.1 Theoretical Findings, Contributions and Failings

The research had the objective of assessing the practicability of applying aspects Goffman’s Interaction Order and derived theories to interpreting and understanding meanings attributed by users of IM mediated backchannels in the context of conference calls and meetings. To date there has been little research into use of RTC within organisations (Frößler & Klein 2007) and little research into the use of Digital Backchannels (Isaacs et al. 2002; Kellogg & Erickson 2006; Cogdill et al 2001; McCarthy & Boyd 2005). Application of Goffman’s theories allow the meanings of Backchannels to be interpreted and understood. Dramaturgical concepts (Goffman 1959) provide a useful means of interpreting the interactions that occur through backchannels within the context of conference calls and meetings. Interaction Ritual (Goffman 1967) and extensions (Collins 2005; Ling 2008) were applied to the empirical data in order to ascertain whether these interactions were indeed ritualistic after Collin’s and Ling’s definitions. Interactions through backchannels fit Collin’s criteria, adapted by Ling, satisfactorily, and these mediated interactions can be loosely termed Interaction Rituals. In this way, it can be claimed the theoretical objective is achieved, and that the study contributes to the literature. It demonstrates that Goffman’s Interaction Order can be successfully applied to interpret interactions through IM mediated backchannels in the context of conference calls. However, the study had set out to claim that use of backchannels generates or maintains social cohesion within groups (through interaction ritual). Unfortunately, it cannot be claimed that this is the case. The reasons for this are twofold. First, for social cohesion to be generated and maintained by interaction ritual, the ritual needs to be repeated frequently (Ling 2008), a concern which is confirmed by Licoppe’s research (2004) into “connected presence”, which shows that it is frequent mediated interaction that helps maintain social bonds. Whilst it is apparent that the average day is quite literally “littered” by conference calls for some interviewees in this research, there is insufficient evidence that they engage in “bitching”, “bouncing” and “brawling” frequently enough. This question was not asked, nor was the evidence volunteered. Second, for the same reason, interviewees did not mention the fact that they find interaction over the backchannel a socially cohesive act, either directly or indirectly. Nevertheless these practices fulfil the criteria of Interaction Ritual which consequently provides evidence that they may contribute social cohesion, especially in conjunction with other factors.

5.2 Methodological Findings, Contributions and Failings

This research had the objective of assessing whether the research method is sufficient to address the research question. The chosen research design was based on obtaining qualitative data from semi structured interviews from two organizations and to present them as two case studies. Sufficient empirical data was collected in order to address the research question regarding backchannels. However, there are shortcomings with the research design and method. The first concerns the fact that research in Goffman’s tradition is based ethnographic observation, rather than case study interviews. Ling (2008), upon which this research is based, used ethnographic observation in order to obtain empirical data for his research into social cohesion, based on Interaction Ritual, generated from use of
mobile telephony. It is debatable whether an ethnographic study would have yielded sufficient data, given the short period of time that was available for field work. A second criticism is that whilst sufficient evidence was acquired to address most of the research question using Goffman’s concepts, there was no independent evidence to back up the claim that Interaction Ritual within backchannels generates social cohesion amongst interviewees. Future research designs must be improved in order that a full contribution is provided.

5.3 Practical Findings, Contributions and Failings

The research had the practice objective of assessing whether the research advances knowledge of how RTC is used within organizations, and whether this is of practical benefit. The process of analysis identified the following four findings. Firstly, in the backdrop of meetings, IM mediated backchannels provides a group of individuals an environment where they can engage in sophisticated interaction, invisible to other meeting attendees. Secondly, backchannels provide a dynamic means for information to be invisibly brought into face to face meetings or conference calls from outside. Thirdly, backchannels allow real time coordination amongst groups within meetings, in a manner invisible to others, which allows them to improvise and react to unexpected events as they unfold. Lastly, mediated interaction through backchannels shows characteristics of Interaction Ritual, which generate an energy amongst those who participate, which may contribute to social cohesion within distributed teams. These findings may be of mixed benefits to organizations. On the one hand interviewees reported how backchannels were of benefit to the way they conducted conference calls. On the other they also reported how backchannels and other examples of concurrent working may be distracting in the context of conference calls and face to face meetings. Indeed some senior managers were introducing policies preventing the use of laptops in meetings, although this would be difficult to enforce in conference calls. Finally, the application of dramaturgy to mediated interactions such as IM interaction through backchannels provides an insightful means of understanding how employees are interacting and how technology is used. In this way this research has contributed to the body of literature within the domains of Distributed Work, CSCW and especially RTC and IM.

5.4 Potential for Future Research

Two areas for future research are immediately identifiable. One is to continue and improve research regarding the use of IM mediated backchannels, and the second is to extend analysis of mediated interactions to other contexts and communications technologies within the organizational sphere. With respect to further research into backchannel interactions mediated by IM, or other technologies, there is a need to correct the methodological failings of this current research. First, an ethnographic approach to collecting data would be more consistent with the methodological approach of Goffman. Second, more evidence is required in order to make claims regarding the creation or maintenance of social cohesion as a result of mediated interactions studied. Last, research into this phenomenon would be more complete if the enquiry extended beyond the act of interaction as agency and also investigated the institutional structural elements and the impact of artefacts, which mediate interaction.

6 CONCLUSION

This research contributes to the analysis of the use of digital backchannels in the context of conference calls and RTC in the areas of practice, methodology and theory.

In practice, it was observed that in the backdrop of meetings, IM mediated backchannels provides a group of individuals an environment where they can engage in sophisticated interaction, invisible to other meeting attendees. Secondly, backchannels provide a dynamic means for information to be invisibly brought into face to face meetings or conference calls from outside. Thirdly, backchannels allow real time coordination amongst groups within meetings, in a manner invisible to others, which
allows them to improvise and react to unexpected events as they unfold. Lastly, mediated interaction through backchannels shows characteristics of Interaction Ritual, which generate an energy amongst those who participate, which may contribute to social cohesion within distributed teams. These findings may be of mixed benefits to organizations. On the one hand interviewees reported how backchannels were of benefit to the way they conducted conference calls, but on the other they reported how it can cause distraction.

Whilst the research design provided sufficient empirical data upon which to analyse the use of backchannels and to identify evidence of social cohesion, it was based on case study interviews rather than on ethnographic observation. Changes to the research design, in order that data is collected through ethnographic observation, would make it more consistent with previous research (Goffman 1983; Ling 2008), upon which it is based.

The application of Goffman’s Interaction Order and derived theories appear to be an innovative lens to interpret and understand meanings attributed by users of IM mediated backchannels in the context of conference calls and meetings. These practices fulfil the criteria of Interaction Ritual, and there is evidence that they may contribute to social cohesion within teams.

7 REFERENCES

Aducci, R., 2008. The Hyperconnected: Here They Come!, IDC.
Kallinikos, J., 2001. The Age of Flexibility - Managing Organizations and Technology First., Academia Adacta AB.
DEVELOPMENT AND VALIDATION OF A MODEL FOR ASSESSING THE SUCCESS OF EMPLOYEE PORTALS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0439.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>DeLone and McLean model, Empirical study, Information system effectiveness, E-collaboration</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
DEVELOPMENT AND VALIDATION OF A MODEL FOR ASSESSING THE SUCCESS OF EMPLOYEE PORTALS

Urbach, Nils, European Business School (EBS), Rheingaustr. 1, 65375 Oestrich-Winkel, Germany, nils.urbach@ebs.edu
Smolnik, Stefan, European Business School (EBS), Rheingaustr. 1, 65375 Oestrich-Winkel, Germany, stefan.smolnik@ebs.edu
Riempp, Gerold, European Business School (EBS), Rheingaustr. 1, 65375 Oestrich-Winkel, Germany, gerold.riempp@ebs.edu

Abstract

Many companies utilize employee portals to improve information exchange and communication between their employees as well as to better support business processes. Owing to limited IT budgets and the need to justify portal investments, assessing employee portal benefits is important. In this article, we propose a model for measuring employee portal success on the basis of the DeLone and McLean IS Success Model as well as an extensive review of employee portal success measurement literature. The resulting model is put forward as the foundation for future empirical work in this area. A card-sorting and item-ranking approach as well as a survey trial is conducted in order to validate the survey instrument. Methods of classical test theory are proposed for the further validation of the survey instrument and the assessment of the measurement model.

Keywords: Employee Portals, Business-to-Employee, IS Success, IS Effectiveness


1 INTRODUCTION

The ways employees handle information, communicate, as well as execute business processes have significantly changed with the emergence of web-based technologies and the subsequent emergence of employee portals. Over the past decade, company intranets transformed from collections of static web pages into highly integrated and interactive information systems (IS). Whereas the first-generation intranets only provided a single interface to information, today’s employee portals enable the integrated support of information, communication, applications, and business processes.

Technically speaking, an employee portal offers a browser-based user interface which provides access to personalized information, resources, and applications. In many cases, an employee portal is the primary tool through which employees do their work. Ideally, employee portals yield different benefits to both organizations and employees, such as reducing information overload, reducing organizational costs, improving corporate communication, and enhancing employee productivity (Tojib & Sugianto & Sendjaya 2006).

Today, many companies, especially large ones, offer their employees portals. A 2006 US study by Forrester Research, Inc. indicates that 46% of large companies offered an employee portal, and another quarter planned to establish one by 2008 (Forrester 2006). The use of employee portals has been growing steadily and, despite restricted IT budgets in many companies, investments in portal solutions are still growing (Remus 2006). Although IT departments and decision-makers have to justify portal investments, a significant number of companies do not assess the actual benefits of their portal implementations (Brown & Mines & Moore & Barnett 2007).

Companies which do assess their portal benefits have often used monetary indicators such as return on investment (ROI) or total cost of ownership (TCO) and other cost-benefit analysis methods (White 2003). These success-measurement approaches do not take into account intangible impacts and intervening environmental variables. Clearly, in order to be comprehensive, success measurement would need to consider both tangible and intangible effects in order to truly judge portal success, to detect potential improvements, and to justify present and future investments in portal solutions.

To date, only a few studies have scientifically investigated single aspects of employee portal success such as user satisfaction (bin Masrek 2007, Sugianto & Tojib & Burstein 2007) and service quality (Cheung & Lee 2005). None of the reviewed studies examined employee portal success comprehensively and in an integrated way. The dominant model for measuring IS success is the DeLone and McLean IS Success Model (DeLone IS Success Model) (DeLone & McLean 1992, DeLone & McLean 2003). It is considered a sound basis for measuring employee portal success, as it is a comprehensive evaluation framework with validated measures and associations; it has also been applied to several types of IS (Urbach & Smolnik & Riempp 2008).

In this paper, we report on our development of a conceptual model, based on the D&M IS Success Model, which considers the specific requirements of employee portals. Where possible, we employ existing measures that have been tested and proved. The following section 2 describes the theoretical foundations of employee portals and a literature review on measuring IS and portal success. In section 3, we explain how we developed our employee portal success model and discuss the characteristics of its success dimensions. Section 4 outlines our steps for validating the survey instrument applying a card-sorting and item-ranking approach as well as conducting a survey trial. The future validation of the survey instrument and the assessment of the measurement model are presented in section 5. To conclude, we summarize the paper’s contribution, refer to the limitations, and present suggestions for additional future research.
2 FOUNDATIONS

The starting point for developing the conceptual model for assessing employee portal success was studying the existing research in the field. We thus reviewed relevant literature on employee portals, information systems success, and existing approaches for evaluating portals.

2.1 Employee portals

Employee portals can be regarded a specific type of the broader concept of enterprise portals. The latter will therefore be discussed first. In this context, we use corporate portal, enterprise information portal, business portal, and enterprise portal interchangeably (Dias 2001).

An early definition of a portal in the corporate context appeared in a Merill Lynch report (Shilakes & Tylman 1998). In this report, an enterprise (information) portal is considered an application that primarily integrates the company’s information and provides users with a single interface to this corporate information. Subsequent definitions tend to include the integration of collaborative applications such as e-mail and calendars (Eckerson 1999). Current enterprise portals not only integrate information and simple tools, but also business applications and processes (Chan & Liu 2007, Daniel & Ward 2005). In summary, enterprise portals have evolved from low-end intranets into highly integrated IS. Today, such portals enable the integrated support of information, communication, applications, and business processes.

Enterprise portals can be classified in terms of their target user group as supplier, customer, and employee portals. Ideally, these three types have the same technical infrastructure. Thus, organizations effectively have one portal with different “windows,” that give each user group access to specific functions that are relevant to it (Riempp 2002).

Before portal technology was available, the web-based intranet was a popular tool for building work force commitment (Azzone & Bianchi 2000). Although it yielded benefits for organizations in this regard, these intranets lacked personalization, offered poor navigation, and did not provide centralized access to information, which often led to losses in productive employee time. To overcome these problems, organizations began to implement employee portals (Tojib et al. 2006).

An employee portal is a web-based interface to access personalized information, resources, applications, and e-commerce options. Employees can access a range of internal and external information from their computers through a network connection. They are provided with relevant proprietary information, displayed in a password-protected setting (Sugianto & Tojib 2006). In addition, business applications are increasingly integrated into employee portals. Thus, in many organizations, the role of the employee portal has become crucial, especially when an entire business process can be completed through the use of the portal. In some organizations, an employee portal is the primary tool through which employees do their work (Tojib et al. 2006).

Ideally, employee portals provide a number of benefits to organizations and employees. These include:
- structured access to enterprise information;
- common and personalized views as well as collections of portal elements;
- improvement of organizational information gathering as well as knowledge acquisition and management;
- improvement in employee productivity;
- improvement of corporate communication; and

Our literature review shows that employee portals have received some attention, but that there is a lack of empirical studies to confirm the assumed benefits of employee portal use.
2.2 Research on measuring IS success

The IS literature provides several definitions and measures of IS success. As DeLone and McLean (1992) state, there are nearly as many measures as there are studies. Obviously, there is no ultimate definition of IS success. Since there are different stakeholders who assess IS success in an organization (Grover & Jeong & Segars 1996), each group has a different definition. From a software developer perspective, a successful IS is completed on time and under budget, has a set of features that is consistent with the specifications, and functions correctly. Users may find an IS successful if it improves their work satisfaction or work performance. From an organizational perspective, a successful IS may contribute to the company’s profits or create a competitive advantage. Consequently, success is always assessed from a certain stakeholder’s point of view. Furthermore, IS success also depends on the type of system being evaluated (Seddon & Staples & Patnayakuni & Bowtell 1999).

In order to provide a more general and comprehensive definition of IS success, one that covers these different perspectives, DeLone and McLean (1992) reviewed existing definitions of IS success and their corresponding measures, and classified them into six major categories. They then created a multidimensional measuring model with interdependencies between the different success categories. The D&M IS Success Model received much attention from IS researchers. Since its publication, many researchers have treated IS success as a multidimensional construct and have measured it as such (Urbach et al. 2008).

Motivated by DeLone and McLean’s call for further development and validation of their model, many researchers have attempted to extend or respecify the original model. A number of researchers have claimed that the DeLone and McLean Model is incomplete; they either suggest that further dimensions should be included in the model, or they present alternative success models (Seddon 1997, Seddon & Kiew 1994). Other researchers have focused on the model’s application and validation (Rai & Lang & Welker 2002). Although some weaknesses have been revealed, the D&M IS Success Model has become a dominant model for measuring IS success (Hu 2003).

Ten years after the publication of their first model, and based on the evaluation of the many contributions to it, DeLone and McLean proposed an updated IS Success Model (DeLone & McLean 2003). The updated model consists of six interrelated dimensions of IS success: information, system and service quality, (intention to) use, user satisfaction, and net benefits. It can be interpreted as follows: A system can be evaluated in terms of information, system, and service quality; these characteristics affect subsequent use or intention to use and user satisfaction. As a result of using the system, certain benefits will be achieved. The net benefits will (positively or negatively) influence user satisfaction and further information system use.

The updated D&M IS Success Model appears to be a sound basis for measuring the success of employee portals because:

- it is a comprehensive evaluation framework;
- the proposed associations have been validated by a large number of empirical studies;
- there are many validated measures for the proposed success dimensions that can be reused;
- it has been applied to several types of information systems;
- web-based systems other than employee portals have been evaluated by applying this model; and
- it is the dominant evaluation framework in IS research (Urbach et al. 2008).

2.3 Research on portal success

Existing measurement approaches to assess portal success in practice usually utilize monetary indicators. Typical examples are return on investment (ROI), total cost of ownership (TCO), or other cost-benefit analysis methods (White 2003). Building business cases on the basis of such indicators is questionable because non-monetary impacts and intervening variables are not taken into account.
There is little documented research on employee portal success measurement. Some studies investigate single aspects of employee portal success, but none of the studies we reviewed took a comprehensive, integrated approach.

In order to measure user satisfaction with employee portals, Sugianto et al. (2007) proposed the model of the B2E Portal User Satisfaction (B2EPUS) on the basis of the End-User Computing Satisfaction measure (EUCS) developed by Doll and Torkzadeh (1988). Another approach for assessing user satisfaction with campus portals was proposed by Bin Masrek (2007) on the basis of an extract of the updated D&M IS Success Model (DeLone & McLean 2003). Focusing on the user-perceived service quality of web portals, Yang et al. developed and validated an instrument based on different conceptual models in the areas of IS and technology adoption (Yang & Cai & Zhou & Zhou 2005).

In our review of the IS success literature, we found no study specifically aimed at comprehensively examining the success of employee portals. Consequently, we built our measurement model for employee portal success on the above-mentioned focused studies. Secondly, we extended our literature review by analyzing studies focusing on the evaluation of web-based systems (WBS), which are similar to employee portals. Additionally, since employee portals have commonalities with web-based knowledge management (KM) systems, we reviewed empirical studies focusing on KM system success with the objective of adapting measures to our conceptual model.

3 CONCEPTUAL MODELING

Since employee portals are widespread, but there is no known comprehensive, integrated theoretical framework for measuring their success, we propose a conceptual model of employee portal success that is based on the D&M IS Success Model (DeLone & McLean 2003).

3.1 Success dimensions

We modified the dimension definitions of the D&M IS Success Model for application in the employee portal context. To this end, we examined the definitions of the original success dimensions, contrasted them with the specific characteristics of employee portals, and merged the different points of view into a revised classification scheme. Consequently, we propose a conceptual model for measuring employee portal success consisting of the following success dimensions:

- **System quality**, which consists of measures of the employee portal as a system in itself. It considers performance characteristics, functionality, and portal usability, among others.

- **Information quality**, which focuses on the quality of the employee portals’ output, (i.e., the quality of the information that the portal provides) and its usefulness for the user. Information quality has been shown to be a prominent success factor when investigating overall IS success, especially in the context of web-based systems (Schaupp & Fan & Belanger 2006).

- **Service quality**, which includes measures of the overall support delivered by the service provider. In the context of employee portals, this success dimension covers aspects such as responsiveness, reliability, empathy, competence, and the overall quality of the portal owner.

- **Portal use**, which measures the perceived use of the employee portal by a company’s staff. To assess use in this context, we propose to measure the perceived time of use of the different functionalities such as e-mail, searching for information, as well as the overall portal usage time. Since we assume that portal use is typically non-mandatory (Tojib & Sugianto & Sendjaya 2008), we believe that perceived use is a sufficient indicator. Thus, we omit measuring the intention of using the portal as a surrogate for portal use.

- **User satisfaction**, which is the affective attitude of the employee who interacts directly with the portal towards the portal (Doll & Torkzadeh 1988, Tojib et al. 2006). User satisfaction is considered one of the most important measures when investigating overall IS success. The
proposed construct evaluates adequacy, efficiency, effectiveness, and overall satisfaction with the portal.

- **Individual portal benefits**, which subsume measures of the perceived individual benefits gained by the employee through the use of the portal. These benefits cover aspects like task performance, job efficiency, improved use and exchange of knowledge, improved communication, and overall usefulness.

With this model, we focus on individual performance impacts as the final dependent variable of interest – instead of organizational performance. The difficulty of measuring the organizational impact of individual IS initiatives has been discussed and demonstrated by many researchers (e.g., Gelderman 1998, Goodhue & Thompson 1995). Thus, we do not consider the organizational impact in our model, although the impact of utilizing employee portals is very likely greater than the individual impact.

### 3.2 Conceptual model and hypotheses

Our research model for measuring employee portal success on the basis of the updated D&M IS Success Model is shown in figure 1. The proposed model assumes that system, information and service quality are linked to user satisfaction and to use an employee portal, and that these, in turn, influence the portal’s benefits for individuals. Furthermore, user satisfaction and portal use are interrelated. In addition, we propose knowledge intensity of tasks, the level of process standardization, and the organizational culture with respect to using the employee portal as control variables, since we expect that these factors may influence some of the constructs and relationships.

![The Conceptual Model](image-url)

*Figure 1. The Conceptual Model*
3.3 Constructs and measures

In order to operationalize the constructs of the conceptual model, we follow the recommendation by various authors (e.g., Bharati & Chaudhury 2004, DeLone & McLean 2003, Sugianto & Tojib 2006) to use tested and proven measures if possible. Thus, we adopted items identified in previous studies and modified them for use in the employee portal context. For measuring system, information, and service quality we used second-order multidimensional constructs as discussed by Petter et al. (2007). The sub-constructs are operationalized with reflective indicators, and the relationships between sub-constructs and constructs under study are formative. Portal use, user satisfaction, and individual portal benefits are measured as reflective first-order constructs.

After surveying the literature for existing instruments, initial item pools were created for each of the constructs. We added additional items where important aspects of the content domain of a construct have not been covered. In doing so, we followed suggested item writing principles found in the literature (e.g., Dillman 2008). All of the constructs used within the study except portal use should be measured using a seven-point Likert-type scale (1 = Strongly disagree, 7 = Strongly agree). Portal use should be measured in terms of actual daily use and the frequency of use. In table 1, we present an overview of the items we have used to operationalize the constructs.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Items adapted from</th>
<th>Original context</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Quality: Navigation</td>
<td>McKinney et al. (2002), 1 new item</td>
<td>Online shopping</td>
</tr>
<tr>
<td>System Quality: Design</td>
<td>Ahn et al. (2004), 3 new items</td>
<td>Online shopping</td>
</tr>
<tr>
<td>System Quality: Usability</td>
<td>McKinney et al. (2002), 1 new item</td>
<td>Online shopping</td>
</tr>
<tr>
<td>System Quality: Functionality</td>
<td>4 new items</td>
<td>N/A</td>
</tr>
<tr>
<td>System Quality: Responsiveness</td>
<td>McKinney et al. (2002), 2 new items</td>
<td>Online shopping</td>
</tr>
<tr>
<td>System Quality: Availability</td>
<td>4 new items</td>
<td>N/A</td>
</tr>
<tr>
<td>System Quality: Reliability</td>
<td>4 new items</td>
<td>N/A</td>
</tr>
<tr>
<td>Information Quality: Understandability</td>
<td>McKinney et al. (2002)</td>
<td>Online shopping</td>
</tr>
<tr>
<td>Information Quality: Usefulness</td>
<td>McKinney et al. (2002), 1 new item</td>
<td>Online shopping</td>
</tr>
<tr>
<td>Information Quality: Reliability</td>
<td>McKinney et al. (2002)</td>
<td>Online shopping</td>
</tr>
<tr>
<td>Information Quality: Completeness</td>
<td>Yang et al. (2005), McKinney et al. (2002), 1 new item</td>
<td>Web portal, Online shopping</td>
</tr>
<tr>
<td>Information Quality: Timeliness</td>
<td>Lin &amp; Lee (2006), 3 new items</td>
<td>Online communities</td>
</tr>
<tr>
<td>Service Quality: Responsiveness</td>
<td>Pitt et al. (1995)</td>
<td>Computer use</td>
</tr>
<tr>
<td>Service Quality: Training</td>
<td>Chang &amp; King (2005), 2 new items</td>
<td>IS function</td>
</tr>
<tr>
<td>Portal Use</td>
<td>Iivari (2005)</td>
<td>Accounting systems</td>
</tr>
<tr>
<td>User Satisfaction</td>
<td>Seddon &amp; Kiew (1994)</td>
<td>Accounting systems</td>
</tr>
<tr>
<td>Individual Portal Benefits</td>
<td>Davis (1989)</td>
<td>Various information systems</td>
</tr>
</tbody>
</table>

Table 1. Measures of the Item Pool

This initial item pool will be revised and developed further in the validation phase of the research project. We will start with a larger item pool than we will use in the final questionnaire and successively reduce the number of items. A generally suggested minimum is three questions per construct (Kankanhalli & Tan & Wei 2005). Because of the relatively high number of constructs in our research model, we consider a number of three items per each of it as reasonable, since a larger number would result in a too long questionnaire that probably leads to low response rates.
VALIDATION OF THE SURVEY INSTRUMENT

Our development of the conceptual model on the basis of theoretical considerations is the first step of a long-term research project that seeks to present a reliable and valid instrument for measuring the success of employee portals. Before it can be applied in practice, it needs further development and validation. We apply methods of classical test theory to validate the survey instrument and assess the measurement model (Churchill Jr 1979). The first step to be taken prior to the field studies is the validation of the survey instrument.

4.1 Content and construct validity

The main focus of the first stage was to ensure content validity of the item pools, i.e. the degree to which the pools reflected all relevant facets of their underlying constructs. We therefore discussed the choice of items with a group of four IS experts. Based on the experts’ feedback both the choice of items as well as the wording has been refined. Finally, we chose four items for each of the sub-constructs and a larger set of items for the other constructs.

Our next step targeted the refinement of the items into scales with a high level of construct validity. We used a card-sorting and item-ranking approach similar to the one adopted by Davis (1989), Moore & Benbasat (1991), and Kankanhalli et al.(2005). In this approach, a group of IS researchers and practitioners, serving as judges, are given cards with the definition of a target construct, as well as cards with one of the items. Each judge is (independently) asked to assign each item card to one of the construct cards. Items that are sorted into the same category by more than half of the judges can be regarded to having face validity for that construct. If items are consistently assigned to one construct, they can also be considered to demonstrate convergent validity with their construct and discriminant validity with the other constructs. Items not meeting the validity criteria can be considered candidates for modification or deletion. In a second stage, each judge is asked to rank the items according to their representativeness in respect of the specific construct. Items considered to represent the construct weakly are also candidates for modification or deletion.

Instead of cardboard cards, we used a computer-based spreadsheet solution to support the sorting procedure. Eight experts (other than the experts that were consulted to develop the items) participated as judges, who were independently asked to assign each item to one of the constructs or an extra “Ambiguous/Unclear” category. Additional space was provided allowing for comments and suggestions on the items. The card-sorting and item-ranking approach was limited to the sub-constructs of system quality, information quality, and service quality for two reasons. Firstly, in contrast to the other constructs, existing items have been modified to a high degree and additional items have been created. Secondly, a sorting and ranking of all items of the initial pool was not acceptable to the judges.

To assess the consistency of the assignments, we used the item placement ratio (IPR), i.e. the percentage to which judges have assigned items to the intended target construct. In order to increase the construct validity and to shorten the questionnaire, we eliminated the one item of each sub-construct with the lowest IPR. In cases where this criterion was ambiguous, we drew on the results of the item-rankings and eliminated the item with the lowest average ranking. The results for both the initial and the revised item pools are reported in table 2. The overall IPR of .91 for the revised pool indicates satisfactory item placement consistency. However, the operationalization of the assurance sub-construct with an IPR of 42% appeared to be problematic, although we used well established items. Accordingly, we modified the wording of the items on the basis of the judges’ feedback.
<table>
<thead>
<tr>
<th>Constructs</th>
<th>Initial item pool</th>
<th>Revised item pool</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of items</td>
<td>% Hits</td>
</tr>
<tr>
<td>System Quality: Navigation</td>
<td>4</td>
<td>78</td>
</tr>
<tr>
<td>System Quality: Design</td>
<td>4</td>
<td>88</td>
</tr>
<tr>
<td>System Quality: Usability</td>
<td>4</td>
<td>81</td>
</tr>
<tr>
<td>System Quality: Functionality</td>
<td>4</td>
<td>97</td>
</tr>
<tr>
<td>System Quality: Responsiveness</td>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>System Quality: Availability</td>
<td>4</td>
<td>91</td>
</tr>
<tr>
<td>System Quality: Reliability</td>
<td>4</td>
<td>91</td>
</tr>
<tr>
<td>Information Quality: Understandability</td>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>Information Quality: Usefulness</td>
<td>4</td>
<td>94</td>
</tr>
<tr>
<td>Information Quality: Reliability</td>
<td>4</td>
<td>91</td>
</tr>
<tr>
<td>Information Quality: Completeness</td>
<td>4</td>
<td>75</td>
</tr>
<tr>
<td>Information Quality: Timeliness</td>
<td>4</td>
<td>97</td>
</tr>
<tr>
<td>Service Quality: Responsiveness</td>
<td>4</td>
<td>66</td>
</tr>
<tr>
<td>Service Quality: Reliability</td>
<td>4</td>
<td>63</td>
</tr>
<tr>
<td>Service Quality: Empathy</td>
<td>4</td>
<td>78</td>
</tr>
<tr>
<td>Service Quality: Assurance</td>
<td>4</td>
<td>34</td>
</tr>
<tr>
<td>Service Quality: Training</td>
<td>4</td>
<td>97</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>68</strong></td>
<td><strong>83</strong></td>
</tr>
</tbody>
</table>

Table 2. Item Placement Ratios

4.2 Test trial

In the next stage, we combined the initial item pool into a draft survey instrument for pre-testing. The questionnaire therefore had to be translated in the target language of the companies that will participate in the upcoming field surveys. The committee (or parallel) approach was used for this translation (Harkness & Schoua-Glusberg 1998). Three translators competent in the target language and in the skills required for survey work translated the questionnaire independently. At a reconciliation meeting, the translators compared the translations, reconciled discrepancies, and agreed on a final version that was considered the best of the independent translations.

In order to ensure the quality of the survey instrument design and presentation, we discussed the draft again with four IS experts and modified it according to their feedback. As a final pre-test prior to using the survey in the field, the draft instrument was trialed with a group of 20 information systems PhD students and faculty members serving as test users. The test users were asked to refer to their institute’s intranet application which acts as an employee portal when completing the questionnaire. One objective of this first field test was to confirm the results of the card-sorting and item-ranking approach as well as getting additional feedback from the respondents. The full model was estimated using Partial Least Squares (PLS) based on the sample data. Although the sample size of 20 is quite small, the model should possess sufficiently high statistical power to evaluate the reliability of the reflective indicators of the sub-constructs.

The indicator loadings of the items on their respective construct were trialed first to assess the quality of the scales. Items with a loading below .500 can be considered too unreliable (Straub 1989). Based on this criterion, we considered four of the items as candidates for deletion. Surprisingly, none of the four items were deletion candidates at the card-sorting and item ranking stage. In addition, the significance of the indicator loadings was tested using a bootstrapping procedure. Almost all of the loadings were significant at the .001 level. However, the loadings of all four items for ‘System Quality: Availability’ did not show significance on the commonly accepted .05 level. Thus, these items have to be modified for the final questionnaire.

Furthermore, we tested the model for convergent validity by analyzing average variance extracted (AVE), composite reliability (CR), and Cronbach’s alpha. AVE indicators above .500 suggested that
all of the factors possess adequate convergent validity. CR and alpha values for all scales were above the generally recommended level of .700 (Bruce & Gary & Terry Anthony 2005).

5 RESEARCH IN PROGRESS

Since the sample size of the first survey trial was very small, we will explore the survey instrument in a further field test. Data will be obtained from the employees of an international development organization through a web-based survey. The organization employs about 10,000 staff in more than 120 countries. In the first stage, the survey will be conducted at the head office, where around 1,000 people are employed. A non-probability convenience sampling method will be used. Even with the disadvantages inherent in a non-probability sample, it is deemed the most suitable method to achieve a response rate that is as high as possible.

Following the validation guidelines of Straub et al. (2004) and Lewis et al. (2005), we will test the measurement model for reliability, nomological validity, convergent validity, discriminant validity, and predictive validity. We will apply commonly employed decision rules. Given an adequate measurement model, the structural model will be analyzed to test the associations hypothesized in our research model. In order to further test and challenge the model, additional field studies will be conducted in several organizations in different countries. Furthermore, a benchmarking study will be conducted on the basis of the validated model.

6 SUMMARY AND CONCLUSIONS

In our review of the IS success literature, we found no study aiming to comprehensively evaluate employee portals. Based upon the D&M IS Success Model, we propose a model for measuring employee portal success. Furthermore, we present our steps for the development and validation of our survey instrument. Research methodology for further developing and validating the model is proposed.

Our research presented here is limited in that the proposed model is merely based on an extensive literature review, on our experiences, and on expert feedback. The results are the foundation for future empirical work in this area. First validation steps are presented in this paper. However, the model needs further development and validation before it can be applied in practice. Future research within this long-term research study will focus on empirically validating the conceptual model and on applying it in practice, as described in this article.

Having finished our research, we believe that our model together with the survey instrument will be valuable for practitioners who want to evaluate employee portals in their organizations. Furthermore, we hope to expand the IS success literature by presenting a complete validation of the updated D&M IS Success Model in the employee portal context.

References


EXPLORING THE IMPACT OF REAL-TIME COMMUNICATION ON MEDIA CHOICE IN THE CONTEXT OF DISTRIBUTED WORK

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0448.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Collaboration, Computer-mediated communication (CMC), Distributed organization / teams, Information Growth</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
EXPLORING THE IMPACT OF REAL-TIME COMMUNICATION ON MEDIA CHOICE IN THE CONTEXT OF DISTRIBUTED WORK

Aleksi Aaltonen, Ph.D. Student
v.a.aaltonen@lse.ac.uk
The London School of Economics and Political Science

Benjamin Eaton, Ph.D. Student
b.d.eaton@lse.ac.uk
The London School of Economics and Political Science

Abstract

This paper presents an exploratory study into the use of real-time communication (RTC) systems to support distributed work. Motivated by the authors’ dissatisfaction with theories assuming individually rational actors, the paper suggests the idea of informed presence to capture phenomena emerging from employees’ increasingly computer-mediated engagement with their work environment. Four case vignettes are presented to illustrate different communicative strategies that develop in response to presence availability updates generated by RTC systems. Drawing from Goffman’s microsociological idea of interaction order and Zuboff’s seminal work on computer-mediated work, the findings indicate the limitations of approaches such as information richness theory in understanding real-time communication in organizational settings. More research is needed to elaborate the implications of informed presence on distributed work and coordination of knowledge workers.

Keywords: communicative strategy, distributed work, presence, real-time communication, RTC
1 INTRODUCTION

This paper presents an exploratory study into the use of real-time communication (RTC) systems to support distributed work. Real-time communication systems are a distinct type of computer-mediated communication tools emerging from the convergence of telecommunications and groupware systems. The applications typically combine a number of synchronous and asynchronous channels such as instant messaging, voice over IP calls, conferencing, and file transfers in addition to incorporating presence availability information (Frößler and Klein 2006). Systems such as Skype, IBM Lotus Sametime and Microsoft Office Communicator are already in use in numerous organizational settings and their proliferation is projected to continue in the coming years (Dewing 2008, Gantz et al. 2007, Gantz et al. 2008, Lyman and Varian 2003).

Previous studies on real-time communication in organizational settings have made several interesting findings regarding the function, content and style of communication. However, paid to understanding how these systems reframe the communicative arrangement employees find themselves at work. We argue that the material makeup of RTC tools differs from antecedent systems (such as email) to an extent that warrants re-examination. More specifically, the systems incorporate real-time presence information unlike any previous media. It is our contention that common approaches such as information richness theory (Daft and Lengel 1984) used to explain media choice are inadequate to understand these differences.

The paper puts forward an idea of informed presence as a way of capturing organizational phenomena revolving around shifts in the interactional arrangements. It also discusses emerging communicative strategies employees harness to cope with interaction in an increasingly digital working environment. We understand communicative strategies as the ways individuals choose a particular media and establish channels in order to communicate with others. The argument draws from the microsociological idea of interaction order (Goffman 1983, Knorr Cetina and Bruegger 2002) and the discussion on information as the substratum of organizing (Kallinikos 2006, Zuboff 1988). In order to substantiate our views we present four case vignettes representing distributed knowledge work, both in small and in large organizations in the information and communication technology industry.

Our contribution is primarily targeted for the study of distributed work and real-time communication in organizational settings, but we believe the findings could as well be useful in Computer-Supported Cooperative Work (CSCW) and Human-Computer Interaction (HCI) research. The reader should, however, be aware of the tentative nature of claims made in the paper. In other words, we do not claim to provide unquestionable evidence for our argument. The empirical data is drawn from two separate, ongoing PhD thesis projects that, nevertheless, have come to point to similar insights. This is at the same time promising but also introduces some methodological difficulties. Moreover, the authors’ shared dissatisfaction with media choice theories based on the assumption of individually rational behaviour was strengthened by our empirical observations and to that end we think it is useful to apply fresh ideas to the topic.

The paper is structured as follows. The next section provides a concise summary of existing literature surrounding real-time communication systems in organizational settings and outlines the idea of informed presence. The third section presents four case vignettes to illustrate emerging communicative strategies through which employees use technological information to be present to each other. The fourth section reviews observations on informed presence and elaborates the communicative strategies that the employees developed across the cases. Finally, we identify possible avenues for future work.
2 LITERATURE REVIEW

It might be assumed that developing communication genres such as instant messaging or voice over IP calls, loosely grouped together as real-time communication, make no significant difference vis-à-vis practices revolving around email. Work-related, interpersonal communication becomes quicker and possibly more efficient, but it is essentially more of the same. In this section we first briefly review extant literature on real-time communication and discuss presence information as a key affordance that sets RTC apart from email. Finally, we sketch informed presence as a perspective to capture organizational implications of real-time communications.

2.1 Literature on real-time communication

The current understanding of the use of RTC systems in organizational settings derives mainly from studies on instant messaging (Cameron and Webster 2005, Huang et al. 2007, Isaacs et al. 2002, Nardi et al. 2000, Quan-Haase et al. 2005). Instant messaging has been found to take place often in a polychronic situation in which the employee engages simultaneously in several different conversations. It can speed up the exchange of information and support new forms of collaboration, but instant messaging has not been found to eradicate social distance or status differences between people. Users typically perceive instant messaging to be more informal than email.

Studying individual communication channels in isolation tends, however, to provide an impoverished view on how technologies mediate social relationships (Licoppe 2004). In the age of internet, employees can choose from a diverse and expanding portfolio of tools and channels to get in touch with each other (Aducci et al. 2008). The variable geometry of work communication easily escapes studies focused narrowly on individual channels. The same employee can take care of equivalent tasks through different media and different employees can use the same media for a variety of different purposes. Most importantly, RTC systems overlay this diversity with situational metadata intended to support individuals in their attempts to instigate and make oneself available for mediated communication with others.

RTC systems provide a status indicator and a short mood message that are used to convey one’s immediate availability to other users. This sets RTC apart, for instance, from email (Quan-Haase et al. 2005). Even relatively simple RTC systems are known to generate awareness of remote colleagues ongoing activity in a number of ways (Riemer et al. 2007), and the sensation of mediated co-presence has been found to play an important role in distributed work (Frößler 2008, Schmidt 2002). Nardi et al. (2000) identified certain outeraction practices, in contrast to the content of the interaction, by which people reach out to others in order to manage their interactions. Outeraction encompasses negotiations about the availability and choice of appropriate media in order that individuals can communicate effectively. Mediated presence has been available in planned occasions using dedicated technologies such as videoconferencing systems. More recently, however, real-time presence information has become increasingly integrated into generic technologies, such as workstations and mobile devices, gradually turning it into a constantly active background condition. RTC applications are often configured to launch and set the user’s online status automatically.

2.2 Computer-mediated presence

Researchers in the field of Computer-Supported Cooperative Work (CSCW), Human-Computer Interaction (HCI) and virtual environments have studied presence awareness since the mid-1980s (Luczak 2002, Makropoulos et al. 2007). The concept refers to the phenomenon of employees effectively coordinating their work without making a cognitive effort to align their interdependent acts (Schmidt 2002). Under such circumstances people are able to take heed of each other’s unfolding activity while focussing on their individual tasks.
Several conceptual distinctions have been put forward to describe the phenomenon. *Presence* is the experience of being in and interacting with a place other than the one is physically located in, and *co-presence* refers to carrying out this process jointly with others (Schroeder 2006). *Connected presence* is the degree to which one’s relationships are mediated in environments in which presence and co-presence are experienced (Licoppe 2004). *Presence availability* refers to practices around the instruments such as the status indicator that are at actors’ disposal to signal their availability and eventually come together in the mediated setting (Frößler 2008).

It is, however, too limited to frame mediated presence merely as a question of efficient coordination. Whenever two or more people enter into each other’s presence they enact a distinctive social institution that microsociologist Goffman (1983) labelled *interaction order*. The basic sociability of humans entails that in order to cope with immediate availability to each other people has to enact a set of institutionalized rules how to co-produce an orderly encounter. Knorr Cetina and Bruegger (2002) observed a similar phenomenon in a technologically mediated, real-time environment of currency traders buying and selling currencies on financial markets. They found that market transactions were embedded in situations distributed across organizations, continents and timezones leading the authors to distinguish between the embodied *face-to-face presence* observed by Goffman and technologically mediated *response presence*. The awareness of other’s presence results in accountability of interaction, as participants are tied to a moral order knowing that everyone is aware of everyone else’s acts. Erickson and Kellogg (2000) describe systems that facilitate normative, immediate commitment to those who are present to us as socially translucent.

### 2.3 Media choice in the context of informated presence

Computer technology differs from most other machineries in that it does not simply automate tasks, but, crucially, renders organizational processes, events and objects visible in a way that opens up new avenues of acting upon them. Zuboff (1988) labelled the latter as informing in contradistinction to the former in her seminal study on computer-mediated work. Processed by complex and to degree unpredictable corporate information infrastructures, the importance of technological information in organizational matters is steadily growing (Ciborra et al. 2000, Kallinikos 2006). The widespread availability of RTC systems holds a promise of going beyond informated organizational processes and objects of work. In RTC systems the availability of employees, allowing them to engage with each other across space, is turned into technologically mediated information. Presence availability updates constitute “a difference which makes a difference” being thus potentially informative (Bateson 2000, Kallinikos 2006). This is what we call *informated presence*.

The ephemeral presence availability updates generated by RTC systems are not persistently stored and have very little semantic content. This may lead an observer to ignore them in favour of the content of interaction. For instance, information richness theory (Daft and Lengel 1984), which posits that some media are richer than others, and managers, by default, prefer richer media for communication, has serious difficulties analysing this type of informativeness. The theory understands the richness of a medium as a function of its capacity to transmit cues such as body language, voice tone, and inflection. Consequently, alternative theories have been put forward to explain media choice. Straub and Karahanna (1998) researched media choice based on actors’ prior knowledge of which media were available to the person being contacted. They suggested a theory based on task closure, so that individuals choose media that they know will likely result in completed communication sequences.

### 3 CASE VIGNETTES

In this section we present four case vignettes to illustrate our argument with communicative strategies employees deploy in the context of informated presence. In order to explore the topics outlined in the previous section we chose four organizational settings in the information and communication technology sector that had adopted real-time communication system for professional purposes. The
aim of the sampling was to generate rich data on practices revolving around organizational RTC while attributing less weight on the representativeness of our sample. We either knew each setting from our previous engagements or we carried out few days of participant observation to familiarize ourselves with them. The primary data corpus resulted from nineteen recorded and transcribed interviews, four to five in each setting, focusing on how employees get hold of each other and how they communicate in the context of distributed work. The nature of work in each case was essentially collaborative. Individuals had to actively coordinate their activities on an ongoing basis with remote others.

Two of the four case studies took place at multinationals, TelCo and AppsCo, both of which had deployed the same corporate RTC system. Since the companies employed tens of thousands of individuals, the interviews focussed on particular teams. IntStart, in contrast, relied by and large on freely available Internet applications. The employees used Skype as well as other RTC applications largely driven by their individual preferences resulting in a remarkably heterogeneous environment. PeriphCo employees had installed Skype on their workstations on the initiative of the company head. IntStart and PeriphCo organizations were radically co-located (Olson et al. 2002) so that their mediated communication was generally oriented outside the company while at TelCo and AppsCo a significant proportion of internal communication was mediated.

In descending order, the general perceived order of time spent on mediated communications activities in TelCo and AppsCo was email, followed by conference calls, instances of synchronous communication (voice calls and instant messaging) and then other modes of communications. All in all, there was within TelCo and AppsCo a significant variation in interviewees’ responses regarding whether they reported using instant messaging more than voice calls or vice versa. Whilst the employees of IntStart considered Skype as a necessary part of their communication toolbox, email and fixed and mobile calls were perceived to be their most important modes of communication. PeriphCo had adopted Skype as the primary communications system and made an ample use of its features including instant messaging, voice and conference calls, file transfers and text messages to mobile phones. Email and other modes of communication were still used in addition to RTC. Despite variation between the entire portfolio of tools and their relative importance between the cases, real-time communication was an integral part of communicative practices in each case.

3.1 Case TelCo

TelCo is a large, traditional telecommunications operator offering consumer and corporate services in several countries. The interviews took place in a team developing solutions for mobile work in large corporations. Informants were a mix of home and office workers whose work required keeping in touch with other sites and occasionally spending time with customers. There were lots of commonalities in the use of mediated communication between TelCo and AppsCo. Amongst those interviewed, email generally remains the primary mode of business communications. It was felt that the asynchronous nature of email suited busy people, because they could respond on their own terms. Email was also perceived to allow an audit trail and facilitate the management of self-expression. The fact that email hands control to the receiver in terms of response meant, however, that questions could be skirted around, misconstrued and manipulated.

In comparison with the other cases, hierarchical control culture came most clearly through in the TelCo interviews. For instance, the informants reported using fixed line communications whenever possible, as the use of mobile was monitored to control costs. Managerial control was also implicated on online status indicator in the corporate RTC system that was used to give the impression of being a good employee. The description of status indicator usage by a lower grade employee expressed a clear awareness of his relative position:

“I’ve never set my presence to ‘Do not disturb’. […] I do see people who have done that and generally it does put a bold message up to say ‘I cannot be contacted’. But I don’t like to do that. I think its fear in a way. I don’t really want my assignment managers or senior
people looking at it and saying ‘Why doesn’t he want to be disturbed, what’s so important that he doesn’t?’ [...] I guess sometimes I have logged out of [RTC application] instead of having it as do not disturb.”

Calls were thought of being more persuasive and suitable for debate and asking someone to get something done than email. Often misunderstandings and ambiguity that occurred through email had to be cleared up by a phone call. The dialogue enabled by a voice call was perceived to provide more control over ambiguity compared to other modes of communication. Calls were typically made from deskphones instead of the RTC application that was used mainly for simple instant messaging exchanges. Instant messaging was often perceived to be best suited for exchanging short bits of time critical information. It was also treated as urgent when received, which generally accorded a quick response. A time-pressed project manager described his use of instant messaging:

“I probably use it maybe at least a few times a day with various different people. It does depend on what work you’re doing at the time. IM tends to be used at the early stages of projects as well and subsequent meetings. Because you’ve got lots of unknowns and you’ve got lots of questions that you have to ask. To save time [...] maybe send out a quick IM [...] It’s very quick, maybe focused questions, but then as the project progresses and the maturity of understanding increases, you probably tend not to have those IM conversations until later on, when things are getting a bit close to completing then you might want to just double-check, you’re finalising what you’re doing at work people are, the wrap up type of stuff.”

Informants noted, however, that busy people avoided taking calls, as it was felt that asynchronous modes such as email gave them more control as to when they responded to instances of communication. The majority of employees had incoming calls routed via their virtual number and few advertised their mobile numbers as their preferred point of contact. The fixed number for contact at home was only given to a select few. The preference for voice calls coupled with the sporadic availability of people to take up calls produced, interestingly, a specific practice around RTC application.

“I quite often, as behaviour, might look on IM to see – track down where somebody is, before I decide whether I ring them or not.”

Despite the more limited use of RTC for the actual interaction at TelCo, it was useful as an outreaction device for assessing and negotiating recipient’s availability in other channels. Individuals were clearly conscious of the availability status they displayed and they said to respect the status of others. In other words, even in the setting in which RTC was used in such a limited manner, being in the system demanded attention to the presentation of self and the others through presence availability information. This was the case both with light and heavy users of RTC.

3.2 Case AppsCo

AppsCo is one of the world’s leading software development companies producing a range of office productivity applications. The interviews took place within the team concerned with the product management and marketing of these applications within the UK. Whilst the majority of interviewees were based in the same office, projects within the company often drew together people from various domestic and international sites. In contrast to TelCo, the company is less hierarchically controlled and encourages flexible working, so many employees often spend a day or two a week working from home. The widespread use of RTC system was implicated in many telecommuting practices. The system was perceived to engage employees with the work environment to a significant degree as is illustrated by the following comment by an administrator within AppsCo:

“If I do take my laptop home, it’s normally because I need – you know I’m waiting for something - a call. So then I’ll just check it a period in the evening. [...] But I do try and switch [RTC application] off because it’s just too tempting for other people to contact you.

Although sometimes you get contacted and it’ll be by somebody saying “what are you doing working at this hour?”
Since AppsCo was a multinational with employees working in different time zones around the globe, there was a realisation that switching on the computer in the evening and becoming present on the corporate RTC resulted easily in being drawn into work with colleagues in other settings. Instigating calls was much less popular at AppsCo than in TelCo as many employees preferred instant messaging over voice calls that were viewed as the least effective means of communication within the company. Interestingly, while people were often simply too busy to pick up the call, email was considered by some too slow. Online presence had a clear impact on media choice as in the TelCo case, but with quite different outcomes.

“If they’re online, in fact, invariably if somebody is online, I won’t send them an email I’ll IM them. I’ll just click on the globe and instead of sending them email, I’ll IM them instead, generally.”

Voice calls were said to be made most often when away from the desk, contacting a person for the first time or reaching people external to the company - especially when handling a sensitive issue. Much like TelCo case, it was felt that voice calls, along with instant messaging, had the advantage of more control, the feedback from intonation and other contextual clues help cut through the confusion that can occur when an email veers off track. There was a sort perception of information richness related to voice calls, but this was not always an advantage. Employees often preferred instant messaging because they found conversations over this medium more succinct avoiding the pfaff of having to waste time going through social grooming in a channel saturated with unnecessary contextual information. The synchronous nature of instant messaging nevertheless provided more control over the course of communication.

“Before, there was a lot of email so my emails have reduced and my instant messaging traffic has gone up. [...] There are probably slightly more [transactions] because they are shorter and probably because you would break up one - would’ve been one traditional email thread, be it perhaps many smaller interactions. The topics, maybe a number of things that you’ve discussed, might be constant. [...] It’s chopped up into smaller, more nimble interactions that we try to communicate with bigger single transactions on email. So the fidelity improves because you get more chances to correct the direction of your conversation rather than once a day, or how often you respond to email.”

The shift from email to instant messaging was perceived to divide communication into ever-smaller segments. This enabled to tap into smaller and smaller in-betweens in the busy managers schedules – often facilitated by presence availability updates. A consensus appeared to be an emerging in terms of which communications mode should be used based on the availability status of the recipient. For instance, when the availability status of a correspondent was set to “In a meeting”, a common strategy was to approach them with an instant messaging. The awareness that that a person is going to a meeting by underground without network coverage in London enabled to send SMS message to be delivered within those few free minutes when the person appears overground and walks to the meeting. This kind of communication needs to be made of very concise messages in order not to overwhelm the recipient. Tapping into such bits and pieces of information may not seem relevant, but, for the following informant, it was precisely what he needed to keep things rolling.

“A lot of the projects I do have lots of dependencies on other people. But the dependencies are quite small. If I just need an answer that’s ‘yes/no’, ‘it’s £27’, ‘here’s a code’, ‘speak to this person’, waiting for everyone to answer that in their own email schedule really can divert your day. And you can wait two or three days for someone to reply to an email saying ‘I’m not the right person to speak to’.”

The bulk of email processing was generally done in the company first thing in the morning and late in the working day. The responses were often received at the end of the business day and it was felt that it was easier to ignore email than other modes of communication such as instant messaging. Waiting all day for a reply from a colleague, whose presence availability you could observe, felt too long. There was also a view amongst some interviewees that an instant messaging dialogue was less intrusive than calling or emailing someone, as it did not require one hundred percent attention. In
TelCo and AppsCo where many people sat in the meetings most of the day it enabled concurrent working in environments such as meetings and conference calls and, as in special cases, backchannel chats between subsets of people engaged in a meeting. In contrast to TelCo, the team interviewed at AppsCo represents corporate heavy users of RTC. Resulting from the different organizational setting the practices revolving around RTC were different, but they equally drew from the affordance generated by presence information.

3.3 Case IntStart

IntStart is a social enterprise founded in 2006 by five co-founders (the informants) with a mission to introduce a radical innovation in the educational sector while aiming to be financially profitable at the same time. The company develops and operates an internet-based platform to facilitate informal learning by matching individuals who are interested in teaching and learning various skills. In a stark contrast to corporations like TelCo and AppsCo, people were relatively free to use whatever tools available in Internet to get their work done at IntStart – including using Facebook and a range of consumer instant messaging systems for work-related communication. While the informants reported a number of idiosyncratic communication patterns and tools, email, both fixed and mobile phone, and Skype were used by everybody.

“It [Skype] is a sort of necessity rather than hey this is a really good idea, let’s get on and do it. It is kind of I need to, I have to, I have no choice.”

Having originally worked from various locations around the UK without a physical office space, the employees valued highly physical presence afforded by the newly acquired office. Resulting from this the importance of mediated communications in running the daily operations had diminished. Great deal of the daily activities revolved around the two big tables the company occupied in an open-plan office hosting several other organizations as well. In terms of advancing their common endeavour, being radically co-located at the same physical space was seen crucial. In order to enable the team to come physically together, the company relied, however, on distributed forms of work and RTC tools.

The informants reported having used real-time communication extensively before they had acquired the office space. Nowadays, RTC was still being used whenever the team members were working remotely or were travelling to meet different stakeholders. Status indicator was then used to signal availability between the remote team member and the core team at the office providing a sense of social cohesion. Long RTC conversations could emerge when distant team members were working simultaneously on the same software development issue or document. Questions about status indicator triggered elaborate accounts on how the informants interpreted each other’s presence information.

“A lot of people don’t respect other people’s business [status indicator state] very much. I like to think I do, but I think that it means that then you ask someone if they can talk, and they don’t have to reply if they don’t want to.”

Individual employees had different habits with respect to how they set their online presence. These ranged from some of the employees being virtually always online whenever their computer was turned on to others preferring to hide their presence availability status. The very possibility of broadcasting and having others observe one’s online presence was, nevertheless, something that employees had to relate to. Staying out from the digital work environment was difficult and, as in other cases, employees felt ambivalent towards the demands computer-mediated presence availability in the local setting.

“Now, partly because [colleague name] is at home, the others maybe not around, and we are using Skype a bit more internally for a sort of pinging documents, doing things like that. I just kind of leave it on. But I hate it because what immediately happens is that I then go off to have a meeting other side of the desk and hour later I found someone tried to Skype me because I forgot to change my status.”

Finally, the nature of start-up organization introduced an interesting twist to use of RTC and computer-mediated presence. The co-founders of IntStart had been and to some degree still are
simultaneously involved in other organizations. Real-time communication has enabled them to share their individual contributions – while sitting together at the office – across several organizations providing vital personal income before the company secured funding. In the case, presence information penetrates into the work of radically co-located team.

3.4 Case PeriphCo

PeriphCo develops computer peripherals such as headsets and webcams, and sells them to distributors throughout the world. The company is a subsidiary of an Asian manufacturing company that has its factory in Mainland China. Located in Northern Spain, PeriphCo office employs 7 people in a relaxed atmosphere. The informants included a product development manager, IT manager and the head of the company. In the fast-moving, highly competitive industry, everyday operations require coordinating projects with product designers, distributors and the main factory across the continents.

At PeriphCo Skype had become the preferred tool for many daily tasks. Employees used instant messaging, voice calls, file transfers and conference calls as part of their everyday tasks to interact with partners in Asia, Europe and North America. Email and mobile phones were still used, but the office had given up desk phones and all the calls were made using Skype. Informants also reported using Skype at the office to talk to a person next room. Voice calls and instant messaging were often used simultaneously, for instance, to clarify details, overcome language barriers and to leave a trail of what was said. The employees used the status indicator in Skype, but just like in the other cases they were ambivalent about it’s meaning because people did not respect it consistently. Mediated presence availability was felt to change the environment in which business was being done. In addition, the value of the system was perceived to increase as more partners adopted it. The head of the company explained how the RTC system intensified certain business relationships.

“I think the most interesting change I find about Skype as a business tool is that you find that when you become a Skype adopter you gravitate towards other people that use Skype. So that business I do, is more intense with other people that use Skype.”

PeriphCo had been evangelizing Skype for its partners as a preferred way of keeping in touch and took some pride in making others to use the system. Given the small number of employees, most of the people in the contact lists were from outside the company. In addition and in contrast to TelCo and AppsCo, employees did not perceive significant differences in how they used RTC with people from inside and outside the company. The fact that mediated presence gave information concerning the availability of others remote from the company offices, sometimes led to local matters taking second place to those at a distance. People had to occasionally disconnect from the system in order to focus on the local task at hand.

“The thing that I did not expect from Skype was that it is another level of ownership I did not expect. Even more so you become owned by the people you work with every day. So [partner name] owns me a little bit more because he knows when I am online.”

Whilst presence availability updates were open to interpretive flexibility, the interpretations had a common orientation: the current situation of an omnipresent other. Like IntStart employees, the informants sometimes reported using RTC to approach others even if they did not seem to be online, as they knew that some people had the tendency to hide their presence when online. Despite the interpretive dance around presence information, people seemed to have no choice but to dance.

“Getting to a buyer through a telephone is quite offensive, it is in English called a cold call. But send them a message over a Skype. You see they are online. They know that you know that they are online so they generally have to respond to it, but they can respond ‘I am busy now can I call you back later’, but at least you get a response.”

In the last excerpt, the head of the company describes, almost word-by-word the accountability mechanism discussed by Erickson and Kellogg (2000). Awareness of the other introduced an obligation, that, in turn, could be harnessed to elicit at least minimal response. Much like in TelCo and
AppsCo, the employees of PeriphCo felt generally that RTC offered more flexibility in terms of choosing and negotiating the right time and channel for interaction, both in terms of the task content and in terms of the situation of receiver.

4 SUMMARY OF FINDINGS AND DISCUSSION

The minuscule digital events informing who is online, in a meeting, at the airport and so forth, both frame employees' attempts to get hold of each other and make their engagement in the work environment visible in a new way. In all of the cases, real-time presence information mediated the engagement of employees with their work environment resulting in different performances depending on, for instance, management culture, spatial arrangements and the type of work. The event of becoming available, change in user’s online status, could perhaps be perceived as a question: "Should I contact this person now when he is available?" Knorr Cetina and Bruegger (2002) argue that the flow of real-time information from the currency traders’ environment was not merely a resource for traders’ decision making, but a self-perpetuating mechanism exciting the market itself. While the empirical evidence does not lend itself to generalizations, the case vignettes provide some insights in respect to study of RTC in organizational settings.

4.1 From media choice to co-configuration of channels

The case vignettes illustrate how different channels were co-configured and switched within broader patterns of mediated communication. Even in the case of TelCo with most limited RTC use, the application was used as an outeraction device to harness other channels appropriately. When instigating mediated interaction, different channels such as instant messaging was often used to negotiate the most appropriate time, for instance, for a call. There was no simple choice between clearly separate channels. Our study adds to the body of literature (e.g. Lee 1994, Markus 1994, Ngwenyama and Lee 1997) presenting contrary evidence to information richness theory that seems inadequate for understanding media choice in the context of real-time communication. Individuals perceived some media occasionally richer than others, but this seemed to have little to do with their choices between different channels.

Employees managed their communications in order to maintain control, maximize the likelihood of achieving their communications objectives and to behave appropriately. Whilst many media choices appear to be consistent with the goal of task closure as posited by Straub and Karahanna (1998), there was also evidence that this is not always the case. For example, it was reported by interviewees that email communication is sometimes favoured as a means of response as issues can be skirted around, misconstrued and manipulated. Rather than focussing on task closure, it would appear that mediated communications may sometimes have other communications objectives, such as task obfuscation. In a similar manner with Knorr Cetina and Bruegger (2002) the order of interaction we observed in the context of mediated presence was not fully amenable to strategic action.

4.2 Obligation to observe the other’s situation

The employees observed each other’s presence and were aware that their own status indicator was being observed. The information was used to figure out what distant others were doing and to save time trying to get hold of someone. If a potential correspondent was freely available, then it was felt that the initiator of communication has a free choice as to which mode of communications to use. There was also evidence that it provided a sense of social cohesion. Presence information was not, however, just a resource for decision-making regarding media choice. It generated a moral obligation between people who shared their online presence to acknowledge each other at some minimum level.

Synchronous communication was perceived to provide the instigator with more control as long as they connect with their destination and, given the awareness of each other’s presence, to impose an
obligation to reciprocate on the side of a recipient. Asynchronous communication was perceived to hand control to the receiver so that they could choose when to get back and how to interpret the message. In terms of receiving communication, interviewees were concerned with being communicated with on their own terms. They often preferred to be communicated with using asynchronous forms that are easier to filter and manage and could be responded to when it was convenient. The appropriateness of handing over or retaining control over the interaction is necessarily a complex issues that transcends any particular communication technology.

5 THE WAY FORWARD

This paper put forward the idea of informated presence as a means to understand organizational implications of real-time communication. Using this lens we analysed four cases and found strategic approaches such as information richness theory and task closure too simplistic in their underlying assumptions to understand mediated situations in which participants enact a form of interaction order through computer-mediated presence information. Employees were not just using presence availability information, but were, to some extent, overwhelmed by it.

Informated presence could be defined as reactions, reflections and orderings that develop out of presence availability updates in organizational settings. In all of the cases it appeared to affect the way interviewees went about their mediated communications. Informated presence forced increasing reflexivity in respect to the recipient’s ongoing situation. Theoretically, the concept is an attempt to synthesize Zuboff’s concept of informing with Goffman’s interaction order. Although this is by no means a modest task, we hope that by describing some of the communicative strategies emerging in the context of informated presence the current study has illustrated the potential of such synthesis.

The limitations of the research design and the scope of study mean that a lot more work remains to be done. It is, for instance, interesting how the findings, so far, resonate with the observations made by Knorr Cetina and Bruegger (2002) on global microstructures. This suggests that further work studying the interplay between real-time communication and emerging coordination mechanisms underpinning distributed work environments could help to elaborate the underdeveloped role of technology in contemporary organization theory (Orlikowski and Scott 2008, Zammuto et al. 2007). Chunking work-related, interpersonal communication into smaller and smaller instances that are punched opportunistically into every possible in-between in employees’ schedules may, at the same time, solve certain coordination problems while making it increasingly difficult to lay out work as a temporally planned succession of activities.

References


Proceedings ECIS 2009


Markopoulos, P., B. de Ruyter and W. Mackay (2007) "Introduction to This Special Issue on Awareness Systems Design", Human-Computer Interaction, 22 pp. 1–6.


Mitigating Response Distortion in IS Ethics Research

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0004.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Behavioural theories, Consumer behavior / choice / demand / empowerment / reviews / consumerism, Ethics, Empirical study</td>
</tr>
</tbody>
</table>
MITIGATING RESPONSE DISTORTION IN IS ETHICS RESEARCH

Abstract

Response distortion attributable to a variety of human motivations has long been a recognized problem for behavioral research relying on self reports by individuals. Researchers studying unethical IS behaviors usually need to solicit self reports because of the secrecy of such behaviors. Unfortunately, the unethical nature of those behaviors often subject self reports to various response distortions such as socially desirable responding. This paper discusses the method of psychometric adjustment for response distortion and empirically examines response distortion due to socially desirable responding in a software piracy research. The boundary conditions of psychometric adjustment are then discussed in depth and the use of randomized response technique, an alternative to mitigate response distortion, in IS ethics research is highlighted.

Keywords: response distortion, socially desirable responding, software piracy.
1 INTRODUCTION

With the rapid development and proliferation of information and communication technology, we are witnessing a growing impact of information system (IS) on our daily lives and social behaviors. More and more concerns pertaining to the ethical uses of IS have come to light. Two decades ago, Mason (1986) identified privacy, accuracy, ownership and accessibility of information as four ethical issues of the information age. These issues are more relevant today than twenty years ago and touch on a wide spectrum of behaviors regarding the ethical (or unethical) uses of IS. With the widespread use of the Internet, many issues such as inappropriate use of IS (e.g., “cyberslacking”) and illicit behaviors enabled by the Internet (e.g., on-line gambling or pornography) are emerging.

Conceivably, individuals who engage in unethical IS behaviors may seek to conceal their behaviors as much as possible because such behaviors may be embarrassing or even subject to legal sanctions in certain cases. Unfortunately, behavioral researchers may sometimes have to solicit self reports of such behaviors from individuals in order to investigate these behaviors in greater depths. Ironically, the trustworthiness of such self reports can often be challenged and this has impeded the scientific investigation of many unethical IS behaviors.

In fact, the general problem of response distortion has long been recognized among behavioral researchers (Himmelfarb 1993). Namely, response distortion occurs when the answer provided by a respondent does not accurately reflect his/her genuine opinion, belief, feeling, intention or behavior. Apparently, such distortion would be more serious when a respondent finds the question embarrassing or when the response might incur legal liability (Locander et al 1976), as are the cases of many unethical IS behaviors.

The existence of response distortion poses serious threats to behavioral research on at least two dimensions. First, researchers are forced to study less relevant proxy variables because the variables of primary interest may be extremely sensitive and are especially inclined to elicit response distortion. Such difficulty has resulted in a general lack of research in areas where objective investigations are most needed. Second, response distortion threatens the conclusion validity of the findings. Substantive conclusions could be biased if response distortion is not taken into consideration.

For example, consider behavioral research on software piracy. Software piracy is an unethical IS behavior estimated to cost the industry over $31 billion a year (BSA 2005). Over the past 20 years, it has attracted a continuous interest among the community of IS researchers. Due to the secret nature of the piracy behavior, many researchers studies the topic using self-reported data solicited from individuals. It is not uncommon for researchers to acknowledge response distortion as a limitation of their empirical findings (e.g. Christensen and Eining 1991; Taylor and Shim 1993; Sims et al. 1996; Cheng et al. 1997; Seale 2002; Limayem et al. 2004; Moores and Chang 2006). However, little effort has been expended to resolve this problem.

The current study uses an empirical study on software piracy to explore the existence of response distortion in IS research on unethical behavior. We outline a methodological approach based on psychometric adjustment in order to cope with response distortion. We will examine the extent to which research conclusions may be misguided if response distortion is not properly accounted for. Furthermore, the boundary conditions of psychometric adjustment will also be discussed in the hope that such approach can be adequately applied in future IS research on ethical or sensitive issues.

---

1 Cyberslacking refers to the practice of employees using the Internet for leisure during work hours.
2 RESPONSE DISTORTION DUE TO SOCIALLY DESIRABLE RESPONDING

The motivations behind response distortion are multifaceted. According to Himmelfarb (1993), people hide their true responses “to protect their privacy, to avoid legal prosecution, to gain economic advantage, to obtain social approval and avoid social disapproval, and to project or protect particular identities” (p. 72).

Among the many forms of response distortion, socially desirable responding has received much attention in research on response distortion in self-reported data (Kline et al. 2000). Socially desirable responding is defined as “responding to items more as a result of their social acceptability than their true feelings” (Podsakoff et al. 2003 p. 882). Ganster et al. (1983) considered socially desirable responding as “a tendency for an individual to present him or herself, in test-taking situations, in a way that makes the person look positive with regard to culturally derived norms and standards” (p. 322).

More specifically, respondents are likely to underreport socially unacceptable opinions, attitudes, and behaviors while they over-report socially acceptable ones. Socially desirable responding is mainly driven by the need for social approval (Crowne and Marlowe 1964), which belongs to one of the motivations of response distortion as discussed by Himmelfarb (1993). In the context of IS ethics research, socially desirable responding may lead the respondent to over-report ethical IS behaviors (e.g. observing the acceptable IS use policy of the company) while underreport unethical IS behaviors (e.g. breaking privacy regulation of the company). As such, it is reasonable to expect that socially desirable responding is an important source of response distortion in IS ethics research.

2.1 Psychometric Measurement of Socially Desirable Responding

Personality psychology researchers have been studying socially desirable responding since the seminal work of Hartshorne and May (1928). Based on the Minnesota Multiphasic Personality Inventory (MMPI), Edwards (1957) provided strong empirical evidence that respondents had a high tendency to admit self-descriptions they regarded as desirable, suggesting that the MMPI score was likely to be contaminated by socially desirable responses. Crowne and Marlowe (1964) further proposed that socially desirable responding is driven by “the need for social approval and acceptance and the belief that it can be attained by means of culturally acceptable and appropriate behaviors”. (p. 109) The authors managed to assemble a collection of 33 behaviors that are either “desirable but uncommon” or “undesirable but common”. These behaviors constitute the M-C SD Scale, a psychometric instrument aiming at measuring a respondent’s propensity for socially desirable responding.

Paulhus (1984) argued that there are in fact two dimensions of socially desirable responding: self-deception and impression management (also known as other-deception). The former refers to the unconscious tendency for one to see oneself favorably while the latter refers to a conscious presentation of a false image, for the purpose of pleasing others. Paulhus (1984, 1991) developed another 40-item psychometric measurement scale for socially desirable responding, called the Balanced Inventory of Desirable Responding (BIDR), which clearly separates the self-deception dimension from the impression management dimension.

The sixth version of the BIDR scale (aka BIDR-6) published by Paulhus in 1991 is shown in the Appendix. BIDR-6 consists of a total of 40 items, with 20 of them measuring the self-deception dimension and the remaining 20 measuring the impression management dimension. All the items are on the possession of a variety of “unlikely virtues” such as “never cover up mistakes” or “always

---

2 The development of the “lie scale” by Hartshorne and May (1928) can be regarded as one of the early attempts to assess socially desirable responding among school children.
appreciate criticism”. The summated score of these items indicates the extent a respondent claims to possess these virtues. Since these virtues are desirable but unlikely, the score indicates the propensity of a respondent to engage in socially desirable responding. The summated score of BIDR-6 was found to correlate highly with the more traditional M-C SD Scale and had very stable test-retest reliability (Paulhus 1991).

2.2 Psychometric Adjustment for Socially Desirable Responding

The psychometric perspective on socially desirable responding posits that respondents differ in their propensities for socially desirable responding and these idiosyncratic propensities can be measured psychometrically. Measurement of idiosyncratic socially desirable responding is usually administered together with measurement of substantive variables of interest. This serves two main purposes. First, researchers can detect if socially desirable responding does exist in self-reported data by checking if there is a significant correlation between the socially desirable responding score and the variable of interest (e.g., Paulhus 1991). Second, if socially desirable responding is indeed detected, researchers may correct or adjust for it using various statistical techniques (Fisher and Katz 2000; Podsakoff et al. 2003).

A commonly known approach of psychometric adjustment applies to the correlation between two variables that are believed to be contaminated by socially desirable responding. The simplest form of adjustment can be accomplished by using the partial correlation controlling for socially desirable responding in place of the zero-order correlation between the substantive variables. A more sophisticated adjustment can be performed by adopting a measurement model that is confounded with measures of socially desirable responding. Podsakoff et al. (2003) provided an excellent summary on such methods.

Apart from adjustment on the correlation between two variables, it is also possible to adjust for socially desirable responding in individual variables. This can be accomplished by regressing the substantive variable of interest on the measurement of socially desirable responding. The resulting intercept term represents the mean estimate of the substantive variable with effect of socially desirable responding removed (Fisher and Katz 2000).

In the following, we will examine psychometric adjustment in an empirical study on software piracy.

3 Psychometric Adjustment for Socially Desirable Responding in Software Piracy Research

To illustrate socially desirable responding, we make reference to the work by Cheng et al. (1997). The authors attempted to identify factors that motivate people to use pirated software through a survey of 340 university students. Students were asked to score their perceived importance of nine commonly considered motivations of software piracy. Their results are shown in Table 1 (ratings range from 1 to 9 with a smaller value indicating higher importance).

Cheng et al. (1997) were careful to employ the indirect questioning method (Robertson and Joselyn 1974; Anderson 1978) in their study. Specifically, respondents were asked to rank the reasons that motivate “people”, not necessarily themselves, to use pirated software. Fisher (1993) provided empirical evidence that indirect questioning helps to reduce socially desirable responding. Also, Simon and Simon (1975) reported that responses solicited from indirect questioning are not totally objective. To a large extent, the responses include the respondent’s own thinking.

<table>
<thead>
<tr>
<th>Reason to Pirate Software</th>
<th>Rating of Importance</th>
</tr>
</thead>
</table>

Proceedings ECIS 2009
Table 1. Ranking of Reasons for Piracy (Cheng et al. 1997)

As shown in Table 1, some of the listed reasons appear to be more socially desirable than do others. Thus, it is plausible that socially desirable responding was not completely eliminated. For example, the reasons “software too expensive” and “can’t afford the software” appear to be more socially acceptable than “little chance of being caught.” Therefore, the tendency for the respondents to overrate the first two reasons and underrate the third reason cannot be ruled out. In fact, the relatively low importance of “little chance of being caught” is surprising and contradicts the findings from other research (e.g., Peace et al. 2003). It is possible that respondents with high propensities for socially desirable responding might have underreported the importance of the chance of being caught.

3.1 Methods

To probe further into the existence of response distortion, we conducted a survey similar to that of Cheng et al. (1997). A total of 612 university students (289 male and 323 female) from a major university in Hong Kong were recruited to participate in the study. Respondents were unaware of the topic of study at the time of recruitment so that bias due to sample selection was minimized. All respondents were requested to answer a set of questions anonymously.

Respondents were first asked to complete the 40-item BIDR-6 (see Appendix). The BIDR items were measured using a 7-point scale ranging from “Strongly Disagree” (1) to “Strongly Agree” (7). The summated score of the 40 BIDR items was used to indicate an individual’s propensity of socially desirable responding as it was found to correlate highly with the more traditional M-C SD Scale and had very stable test-retest reliability (Paulhus 1991).

Then, they were asked to rate the importance of the nine reasons for people to pirate software, as listed in Table 1. The reasons for pirating were rated with a range from 1 to 9 with a smaller value indicating higher importance.

3.2 Detection of Socially Desirable Responding

We detected traces of socially desirable responding in the correlation between respondent’s socially desirable responding (SDR) score and the corresponding ratings of piracy reasons as shown in Table 2.
Our results show that the socially desirable responding score does not have significant correlation with the rating of all reasons except Reason 7 (i.e., “little chance of being caught”) and Reason 8 (i.e., “most people I know copy software”). In both cases, the correlation is positive, suggesting that respondents with higher propensities for socially desirable responding indeed reported a higher rating (and thus lower importance) for these two reasons. Although the correlations (0.13 and 0.127) for Reasons 7 and 8 are only moderate, they are significant (p < 0.01) and may have an impact on the relative ranking of the listed reasons.

### 3.3 Adjustment of Socially Desirable Responding

To illustrate the sensitivity of the findings to socially desirable responding, the nine reasons were ranked again after adjusting for socially desirable responding based on the BIDR instrument. This is accomplished by regressing each rating score on the socially desirable responding score. The resulting intercept term represents the mean estimate of the rating with the effects of socially desirable responding removed (Fisher and Katz 2000).

An assumption behind this technique is that socially desirable responding is unrelated to the factors leading to one’s genuine judgment of a reason’s importance. Technically speaking, the simple regression of the importance score on the socially desirable responding score is subject to the problem of omitted independent variables (see Cohen et al. 2003 pp. 127). However, as we assume that socially desirable responding is orthogonal to the underlying factors affecting a person’s evaluation of the importance of a pirating reason, the estimate of intercept term is still unbiased (Greene 2003) by the assumption of orthogonality.

Table 3 shows the mean estimates and the 90% confidence intervals of the nine reasons before and after the psychometric adjustment. It can be seen that the ranking is quite different after adjustment for socially desirable responding. While Reason 1 (“software too expensive”) still maintains the top rank on the list, Reason 7 (“little chance of being caught”) rises to become the second most important reason for software piracy. Also, Reason 8 (“most people I know copy software”) advances to the fourth place after adjustment. The revised ordering is consistent with our expectation that respondents do not reveal their true responses.

<table>
<thead>
<tr>
<th>Reason to Pirate Software</th>
<th>Rating (Mean [90%-CI])</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cheng et al (1997)’s Study</td>
</tr>
<tr>
<td>Reason1: Software too expensive</td>
<td>1.90 [1.78,2.02] Rank 1</td>
</tr>
<tr>
<td>Reason3: Can’t afford the software</td>
<td>3.29 [3.08,3.49] Rank 3</td>
</tr>
<tr>
<td>Reason4: Only use it for a short time</td>
<td>3.31 [3.11,3.51] Rank 4</td>
</tr>
</tbody>
</table>

3 Apart from the technical perspective, we will also discuss more about this independence assumption from the theoretical perspective after our empirical findings are presented.
Table 3. Ranking of Reasons for Piracy

<table>
<thead>
<tr>
<th>Reason</th>
<th>Rank 1</th>
<th>Rank 2</th>
<th>Rank 3</th>
<th>Rank 4</th>
<th>Rank 5</th>
<th>Rank 6</th>
<th>Rank 7</th>
<th>Rank 8</th>
<th>Rank 9</th>
</tr>
</thead>
</table>

This study serves to demonstrate how socially desirable responding affects any substantive conclusions of software piracy studies based on self-reported data. It is observed that the slim likelihood of being caught is an important motivation to engage in software piracy behavior. However, our respondents did not directly admit to this behavior. A plausible explanation is that university students, being the more educated and elite class in the society, may feel embarrassed to admit that deterrence is an important consideration of whether to pirate.

The implementation of any anti-piracy strategy or policy is likely to incur substantial business, social and political costs. In order to formulate the most effective countermeasures against software piracy, an in-depth understanding of the importance of motivating factors behind piracy behavior is essential. For instance, the software industry has been calling for more stringent law enforcement whereas critics maintain that reduction in price may be more effective. Understanding whether people would be dissuaded from piracy by a reduction in software price or by more stringent law enforcement would provide valuable insights for reaching broad consensus among different stakeholders. Our results suggest that both price cutting and law enforcement would be effective and important countermeasures to combat piracy. Had we ignored the response distortion in our studies, we would have missed out an important motivation behind piracy and an effective countermeasure to the problem.

4 DISCUSSIONS

Our empirical study is based on a sample of university students in Hong Kong and thus the substantive findings should only be interpreted within this confine. Nevertheless, we highlight the methodological implications of ignoring response distortion in similar empirical studies. In the following, we further consider the boundary conditions of the approach used.

First, we assume that the propensity for socially desirable responding is unrelated to the underlying substantive variables of interest. It follows that the correlation between socially desirable responding and the variables of interest can be attributed to socially desirable responding alone, and thus can be safely removed. Otherwise, adjustment is unwarranted and will reduce the content validity and predictive power of the measured variable (Zerbe and Paulhus 1987). Paulhus (2002) commented on the importance of distinguishing measurements of socially desirable responding that focus on the “style” of the response (i.e., response style) from measurements that tap on the “content” of the response (i.e., relevant components of the variable in concern). Adjustment is not necessary in the latter case.

For instance, Fisher and Katz (2000) argued against adjustment for socially desirable responding on self-reported personal values because of a conceptual overlap between the “need for social approval”, the motivation of socially desirable responding, and the intrinsic values of a person. In our study, we were cautious about the existence of such a conceptual overlap when responses on personal involvement (e.g., evaluative variables such as attitude toward piracy) are solicited through direct questioning. For instance, if respondents feel that purchasing licensed software would make them
appear stupid before their peers (i.e., social disapproval), such a sentiment may also be reflected in their self-reported attitude toward software piracy. Adjustment on the self-reported attitude for socially desirable responding scores would therefore be redundant and would distort the true attitude held by the respondent.

Second, psychometric adjustment for socially desirable responding is not applicable in cases when respondents feel threatened (e.g., due to risk of prosecution) even though social sanctioning does not exist or is largely absent. Such situations are best exemplified by discrepancies between law and social norms. Nederhof (1985) pointed out that norms, instead of laws, are the most important determinants of socially desirable behaviors. His example to illustrate this point was gang members who admitted to more violations of the law than actually committed in order to impress others. In the context of software piracy, using pirated software may not be judged as undesirable in certain social groups. In this case, underreporting of piracy intention (or behavior) would not be motivated by socially desirable responding. Instead, it is triggered by the fear of legal sanctions.

In summary, response distortion in self-reports of software piracy behavior may not be entirely attributed to socially desirable responding alone and therefore psychometric adjustment may not always be effective.

4.1 Confidentiality Assurance to Respondents

In view of the limitations of psychometric adjustment, it would obviously be more desirable if motivations behind untruthful responding can be eliminated as much as possible rather than being adjusted for after the fact. Conceivably it serves as a more fundamental cure to response distortion than psychometric adjustment.

Traditionally, researchers soliciting self-reports on sensitive topics would provide confidentiality assurance to respondents so as to encourage truthful responding. Usually this is accomplished by convincing the respondents that a survey is completely anonymous or confidential. Over the years a number of good survey practices (e.g. physical separation of respondents, promises of confidentiality, emphasis on truthful responding rather than a “right” answer, etc.) has been identified for different types of survey administrations (see Paulhus 1991). Confidentiality assurance is a very powerful strategy in the sense that, if successful, should naturally eliminate most of the motivations behind response distortion (including socially desirable responding). The success of this strategy hinges on whether the assurance does look convincing to the respondent. When questions are considered threatening, credibility of such confidentiality assurance would be of prime importance as respondents may demand a more convincing assurance.

4.2 Randomized Response Technique and Research on Unethical IS Behaviors

Randomized response technique (RRT) is a questioning method that incorporates respondent confidentiality by design, and was developed to encourage truthful responding on sensitive topics. First invented by Warner (1965), various forms of RRT have been designed (Warner 1965; Greenberg et al 1969; Greenberg et al 1971; Himmelfarb and Lickteig 1982; Fox and Tracy 1986) to date for soliciting truthful responses to evasive or embarrassing questions. The core idea of RRT is to assure complete confidentiality of a participant’s response by contaminating it with a random “noise” value with known statistical properties.

For instance, in the unrelated-question randomized response model (Greenberg et al 1969), participants are instructed to choose among a set of questions to answer according to a randomizing device controlled by them privately (e.g. as simple as flipping a coin). Each sensitive question is paired with another unrelated and innocuous (i.e. non-sensitive) question such that the participants would answer either one of them depending on the outcome of the privately performed randomization procedure. As the researcher has no way to know exactly which question was answered, complete
confidentiality can be assured and this should in turn help eliminate any possible stigma or embarrassment caused by the sensitive question, resulting in more truthful responses. On the other hand, as the probabilistic property of the randomization procedure is known, useful aggregate properties of the answer to the sensitive question can be estimated using statistical methods. For instance, if coin-flipping was used as the randomization procedure, researchers would know that about half of the respondents should have actually answered the sensitive question and may derive useful research conclusions.

Originally, RRT was limited to two-choice questions only. It was later extended to cover multiple-choice questions as well as quantitative answers. As the majority of literature on RRT focuses on univariate analysis (e.g. proportion of people committing a certain criminal behavior, mean value of a certain sensitive characteristic, etc.), there is a common misunderstanding that RRT is limited to univariate analysis. This is in principle not true although more sophisticated statistical estimators would be required for multivariate analysis.

In particular, Kwan, So and Tam (forthcoming) demonstrated an innovative method to deploy RRT for correlational studies. By representing randomized responses in a mathematical model consisting of indicator variables, estimators for the mean, variance and covariance of randomized responses can be derived based on the method of moments. The method has been empirically tested using two empirical studies on software piracy using online survey data. The applicability and usefulness of RRT for studying unethical IS behaviors have been clearly demonstrated.

Comparative studies have shown that RRT leads to higher admission of sensitive characteristics. Validation studies deploying participants whose sensitive characteristics were known in advance have also revealed that RRT does outperform other techniques in soliciting valid responses to sensitive questions (Umesh & Peterson 1991; Scheers 1992; Hosseini & Armacost 1993).

However, there are a number of issues with RRT that must be noted. First, there is an additional sampling error due to the randomization procedure. This would inevitably reduce the overall measurement reliability and a larger sample size would normally be needed. As the loss of statistical power would depend on (i.e. decrease with) the probability of respondents answering the sensitive question, researchers have to trade off between a higher perceived confidentiality protection and a higher data collection efficiency (Fox & Tracy 1986). Hosseini & Armacost (1993) thus concluded that RRT should only be used for questions of sensitive nature. We believe that questions pertaining to unethical IS behaviors should belong to this category.

Second, the complexity due to randomization procedure may not be completely understood by respondents, especially those less educated ones, resulting in incorrect responses or missing data. Those respondents who do not understand the procedure well may not be convinced of the confidentiality protection that is built into the method, further undermining the benefit of RRT. In most research studies on unethical IS behaviors, it should be reasonable to assume the target population (i.e. IS users) should belong to the more educated class. We thus believe this should be an appropriate context to apply RRT.

Third, it should be noted that RRT is a procedural improvement (Podsakoff et al 2003) on increasing the credibility of response confidentiality. Even if RRT is successfully applied in a survey, there may still be untruthful responding although it is reasonable to assume the amount of distortion should be much less than that when direct self-reports are solicited. In fact, a recent RRT validation study conducted by van der Heijden et al (2000) compared the proportions of respondents admitting income fraud when surveyed using different questioning methods. The study revealed that RRT could help boost the percentage of truthful response to 43%, up from 25% with face-to-face direct questioning and 19% with computer-assisted self interview. Although it is not a complete cure to the problem of response distortion, this clearly represents a significant improvement.
5 CONCLUSION

We have empirically examined response distortion in self-reports based on an empirical study on software piracy. Despite the use of an indirect questioning technique to minimize socially desirable responding in our study, we still observed evidence of response distortion. Our findings substantiate the worries of many previous researchers about response distortion in software piracy research.

We demonstrated how substantively different conclusions about the motivating factors behind software piracy can be reached with or without psychometric adjustment for socially desirable responding. In particular, we found that the slim possibility of being caught committing piracy could be an important reason for piracy behavior although respondents would not admit this directly. Our findings demonstrate how response distortion could mislead researchers into erroneous conclusions.

We also discussed the limitations of psychometric adjustment for socially desirable responding. First, if the substantive variables of interest are actually related to socially desirable responding in theory, psychometric adjustment would be inappropriate because it would reduce the content validity of the measured constructs. Second, psychometric adjustment would be irrelevant when the anticipated distortion is not attributable to socially desirable responding. Lastly, we highlighted the use of randomized response technique for assuring confidentiality and mitigating response distortion when psychometric adjustment is inappropriate.

In conclusion, we find response distortion to be a real threat to research that relies on self-reports of unethical IS behaviors. Although our empirical study was conducted in the domain of software piracy, we believe the problem is also threatening research in ethical or unethical IS behaviors such as data privacy, cyberslacking, system hacking, online gambling and pornography viewing. Psychometric adjustment for socially desirable responding should help alleviate the problem in certain contexts but may not be useful in many others. We encourage researchers to employ RRT as a more fundamental remedy for tackling response distortion in research related to sensitive topics.

References


APPENDIX: BIDR-6

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Variable</th>
<th>Measurement Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Deception</td>
<td>sde1</td>
<td>My first impressions of people usually turn out to be right.</td>
</tr>
<tr>
<td></td>
<td>sde2*</td>
<td>It would be hard for me to break any of my bad habits*.</td>
</tr>
<tr>
<td></td>
<td>sde3</td>
<td>I don’t care to know what other people really think of me.</td>
</tr>
<tr>
<td></td>
<td>sde4*</td>
<td>I have not always been honest with myself*.</td>
</tr>
<tr>
<td></td>
<td>sde5</td>
<td>I always know why I like things.</td>
</tr>
<tr>
<td></td>
<td>sde6*</td>
<td>When my emotions are aroused, it biases my thinking*.</td>
</tr>
<tr>
<td></td>
<td>sde7</td>
<td>Once I’ve made up my mind, other people can seldom change my opinion.</td>
</tr>
<tr>
<td></td>
<td>sde8*</td>
<td>I am not a safe driver when I exceed the speed limit*.</td>
</tr>
<tr>
<td></td>
<td>sde9</td>
<td>I am fully in control of my own fate.</td>
</tr>
<tr>
<td></td>
<td>sde10*</td>
<td>It’s hard for me to shut off a disturbing thought*.</td>
</tr>
<tr>
<td></td>
<td>sde11</td>
<td>I never regret my decisions.</td>
</tr>
<tr>
<td></td>
<td>sde12*</td>
<td>I sometimes lose on things because I can’t make up my mind soon enough*.</td>
</tr>
<tr>
<td></td>
<td>sde13</td>
<td>The reason I vote is because my vote can make a difference.</td>
</tr>
<tr>
<td></td>
<td>sde14*</td>
<td>My parents were not always fair when they punished me*.</td>
</tr>
<tr>
<td></td>
<td>sde15</td>
<td>I am a completely rational person.</td>
</tr>
<tr>
<td></td>
<td>sde16*</td>
<td>I rarely appreciate criticism*.</td>
</tr>
<tr>
<td></td>
<td>sde17</td>
<td>I am very confident of my judgments.</td>
</tr>
<tr>
<td></td>
<td>sde18*</td>
<td>I have sometimes doubted my ability as a lover*.</td>
</tr>
<tr>
<td></td>
<td>sde19</td>
<td>It’s all right with me if some people happen to dislike me.</td>
</tr>
<tr>
<td></td>
<td>sde20*</td>
<td>I don’t always know the reasons why I do the things I do*.</td>
</tr>
<tr>
<td></td>
<td>im1*</td>
<td>I sometimes tell lies if I have to*.</td>
</tr>
<tr>
<td></td>
<td>im2</td>
<td>I never cover up my mistakes.</td>
</tr>
<tr>
<td></td>
<td>im3*</td>
<td>There have been occasions when I have taken advantage of someone*.</td>
</tr>
<tr>
<td></td>
<td>im4</td>
<td>I never swear.</td>
</tr>
<tr>
<td></td>
<td>im5*</td>
<td>I sometimes try to get even rather than forgive and forget*.</td>
</tr>
<tr>
<td></td>
<td>im6</td>
<td>I always obey laws, even if I’m unlikely to get caught.</td>
</tr>
<tr>
<td></td>
<td>im7*</td>
<td>I have said something bad about a friend behind his or her back*.</td>
</tr>
<tr>
<td></td>
<td>im8</td>
<td>When I hear people talking privately, I avoid listening.</td>
</tr>
<tr>
<td></td>
<td>im9</td>
<td>I have received too much change from a salesperson without telling him or her*.</td>
</tr>
<tr>
<td></td>
<td>im10</td>
<td>I always declare everything at customs.</td>
</tr>
<tr>
<td></td>
<td>im11*</td>
<td>When I was young I sometimes stole things*.</td>
</tr>
<tr>
<td></td>
<td>im12</td>
<td>I have never dropped litter on the street.</td>
</tr>
<tr>
<td></td>
<td>im13*</td>
<td>I sometimes drive faster than the speed limit*.</td>
</tr>
<tr>
<td></td>
<td>im14</td>
<td>I never read sexy books or magazines.</td>
</tr>
<tr>
<td></td>
<td>im15*</td>
<td>I have done things that I don’t tell other people about*.</td>
</tr>
<tr>
<td></td>
<td>im16</td>
<td>I never take things that don’t belong to me.</td>
</tr>
<tr>
<td></td>
<td>im17*</td>
<td>I have taken sick-leave from work or school even though I wasn’t really sick*.</td>
</tr>
<tr>
<td></td>
<td>im18</td>
<td>I have never damaged a library book or store merchandise without reporting it.</td>
</tr>
<tr>
<td></td>
<td>im19*</td>
<td>I have some pretty awful habits*.</td>
</tr>
<tr>
<td></td>
<td>im20</td>
<td>I don’t gossip about other people’s business.</td>
</tr>
</tbody>
</table>

Note: All items are in 7-point scale and those marked with * are in reverse. Unless specified otherwise, items are measured in 7 points from “not true” (1) to “very true” (7)
# Identity in Information Systems

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0514.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Identity, Literature review, Organisational identity, Organizational theory</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
IDENTITY IN ORGANIZATIONS: A REVIEW OF INFORMATION SYSTEMS RESEARCH

Gal, Uri, Aarhus School of Business, University of Aarhus, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark, urig@asb.dk
Kjærgaard, Annemette, Copenhagen Business School, Howitzvej 60, DK-2000 Frederiksberg, Denmark, amk@cbs.dk

Abstract

Research into the role of identity in organizations has become increasingly popular in recent years. Scholars from different organizational and management disciplines have applied the concept to address a wide variety of issues. In light of the concept’s broad applicability, we conduct a review of six prominent IS journals to examine the extent of the concept’s use in our field. Our review reveals that the concept is scantly employed by IS researchers, which, we believe, constitutes a missed opportunity. In light of this finding, we outline possible directions and opportunities for using the identity concept in IS research with the aim of enriching and diversifying the conceptual discourse in our field.

Keywords: Organizational Identity, Information Systems, Literature Review.
1 INTRODUCTION

Research into the role of identity in organizations has become increasingly popular in recent years (Albert et al. 2000; Corley et al. 2006; Cornelissen et al. 2007; Gioia et al. 2000; Haslam and Ellemers 2005). The interest in identity is quite diverse, reflecting perspectives that originate in the fields of organization studies, corporate communications, social and organizational psychology, personnel and human resources, and strategy and marketing (Cornelissen et al. 2007). Furthermore, theoretical and empirical research has explored identity and identification processes as well as their organizational outcomes at various analytical levels from personal to social to organizational (Ibid.).

The growing interest in identity issues reflects the concept’s centrality to the way scholars from different disciplines understand and theorize about organizations. In addition it underscores the importance and practical relevance of the concept to a variety of organizational areas including strategy (Dutton 1997), management and leadership (Gioia and Chittipeddi 1991; Pratt and Foreman 2000), inter-organizational collaboration (Beech and Huxham 2004), and corporate communication (Cheney and Christensen 2001; Schultz et al. 2000).

In light of the broad applicability and use of the concept in general management and organizational studies, its scant utilization in IS literature stands out. Accounts of the relationship between IS and identity in organizations have been sparse (e.g., Barrett and Walsham 1999; Lamb and Davidson 2005; Walsham 1998). This is surprising given the significance of identity to a variety of issues that have received considerable attention from IS researchers such as group and organizational sense-making (Weick 1995), the shaping of organizational practices and change (Corley and Gioia 2004; Gioia et al. 2000), organizational learning (Corley and Gioia 2003), and knowledge work (Nag et al. 2007).

Although identity research in management studies does not focus on technology in general or IS in particular, we believe that focusing on identity issues in organizations can help IS scholars to produce thoughtful and meaningful insights into individual and collective self-constructions in organizations and into the interactions between the implementation and use of IS, and organizational processes and outcomes. Motivated by this premise, we set out to review and characterize the use of the concept of identity in organizational settings in existing IS literature.

Based on this review we aim to point at possible directions for employing the concept in future studies and indicate how it can contribute to explanations of the organizational implications of and reactions to IS, as well as the interaction between IS and users of systems in organizational contexts.

Next we present the concept of identity, in particular with respect to its application in organizational settings. In section three we present the research methodology which is followed by the findings from the review of the use of the identity concept in IS literature. We continue by outlining possible avenues for future IS research employing the concept of identity, and conclude by emphasizing the versatility of the concept and encouraging researchers to make use of it in their future work.

2 THE CONCEPT OF IDENTITY IN ORGANIZATIONS

The increase in theoretical and empirical identity research in organizational settings can be attributed to the richness of the concept and the opportunity that it provides to explore a wealth of issues that are of interest to scholars from multiple fields. In the organizational domain, this research spans several levels of analysis, ranging from individual or personal to organizational.

Personal identity typically refers to unique individual attributes that are assumed not to be shared with other people and which do not indicate or derive from group membership (Alvesson et al. 2008). Different from personal identity, social identity refers to an individual’s perception of him or herself, resulting from his or her membership in a social group (Tajfel and Turner 1979). Moving up the
analytical scale, organizational identity is generally understood to be the members of an organization’s collective understanding of the features that are presumed to be central, distinctive, and relatively permanent about the organization (Albert and Whetten 1985; Dutton et al. 1994). Common to most theoretical and empirical accounts of organizational identity is the view that identity is rooted in a deep cultural level of the organization (Gioia et al. 2000), residing in interpretive schemes that organizational members collectively construct to provide meaning to their shared history, experiences, and activities (Gioia 1998; Ravasi and Schultz 2006).

Despite the apparent distinctions separating the different levels of analysis, several scholars have emphasized their similarities and called for a more holistic understanding of identities in organizational contexts. For example, Alvesson et al. (2008) maintained that “despite the appeal of persistent distinctions between personal and social identities… we also wish to resist the often arbitrary clarity of such divisions. Instead… we develop a sharper eye for the diverse and fine-tuned ways in which the personal-social relation might be configured in identity research” (p. 10). The authors observe the role that personal and social identities play in each other’s construction. On the one hand, “personal identities necessarily draw on available social discourses or narratives about who one can be and how one should act” (p.11). Furthermore, self conceptions emerge and develop in reference to a range of associations, roles, and behaviors that tie the individual to his or her social surroundings. On the other hand, social identities cannot be formed without individuals that engage in action and interaction that are informed by some notions of the self. Thus, the two forms of identity are intimately intertwined in a way that makes it hard to examine or understand one in complete separation from the other.

In accordance with this line of argumentation, several researchers have attempted to highlight the common features that personal, social and organizational identities share. Some have done this by stressing the relational aspects of identity. As pointed out by symbolic interactionists, personal identity is inherently relational (Sluss and Ashforth 2007); one’s self-conception as a powerful leader is difficult to achieve without the presence of followers. Social identity is similarly relational; it is through ongoing relationships, interactions, and comparisons with various out-groups that the in-group becomes a salient locus for individual identification and attachment. Organizational identities are also relational as they are constructed not only against a backdrop of members’ shared histories and experiences but also in the context of multiple interactions in which the organization is involved with a variety of outsiders such as costumers, competitors, and suppliers (Ashforth and Mael 1996; Gioia et al. 2000).

Another characteristic of identity in organizations is its fluidity. Although much of the literature has played up and focused on the seemingly stable and permanent features of identity, acknowledgement of its potentially changing character can be found in recent research on the topic (Gioia et al. 2000). For instance, personal identity is seen as a social construction deriving from changing interactions with others. As Weick puts it, “identities are constituted out of the process of interaction. To shift among interactions is to shift among definitions of the self” (Weick 1995, p.20). Social identity is also flexible; an individual’s representation of in-groups and out-groups is likely to change as features of the comparative and normative context undergo transformations (Cornelissen et al. 2007). Flexibility is also characteristic of organizational identity. Changes in the organization’s environment and relationships with other organizations are likely to require modifications to the way members interpret what is central and distinctive about their organization. That is, organizational changes will require members to actively reinterpret and develop new representations to symbolically characterize their organization (Fiol 1991).

Two additional qualities that characterize identities in organizational contexts are the role that they play in informing individual and collective action and their embeddedness in social discourse and communication. Firstly, individual actions are performed by actors with certain dispositions and preferences that derive from their self-conception. Likewise, social identity orientates the behaviors of individuals based on inter-group comparisons and relationships and the construal of social information. Organizational action is informed by organizational identity that provides a basis for
sense-making and renders a particular repertoire of behaviors appropriate; a “green” organization is likely to take certain actions to reduce operational costs and be associated with relevant industry and environmental groups to justify its green identity.

Secondly, most researchers agree that identity in organizational settings is produced and reproduced through ongoing communicative activities that take place within and across people and organizations. For example, social construction theorists maintain that personal identities are created, negotiated, and changed through ongoing interactions among multiple actors (Alvesson et al. 2008). Organizational identity is also a product of social communication; organizational members negotiate, through continuous interactions, a shared symbolic representation of their organization that gives a sense of meaning to the organization’s actions, objectives, and existence, and that distinguishes the organization from other social entities in its environment (Gioia 1998; Gioia et al. 2000).

To sum, identity issues are experienced by social actors embedded in organizational settings. The concept of identity provides a lens for studying how organizational members give meaning to their experiences as a basis for individual and collective action. Therefore, it offers an opportunity to explore the interrelationships between the symbolic and the concrete organizational domains, as well as to examine the reciprocity of micro activities and macro phenomena. The recognition that identity is a foundational notion that is essential to understand multiple organizational processes and experiences is evident in the wealth and diversity of research that has employed the concept. In what follows, we examine the utilization of the identity concept in IS research and characterize its application.

3 RESEARCH METHODOLOGY

3.1 Defining the Scope of the Review and Searching the Journals

To examine the use of the identity concept in IS research we conducted a review of some of the primary journals in the field: MIS Quarterly (MISQ), Information Systems Research (ISR), Journal of the Association of Information Systems (JAIS), Information Systems Journal (ISJ), European Journal of Information Systems (EJIS) and Journal of Management Studies (JMIS).

In selecting journals for review we wanted to include top ranked, North American journals (MISQ, ISR, and JAIS), top ranked European journals (EJIS and ISJ), and a journal which we thought is likely to publish identity research (JMIS). As we focus the on use of the identity concept in academic research, we did not include practitioner-oriented journals in our sample.


3.2 Selecting and Examining the Articles

Our search of the databases produced a sample of 395 articles, which matched our keyword. These articles constitute our Sample A. Next, both authors read the abstracts and keywords of the papers in Sample A to determine whether the concept of identity was employed theoretically and/or to interpret empirical data. After going through a portion of the articles in the sample we compared our analyses to align our sorting criteria and establish inter-rater reliability. The result of the sorting was a reduced
Sample B of 27 articles. We then read the articles carefully to assess each article’s use of the identity concept. In reviewing the articles, we examined two issues: first, whether the articles’ use of the identity concept was substantial and; second, whether they addressed the topic of identity in organizational settings, which is the focus of our study. In this process, we excluded articles that did not use the concept of identity in a theoretically-informed way. For example, articles that only mentioned identity in the introduction but not in other parts of the paper, or articles that used the concept as a peripheral idea. Furthermore, we excluded articles that used the word identity in a completely different connotation from our research interest of studying identity in organizations. This produced a (surprisingly) small sample of 11 articles, which we refer to as Sample C (a list of articles excluded is provided in Appendix A).

<table>
<thead>
<tr>
<th>Table 1: Sample sizes and distribution of articles on journals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Sample A</td>
</tr>
<tr>
<td>Sample B</td>
</tr>
<tr>
<td>Sample C</td>
</tr>
</tbody>
</table>

4 FINDINGS: IDENTITY IN ORGANIZATIONAL SETTINGS IN IS RESEARCH

The remaining 11 articles were read carefully with the aim of constructing categories based on their use of identity. Using techniques from grounded theory (Glaser and Strauss 1967; Strauss and Corbin 1998), we constructed categories of the use of the identity concept based on a process of naming and comparing. We constantly alternated between reading the articles and naming first order concepts that were then compared to the text and finally combined into second order categories which formed the categories of use.

Five of the articles address the interrelationship of technology and identity in the context of IS implementation. Van Akkeren and Rowlands (2007) examined the assimilation of a new ICT in a radiologist practice and drew on social actor theory to analyze the relationships among the radiology practitioners, the technology, and the context. Findings from the case study showed that user-identity can both inhibit and enable IS assimilation. Gal et al. (2008) studied the implementation of 3D technologies into the architecture industry. They proposed a model to outline the relationship among IS, information infrastructures, and organizational identities and suggested that IS help to form organizational identities and enable cross-organizational change. Similar to Gal et al, Alvarez (2008) also emphasized the co-construction of identity and IS and argued that technology, structure, and identity are mutually constitutive. Critically examining the implementation of an enterprise system, Alvarez discussed users’ power relations, experienced loss of autonomy, isolation, and fragmentation during the implementation process. Barrett and Walsham (1999) studied the implementation of an electronic trading system in the London Insurance Market and drew on work by Giddens to examine its impact on users’ self-identity. Finally, Barrett and Scott (2004) also utilized Giddens’ concept of

---

1 Most articles in Sample A used the word “identity” without actually engaging the concept as, for instance, in exploring identity-matching for crime-solving.
self-identity and examined how reflexive self-identity is impacted by increased globalization and ICT during the adoption of an e-trading system.

Articles in the second category examined the influence of the presentation and perception of identity on online behavior. Winter et al. (2003) conceptualized identity as the way an organization is perceived externally by its customers and discussed how managing a website as a symbolic representation of a company influences the customers’ perception of the company and thereby their purchasing behavior. Related to this is a paper by Forman et al. (2008) about online-shopping and consumers’ identity. Drawing on theories of information processing and partly based on social identity theory, the paper suggests that self-disclosure of consumers’ identity affects the behavior of other shoppers and is positively related to sales. Finally, Ma and Agarwal (2007) studied the impact of community infrastructure design and identity verification in computer-mediated communication. Their findings suggested that identity verification is positively impacted by IT artifacts, and leads to satisfaction and knowledge contribution in online communities.

The third category focuses on the role of identity in enabling collaboration and coordination across groups and organizations. Sarker and Sahay (2003) propose a theoretical model that relates the concepts of communication, virtual team development, and collaboration to understand how virtual teams develop over time. They suggest that the development of an ‘integrative identity’ across teams is an enabler of successful collaboration. The second paper in this category by Dickey et al. (2007) is a study of how customers and customer service representatives build a shared context in chat communication. Identity is viewed as the interpretations that customers have of the company representatives’ appearance in the chat session. The paper describes how improvements in peoples’ articulation of intention and creation of a shared frame of reference may be valuable in enabling coordination between customers and customer service representatives.

The fourth and final category is represented by a single paper by Merali (2002) that examined the role of organizational identity in enabling organizations to survive. Looking at inter-organizational relationships in managing organizational transformation, Merali views identity as the abstract characterization of the organization that persists over time. It is the embodiment of this identity through action that impacts the world and influences boundary management between organizations. The paper draws on Maturana and Varela’s concept of autopoietic unity as a metaphor and provides a framework for exploring issues of organizational identity and integrity, focusing on boundary phenomena to explain organizational knowledge processes.

5 INCORPORATING IDENTITY INTO IS RESEARCH: POTENTIAL DIRECTIONS AND OPPORTUNITIES

Our review of the leading IS journals revealed that the concept of identity work in organizations is seldom used by researchers in the field. When it is employed, it is mostly done without theoretically engaging the concept or leveraging the insights that could be yielded from its application. This finding stands in stark contrast to the overall trend in other management disciplines. As noted, the concept of identity has become increasingly popular in recent years and been used to investigate a wide array of organizational and managerial issues such as workers’ motivation, organizational strategic behavior, group functioning, power relations and politics, entrepreneurship, loyalty, and emotions at the workplace (Alvesson et al. 2008).

In light of the broad applicability of the concept, we believe that the lack of attention to the importance of identity in organizational settings constitutes a missed opportunity. Bringing identity issues into the fore can help develop novel theoretical accounts and produce rich empirical analyses that capture the complexity of organizational life and of the interactions between people and technology in organizational settings. In what follows we outline possible avenues for using the concept of identity in IS research.
Organizational research on identity has been previously classified according to three philosophical orientations (Alvesson et al. 2008): functional, interpretivist, and critical. We build on this classification to describe different theoretical areas in which the identity concept can be utilized to inform IS research and identify research opportunities that characterize each area.

**Functional**

Research in this area typically aims to find causal relationships between different organizational phenomena in order to enhance control over organizational processes and increase organizational functioning and efficiency. Identity is treated as a tangible construct that can be gauged, assessed, and intentionally manipulated through management interventions with real organizational consequences. For example, some research argues that members’ level of identification with their organization impacts decision-making processes, group cohesion (Ashforth and Mael 1989), and commitment (Sass and Canaray 1991). Other research maintains that the emergence of a collective identity influences the way members interpret and react to issues facing the organization (Dutton and Dukerich 1991; Gioia and Thomas 1996) by influencing the importance that members assign to them.

The theoretical framework most often associated with the functional approach is social identity theory (Tajfel and Turner 1979). The theory’s main contention is that people’s identity derives from the groups to which they believe they belong. Because people strive to maintain a positive self-identity, they will tend to favorably compare their in-group to external out-groups along valued dimensions. Applying these ideas to the organizational domain, it has been argued that organizational identity is merely a form of social identity, one that is associated with perceived membership to an organization (Haslam 2001). Therefore, one’s degree of identification with the organization provides the basis for a range of organizational behaviors such as leadership, group motivation, and the willingness to take on organizational roles and exercise collective power (Turner 2005).

Opportunities for IS researchers in this approach are varied. For example, one could look at the role of technology in facilitating the creation and maintenance of organizational identification among group members. As identification with the organization is assumed to have a decisive influence on a range of organizational actions, being able to control and manipulate the identification process becomes an important managerial issue. Therefore, research that looks at the way technology can assist in accomplishing this in different organizational situations, such as geographically distributed or virtual teams, can be particularly valuable (these ideas are reminiscent of those explored in papers in the third category we identified above). Another research possibility lies in examining the impact of members’ identification levels on their willingness to accept new technologies. Technology acceptance research is one of the most substantiated in the IS field. This area can be considerably enhanced by examining how the emergence and strength of social identities influence the propensity of users to adopt and use a new technology.

**Interpretivist**

Whereas functional research aims to target and utilize identity to produce effective organizational behavior, interpretivist researchers are primarily interested in understanding human experience in organizational settings. Rather than directly serving organizational interests, researchers in the interpretivist tradition look to gain in-depth insights into people’s subjective reflections on who they are and what they do (Alvesson et al. 2008). Interpretivist approaches focus on how people weave organizational narratives with personal experiences to construct identities that provide a sense of meaning and continuity over time and across geographical locations (Ravasi and Schultz 2006). This construction process is often referred to as ‘identity work’, a term that is meant to emphasize the continual and dynamic nature of identities in organizational settings and their capacity to change and adapt to accommodate transformations that take place within or outside the organization.

This idea is demonstrated in a number of studies such as Fiol’s examination of an organization named Tech-Co. During the 1970’s and 1980’s Tech-Co had a stable organizational identity as an engineering-driven data storage company. However, during the 1990’s, the computer storage industry
as a whole was undergoing significant changes from a primary hardware, engineering mindset to mindset of information management and storage solutions. Fiol followed the transformation in Tech-Co’s identity as the company attempted to adapt to the changes in its environment (Fiol 2002).

Some IS research that used the concept of identity has been interpretive in nature. For instance, Lamb and Davidson (2005) described the transformations in the professional identities of groups of scientists associated with the introduction of new IT. Similarly, Gal et al. (2008) studied the transformations in the identity of an organization as it adopted new IS. Walsham (1998) and Barrett and Walsham (1999) explored the links between the introduction of new IT and changes in the identities of groups of professionals in the London insurance market.

Thus, interpretivist IS research can enrich our understanding of how individuals, groups, and organizations incorporate technology-enabled changes in their environment into ongoing identity work; how new technologies get interpreted and fed into the way people perceive themselves and their organizations; how ongoing enactments of organizational interactions, practices, and identities are influenced by the introduction of new technologies; and what role existing identities play in sense-making processes of new technologies.

**Critical**

The critical approach focuses on power relations and repressive discourses that exist within and across organizations. These relations and discourses impose certain normative demands, behavioral scripts, and cognitive frames that shape individual, group, and organizational identities, both explicitly and implicitly. A critical perspective challenges some of the basic assumptions that characterize the functional and interpretivist approaches, most importantly, that individuals and groups freely construct their identities and (challenging the functional approach) that these identities will have beneficial outcomes both for the individuals involved and the organization (Alvesson et al. 2008).

A prominent theme in critical identity research is managerial interest in controlling employees through the regulation of their identities. Efforts to establish a rigid organizational environment that funnels identity construction in specific ways are given prime consideration. Attention shifts to the role that organizational elites play in generating discursive regimes and material arrangements that pose strict limitations on identity construction in ways that are deemed congruent with broad managerial objectives. For example, discourses of quality management, service management, and knowledge management provide a rich vocabulary and conceptualize the organization and its relationship with its members in ways that form and define certain identities, such as ‘the knowledge worker’. Identities can also be constituted by reference to their location within a broader organizational or inter-organizational scheme and in terms of their relationships to others. For instance, in a study of an advertising agency, Alvesson found that reference to other agencies as amateurish and insincere tended to be interpreted as communicating professionalism and honesty as desirable attributes to be possessed by members of the researched agency (Alvesson 1994).

IS identity research that adopts a critical stance can build on existing work in the field that has examined the repressive impact of IS on people’s privacy (Zuboff 1988), capacity to exercise their agency (Kallinikos 2004), and on the way organizational action and discourse are induced through the implementation of new technologies (Doolin 2002; Doolin 2003). Future research can examine how IS are used to impose certain discourses that facilitate the construction of particular identities; the role that IS play in the distribution of material and symbolic resources within and between organizations and the way these resources are used to construct different identities; and the mechanisms through which IS structure communicative activities within and between organizations and how these communicative activities (that may be power-laden, asymmetrical, or exploitive) are incorporated into identity construction processes.
6 CONCLUSION

Our review of the use of the identity concept in IS research has shown that term ‘identity’ has been cited substantively in the dominating IS journals to date (395 times, to be exact). At face value, this would seem to indicate a significant level of interest in the concept in the IS field and point to the existence of a high number of IS researchers seeking to engage seriously and constructively with the identity concept.

From an IS perspective, it would seem that the notion of identity work in organizations offers a number of distinctive concepts and perspectives that are appealing to researchers in the field. Among these are the possibility of using IS to develop members’ collective identities and enhance organizational performance; and developing an identity-based view of organizations which can open a conceptual platform to address IS implementation challenges, IS-induced organizational change processes, and the ongoing adaptation of IS to users’ practices.

While 395 articles in the sampled journals used the term ‘identity’, only 11 of them had examined the concept in organizational settings to develop or expand theoretical insights, interpret empirical data, or both. The vast majority of IS researchers have adopted the identity concept in a non-theoretical fashion. Nonetheless, the IS researchers that used the identity concept substantively, have done so in a wide variety of ways, demonstrating its broad applicability. A number of them focused on the relationship between technology and identity and examined how this relationship plays out during the implementation of IS. Another group of researchers have investigated the formulation, perception, and presentation of identity to enrich current understandings of online behavior. Others have looked at the role of identity in facilitating collaboration and coordination within or across organizational boundaries.

In light of the concept’s versatility, we would encourage future IS identity-researchers to follow suit and seek to explore the full scope of identity theory by borrowing from other disciplines rather than restricting themselves to building on the limited understanding of the concept that is apparently most directly applicable to their work. Based on Alvesson et al.’s classification, we outlined possible directions for such research, which fall into the functional, interpretivist, or the critical perspectives.

From a functional standpoint, identity is a tangible mechanism that can be utilized and leveraged to solve organizational problems, impact a range of organizational behaviors, and promote organizational effectiveness, e.g., through the timely implementation and adoption of technology. An interpretivist understanding of identity in IS research views IS users as social actors whose actions are underlined by their self and group conceptions. Therefore, in order to understand how and why people respond to and interact with technology in certain ways, we necessarily need to understand their identities. An in-depth examination of the communicative and symbolic activities that underscore identity construction can therefore provide a better understanding of a range of organizational events, e.g., why a particular technology is rejected by its intended users, or why users have difficulties integrating a technology into their everyday practice. From a critical point of view, the understanding of identity can produce important insights into the power structures and opposing interests often involved in the interaction between technology, social institutions, and actors in organizations.

Finally, the IS field has a lot of catching up to do with other management disciplines in terms of addressing identity work in organizations. As we noted in the beginning of the paper, in recent years identity research has flourished in most management disciplines to address numerous issues and enrich the discourse around a variety of managerial and organizational topics. IS researchers, for the most part, have yet to capitalize on this opportunity to further develop and diversify their conceptual playing field. This paper is meant to provide an initial exploration into some of the opportunities and possibilities for carrying this out.
REFERENCES


8 APPENDIX A

The following papers were excluded from sample B to form sample C: Erat et al. (2006); Hanish and Corbitt (2007); Huang et al. (2001); Hwang (2005); Jensen and Aanestad (2007); Koufaris (2002); McCoy et al. (2007); McGrath (2002); Adam et al. (2006); Gregg and Walczak (2008); Levina and Vaast (2005); Light (2007); Maes and Huizing (2005); Otjacques et al. (2007); Robey (2003); and Xu et al. (2007).
# A Situated Knowledge Work Context Perspective on Knowledge Management Software-Organisation Misalignments

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0322.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Knowledge Management Systems, Adoption, IT Business Alignment / Value, Practice</td>
</tr>
</tbody>
</table>
A SITUATED KNOWLEDGE WORK CONTEXT PERSPECTIVE
ON KNOWLEDGE MANAGEMENT
SOFTWARE-ORGANISATION MISALIGNMENTS

Khuong Le-Nguyen, Royal Holloway College, University of London, Egham, Surrey, TW20 0EX, UK, Email: k.le-nguyen@rhul.ac.uk
Romano Dyerson, Royal Holloway College, University of London, Egham, Surrey, TW20 0EX, UK, Email: r.dyerson@rhul.ac.uk
G. Harindranath, Royal Holloway College, University of London, Egham, Surrey, TW20 0EX, UK, Email: g.harindranath@rhul.ac.uk

Abstract

Underpinned by the work of Schultze and Boland (2000) on situated knowledge work context, and using Bourdieu’s (1990) theory of practice, this paper draws attention to potential sources of likely misalignment between Knowledge Management (KM) software and the implementing organisation. Using the case of a global IT-Management consultancy firm, the paper elucidates KM software-organisation misalignments as the consequence of differences between organisational and KM technology developer’s ‘situated’ knowledge work contexts with due consideration of the case organisation’s large and diverse user base. Such contextual differences reflect the differing ‘situated’ knowledge work structures that KM technology developers and adopting organisations deliberately acquire and evolve over time. Theoretically, by giving a more comprehensive account of how knowledge work structures were emergent and then embedded into the organisation that either develops or implements the technology, we wish to provide readers with insights into the sources of misalignment, thereby enriching theory of KM technology implementation. Practically, studies of this nature would help make organisations more aware of the factors that can influence KM software package implementation, particularly in the case of large firms that are characterised by high-value, text-based knowledge for decision-making, such as consultancies, R&D, healthcare or legal organisations.

Keywords: Knowledge Management, software package, theory of practice, knowledge work context, global IT-Management consultancy firm.


1 INTRODUCTION

An increasing number of knowledge management (KM) projects have not achieved the expected benefits from the technology and even worse, suffered critical failures (Malhotra, 2005). As Roell (2004: 1) argues: “tools that have been developed in KM focused on information management and do not support many of the key knowledge work processes”. Hence, if KM technology is not appropriately aligned with organisational and people ways of working, then even with state-of-the-art technology, organisations would not be able to provide the ‘right knowledge’ to the ‘right people’ at the ‘right time’. Our reading of the literature highlights that such low success rates were in part attributable to misfits arising from the different situational contexts of the adopting environment and the software developer. In other words, poor understanding of situated work practices often lies at the heart of failed KM implementation projects (e.g. Grudin, 1994; Suchman, 1995). More specifically, vendors and adopters alike should create solutions that are consistent with the situated knowledge work practices of the intended user communities (Schultze and Boland, 2000). This requires them to follow a practice-oriented approach, which, as Townley (1993: 235) argues, should focus not only on what people ‘actually’ do, but also on ‘what doing it does’. Suchman (1995) suggests that this approach is particularly suitable for KM because system developers do not have accepted models for the largely invisible and complex nature of knowledge work.

Although the concepts of situated knowledge work context and practice have been central in the writings of Orlikowski (1996), Majchrzak et al. (2000), Schultze and Boland (2000) and Orlikowski (2002), little is known of the sources of misalignments in terms of such situated context as well as how emergent knowledge work structures are surfaced on a large-scale. Notably, of the four studies, only Schultze and Boland opted for Bourdieu’s theory of practice (1990 and 1998) whilst the remaining employed Gidden’s (1984) structuration theory. We find that while both theories emphasise knowledge work practice as an on-going accomplishment, constituted and re-constituted in daily practice (Orlikowski, 2002), Bourdieu’s articulation of elements such as field (structures), habitus and practice in his theory helpfully reflects the nature of knowledge work practices as we observed it in our research setting to elaborate on the nature of misalignment. Further, Schultze and Boland’ study also exposes some limitations. Their study focused only on a small group of knowledge workers, and particularly a specific practice named ‘gate-keeping’, which we argue does not reflect a sufficiently diversified range of knowledge works. Their research setting – a large building material manufacturing firm – may also not truly represent a ‘knowledge-intensive firm’ consisting of communities of practice (Alvesson, 1993). Most importantly, the technology implementation project ‘did not move beyond its initial pilot stage’ (Schultze and Boland, 2000: 197), raising questions in our minds as to unexplored potential misalignments occurring in actual implementation.

Specifically, in this paper, our objective is to understand the sources of misalignment between KM technologies and organisations. To do so, we use a case study of an IT-management consulting company ranked among the world’s top 20 companies and whose offices are located throughout the world. Our discussion proceeds as follows. Section 2 uses Bourdieu’s theory to review the misalignments between the context of KM software and that of the adopting organisation. Section 3 introduces our research methodology and the framework. Section 4 introduces briefly the case company and then presents data analysis and findings. The last section discusses the key issues arising from the findings and presents some conclusions and suggestions for future research.

2 UNDERSTANDING THE MISALIGNMENT OF SITUATED KNOWLEDGE WORK CONTEXT

For this study, understanding the ‘situated’ nature of knowledge work practices is of paramount importance to help analyze sources of misalignment between KM technology and the adopting
organisation on the following grounds. First, users are usually unaware of developers’ context and the embedded assumptions and rules of the surrounding world (Orlikowski, 1992; Latour, 1992). Indeed, the context of knowledge work, along with their assumptions and rules as well as knowledge attributes such as its “sticky” and contextualised nature (Szulanski, 1996), is the factor that makes knowledge work practices different from other work practices (Schultze, 2000). In this context, a mere understanding of what ‘people do’ is often not adequate in helping organisations to anticipate the long-term impacts of technology, both intended and unintended (Schultze and Boland, 2000).

Second, knowledge, be it object, cognition or capability, must be enacted from people’s practices and reside in a particular work context (Orlikowski 2002). In this regard, to understand KM problems, analyzing what experts do in situ is crucial for studying how their practices are embedded in and shaped by work contexts (Hsiao et al, 2006). In other words, people and their practices can not be separated from the embedded work contexts (Lam, 1997; Tyre and von Hippel, 1997). Neither is it possible to analyze KM problems without considering the contexts where these practices were acquired (Lave and Wenger 1991; Orlikowski 2002). An ignorance of how practices are shaped within work contexts also induces operational problems of KM initiatives and results in difficulties with regard to technology adoption.

Third, different expert groups may employ different types of knowledge under different work contexts (Bogenrieder and Nooteboom, 2004). Alternatively, to acquire a ‘situated’ knowledge for a particular project, experts must participate in different working situations (Lave and Wenger, 1991; Tyre and von Hippel, 1997). Under different work contexts, knowledge can reside in physical processes, social communities and industrial contexts (Lam, 1997; Tyre and von Hippel, 1997). Therefore, detached from work contexts, the analysis of misalignment will only show ‘how information is managed’ without elaborating on why one group can employ technology to collaborate, share, and reuse knowledge effectively whilst another group struggles with the same technology (e.g. Huber, 2001). Hence, unattached to practices, transferring knowledge as object from one place to another or sharing it as individual cognition is out of the question (Hsiao et al, 2006). Transferring knowledge as a capability initiates a learning and evolving mechanism within individuals to enact ‘doable practices’ in any particular context (Orlikowski, 2002).

Our growing awareness of situated knowledge work context implies that anticipating the impact of KM technology on both organisation and individuals alike would require us to understand the various processes and relationships through which work practices enact and re-enact the ‘objectified social structures and the conditions’ where they occur (Schultze and Boland, 2000: 195; See also Orlikowski, 2002). We employ Bourdieu’s theory of practice as we believe it fits the analysis of ‘situated’ knowledge work practices in our research setting.

For instance, elements such as field (structures), habitus and practice described in Bourdieu’s theory of practice (1990 and 1998; Bourdieu and Wacquant, 1992) are very much relevant to knowledge work practices in any organisation (See figure 1 in section 3). To illustrate, within a particular field, knowledge workers struggle for power to obtain the intellectual capital. These multiple fields then define the (social) knowledge structures characterising a ‘context’. Such (social) knowledge structures are then internalised into knowledge workers’ minds and bodies as habitus acting as a symbolic template for their conduct, thoughts and judgments. These dimensions of habitus, underpinned by their own cognitive and motivating principles and shaped by their own collective experiences (Bourdieu, 1990), then decide their repeated and patterned knowledge behaviours and practices. Notably, understanding two important qualities of habitus is significant to know why there may be existing differences between the knowledge work practices between two individuals or two organisations. One is the habitus’ durability that explains why people adhere to certain knowledge work practices regardless of the working conditions. The other is the transposable quality that facilitates the knowledge worker’s ability to cope with ever-changing situations and to innovate in the face of unforeseen circumstances. Such an improvisation in practice is critical for organisational learning, innovation and change (Barrett, 1998). Deducing from Bourdieu’s theory, if a KM technology is developed in a context where people, for some reason (e.g. history, culture, experience, etc.) become
familiar with a particular ‘style’ of knowledge work practice that favours resistance to change practices, then this particular KM technology may not be appropriate for a wider market where requirements for technology customisation is mandatory. Alternatively, in a context such as that of a global management consultancy firm, knowledge workers, in some cases, may need to adapt to certain types of knowledge work practice in order to be able to work with international clients. Such ‘situated’ knowledge work practices may inevitably be different from what was assumed and inscribed by the KM technology developer. In sum, since habitus is a concept resembling culture (Harker, 1990) and ‘socialised subjectivity’ (Bourdieu and Wacquant, 1992: 126), we believe that country-level as well as industry-level differences are among pivotal factors in designing, developing, adopting and implementing KM technology.

As a result of a dialectical relationship between a specific circumstance in a field and habitus, knowledge work practices are effectively improvisational actions enacted by dynamically combining past experience, the present situation, and the implicit anticipation of the future consequences of these very practices (Bourdieu, 1973). The combination of all three elements varies from one organisation to another. Although two organisations may have the same assumptions and visions for their current contexts of knowledge work together with the implicit prediction of the outcomes of such practices, differences in their past experiences (i.e. behaviours, feeling and judgement, success or failure) may eventually result in somewhat differing situated practices performed by each organisation. For instance, Lam’s (1997) study exploring knowledge transfer between British and Japanese engineers found that British engineers worked quite independently in their social network and relied on explicitly codified practices (i.e. electronic repository) for disseminating knowledge. In contrast, as the Japanese engineers collaborated in a close-knit social network, knowledge transfer was mostly dependent on their intensive socialisation, which cannot be easily made explicit unlike formally codified rules. In this example, we contend that the difference in habitus, where British engineers work individually whilst Japanese engineers work collectively, eventually leads to the difference in knowledge work practice where Japanese engineers rely more heavily on their team-mates for solutions. We would therefore argue that a specific KM software package, if produced by developers in the UK and sold to the Japanese market, should additionally reflect such unique knowledge work practices to boost collaboration and capture necessary knowledge for re-use. The differing ‘situated’ knowledge work contexts can be also reflected via industry-level differences. Consider the pace of innovation activities for example. In the context of an innovative but stable industry, the development of a new product often follows a continuous technological roadmap (e.g. microprocessors). However, in the context of fast-moving industries such as biotechnology, aerospace, nano or computing, the ever-increasing competition in the marketplace may activate a rapid, discontinuous innovation in product and process (Hsiao et al, 2006). Eventually, lessons learned must be updated frequently, rendering it increasingly difficult for experts to spend sufficient time submitting their reflections and experiences (Hsiao et al., 2006). The differing nature of the two industry contexts reflected by the propensities (habitus) towards producing products, suggest that the resultant practices would require organisations to acquire differing knowledge work structures. Hence, employing the same technology to support knowledge works in two completely differing ‘situated’ contexts could be problematic. We would expect for example a KM software package employed in the context of fast-moving industries to incorporate special structures to enable experts to capture, classify and extract contents for re-use more efficiently and in a more timely fashion than in slower moving industries. Pentland (1995) for instance reported in his research that the decision-support system designed to share energy-audit engineers’ practices was significantly limited because the software algorithms became quickly outmoded due to the increasingly diversified energy-audit regulations in the US and the consequent changes to the audit processes.

3 THE RESEARCH FRAMEWORK AND METHODOLOGY

Our literature review has suggested that the misalignments of knowledge work context between KM software and organisations can be conceptualised as in figure 1. Additionally we use Markus and
Tanis’s (2000) four-phase view of the enterprise system experience cycle to explore misalignments. Markus and Tanis’s phase view helps us to clarify the business or operational context where misalignments arose. Such a context is also essential to understanding how KM practices occur throughout the technology lifecycle. Based on our case firm’s actual software implementation, we focus on the first three phases. Phase I, the chattering phase, involves gathering and defining the business case and solution constraints. Phase II, the project phase, involves getting system and end-users ‘up and running’. Phase III, the shakedown phase, involves getting the system and users to be in normal operation after the system’s roll-out. Each of these phases occurs sequentially with no time lags in between the phases.

We adopted a case study strategy (Yin, 1994) to investigate the likely misalignments between KM software package and the implementing organisation. This is a well-accepted approach to study the complex phenomena of IT implementation in an organisational setting (Orlikowski and Baroudi, 1991; Yin, 1994). We analysed the Change Request Forms (CRFs) submitted by the case study organisation from phase I to III to identify instances of misalignments during the course of a KM portal implementation project. A misalignment was defined as any instance where the KM group identified an organisational requirement that they felt was not being addressed by the KM software package. To exclude minor or trivial misalignments only those reviewed and accepted by the KM team and recorded in the standardised CRFs were analysed. The KM group and IS professional discussed with each other to decide whether to accept, put on hold or reject the request, and then with vendor’s consultants to decide whether to adapt to or customise the portal. Following this analysis, semistructured interviews were used to confirm and further clarify the contexts and rationale of identified and potential misalignments. The field work (lasting about 9 months) comprised of familiarising ourselves of the company’s business processes, reviewing the KM system and documentation (e.g. the technology, company’s documentation, contracts, project files, meeting minutes, requirements analysis, issues logs and video recordings of the negotiation between the KM group and vendor’s consultants), and interviewing 19 key internal stakeholders to gain insights from many different perspectives including all members of the KM group, i.e. KM programme manager, KM roll-out project managers, content managers, IS professionals, and especially two consultants representing the vendor. Additionally, informal meetings formed further sources of data gathering. Care was taken to consider all pertinent case evidence to reduce the risk of research bias. Contextual data that further explain the observations made were taken into account (Klein and Myers, 1999). Each interview lasted about an hour and was then transcribed and checked with interviewees for accuracy. The interviews were analysed to identify the sources of misalignments. The conceptual framework is used to examine the data, and to consider some ‘what if’ questions arising from problems reported by the interviewees or anticipated during the case. Findings were also discussed with two industry experts with rich experience of KM implementation. In all, 83 distinct instances of misalignments were
identified. However, since the focus of this paper is on misalignments of situated knowledge work context, we therefore analysed only 43 instances that are related to this issue. Our analysis was conducted as an iterative loop to uncover the factors underlying each instance. We found that of 43 instances, 32 surfaced during phase III. This means that the majority of misalignments only surfaced during actual interaction with the business or operational context. The remaining instances submitted during phase II were requested because of the consultants’ experiences.

4 CASE STUDY ANALYSIS

4.1 Case company’s profile

*KLN (pseudonym)*, headquartered in Europe and currently ranked within the world’s top 20 companies in terms of revenues, is a global IT-Management consultancy firm and employs tens of thousands of people in over 35 countries. KLN provides business consulting, systems integration and IT and business process outsourcing services across diverse markets including telecoms and media (T&M), financial services (FS), energy and utilities (E&U), industry, distribution and transport (IDT), space and defence (S&D) and the public sector. The company formally launched its KM initiative in 2005 although it had used portal technology from mid 2003. From late 2006, the company decided to migrate onto the latest version of the portal technologies with many enhanced features and a new design. The company’s portal is divided into three environments, namely *My Information*¹, *Workspaces*² and *Our Knowledge*³. This organisation was chosen based on a combination of accessibility and representativeness: KLN is a global consulting firm which is commonly discussed as the archetype of knowledge-intensive firms consisting of multiple communities of practice (Alvesson, 1993); KLN has implemented its KM initiative, including its portal, in a substantial way; KLN has a large-scale technology infrastructure in terms of its user base; KLN’s global operational context could present some interesting challenges affecting its technology implementation.

4.2 Data analysis

As far as the ‘situated’ knowledge work context misalignments are concerned, we have identified and classified them into six main themes of misalignment in accordance with our conceptual framework. The themes are related to access control for external workspaces, assistance for working with smart documents, enhancing the report function in *Our Knowledge* environment, searching and grouping results according to content types, capturing organisational metadata, and re-using workspace’s default components.

The first misalignment theme was about strengthening the access control mechanism. One of the key objectives of this case company’s KM initiative is to enable its clients to access workspaces outside the company conveniently. Since clients can only access ‘external workspaces’, which are similar to the ‘actual’ workspaces inside the company, it is important to effectively control clients’ access to these external workspaces. Upon reviewing the current protection and authentication method, some consultants found that the controlling mechanism must be strengthened to ensure the integrity of the organisation’s intellectual assets. For example, the current authentication method requires only a username and password check to fully access the system. However, the requesters were attempting to implement a more secure access control process that required the verification of a client’s identity and

---

¹ *My Information* provides private and shared information relating to each staff. It features a personal storage space for shared and private documents and allows for the management of an individual’s company profile.

² *Workspaces* are web-based environments that enable teams to share and collaborate on information.

³ *Our Knowledge* is a central store of publications, shared with all company staff. The information is stored in knowledge areas by subjects, such as sales and marketing and market intelligence.
purpose of access, in order to facilitate a more differentiated approach to access levels for various
clients. To progress between levels, clients must interact with a series of prompting questions on their
understanding of the projects, i.e. their roles and status. All these practices are monitored by the local
workspace’s manager in real-time. Viewed from Bourdieu’s theory of practice, this ‘situated’
knowledge work practice led to a misalignment that can be explained as follows. First, for the position
in field (structures), the primary concern of this global consultancy firm, in general and the KM group
in particular, is dealing with risks of breaching information security, ensuring knowledge integrity and
protecting intellectual assets. Specifically, the risk of breaching information security is higher than
ever because of KLN’s global operational context. The problem is getting more acute since KLN has
pursued an acquisition strategy in recent years, and thus experienced an expanding pool of
international projects. In this context, the fact that ever more clients have access to external
workspaces further threatens the integrity of KLN’s intellectual asset. These facts resulted in the
habitus that led the KM group to apply a more restrictive access control mechanism. Thus, by
suggesting the above-mentioned practices to verify and authorise the clients’ access level, consultants
believed organisational knowledge would be better controlled. Misalignment arose as the portal was
not intended to support this sophisticated mechanism.

The second misalignment theme was related to assisting content managers and experts in classifying
documents more efficiently. Specifically, what the requestors suggested doing was to equip the portal
with structures to easily capture contextual organisational metadata from produced or received
contents. For example, regarding the E&U industry, many projects and requirements from
international clients have produced a very high volume of documents, prompting the need to classify
these contents more efficiently. According to the KM Co-ordinator in Holland: “Tim [consultant
based in Holland] told me a couple of times that the [current] metadata could not help classify the
documents effectively. [Here is] an example. Tim is the author of two documents on Intelligent
Transport Systems (ITS) in Holland and Singapore. There are certainly many common attributes
within [these documents]. However, the context [between them] is slightly different. I mean, users
need to know how, what and why behind the work [of ITS] is being undertaken in those countries or
[they] need to know for each work’s context, what decisions were made in the first place that our
company can begin and continue, to learn”.

One way to deal with the above problem is to capture more ‘information in context’ for the meta-data.
This kind of ‘information in context’, which seems to be unique or situational, is normally defined and
reviewed in a meeting held every month for leading experts in each industry sector to find out what
have been the prominent or newly-emerging issues to be included. Then, such ‘information in context’
would be cross-checked by other industry sectors to remove duplication or confusion and to increase
transparency. Eventually the portal should be able to detect certain keywords which appear in
frequencies defined by this ‘contextual information’. Viewed from Bourdieu’s theory of practice, this
’situated’ knowledge work practice led to a misalignment that can be explained as follows. First, for
the position in the field (structures), being empowered by portal technology, content managers must be
able to classify an increasing volume of contents accurately. However, what is worth mentioning is the
fast expanding pool of international clients and projects that causes difficulties and complexities for
classifying the contents effectively and quickly. Given this context, content managers are prone to
(habitus) capturing more valuable data describing the projects’ circumstances, thereby continually
updating the metadata and taxonomy. Adopting the above-mentioned practice can strengthen content
managers’ positions in the field. Misalignment arose since the portal was not assumed to have the
capability of capturing such ‘situational’ data.

The third misalignment theme touched upon enhancing the reporting function in the organisational
repository. Because lack of reporting (i.e. on pushing and pulling knowledge items to and from Our
Knowledge environment) led to resistance in terms of KM implementation in some parts of the
business, content managers in Bangalore (India) and UK suggested upgrading the reporting function.
On the one hand, this upgrading is crucial for content life-cycle management to ensure the
effectiveness, quality and reliability of the content for users. On the other hand, upgrading the
reporting function could provide content managers with competent tools to encourage users’ participation in contributing knowledge to a particular area such as sales and marketing. Content managers could also measure the level of contribution from different users or establish the most visited or frequently downloaded documents in a particular country, business unit or service sector. This information could also be useful for the HR department as part of the KM performance indicator (KPI). This context was clarified by a KM programme manager: “One of the key things is to provide them [with] the feedback in terms of what is actually happening. If we ask the Marketing department in a country to do something, then we need to show them what they have actually achieved so far and give them some targets, for example and allow them to compare how they are doing with another company in other parts of business...So, we will know how much content is being provided by each part of the business and what the usage level is like...In other words, there are two sides. One is [that] we want to have reporting on what is in there [Our Knowledge]. The other is [that] we want to have reporting on what has been taken out. That means we have the publication side and the usage side”.

This misalignment reveals a difference in ‘situated’ knowledge work context. In accordance with Bourdieu’s theory, for position (structure) in the field, to support the overall KM initiative, content managers need to assess how users across the company pull and push knowledge from and to Our Knowledge environment. Their roles to support the knowledge contribution and use then lead to their tendency (habitus) to measure the number of documents to be downloaded and uploaded in each business unit or sector. The portal was not designed to assist content managers in capturing and reporting data on knowledge use in Our Knowledge environment. Therefore, by adopting the above suggested practices, content managers believe their positions in the field could be strengthened.

The fourth misalignment theme regarded to enabling the project teams to work more productively. From a knowledge-based perspective, assisting knowledge workers in creating contents is a good way to increase responses to fast changing business environment (Nonaka et al., 2000). There could be certain ways to automatically provide staff with data, information and reports for analyses. One way was to request to adopt a process whereby consultants are able to work with smart documents in workspaces. Specifically, by trying to embed a special tool into the portal, consultants could start a new project based on time-lines. Each of the project’s time-lines links with appropriate templates for documents and analyses that need to be created at each point of the project life-cycle. To do the analysis, topic or context-related reports for each of the project’s time-line are supplied by an automated procedure that collects data from Business Intelligence (BI) software. This kind of procedure is based on profiles defined by consultants before and during their project. Such knowledge works are particularly useful in the FS, T&M and E&U sectors. According to a senior consultant in IDT and FS in Germany: “Because of our recent acquisition, within six months, the number of projects and clients has increased remarkably. The deals are getting more complex and uncertain. To ensure our powerful delivery capability, we must be able to cope with abrupt changes and complexities...Internally, part of our competitive advantages is a strong BI [Business Intelligence] system continually providing us with quality and reliable data about our competitors and the [global] market. Our approach is to connect the output of BI with each of the project’s time-lines. That means [that] we are trying to make things [reports and templates] ready for them [consultants] to analyse throughout their projects. They would no longer spend too much time reading and manually importing things from BI [system]”.

Viewed from Bourdieu’s theory of practice, this misalignment can be explained as follows. First, for the position in the field (structures), there are several unforeseen or fast-changing circumstances occurring in sectors such as FS, T&M or E&U. To survive in the field, team members need to equip themselves with tools and data to cope with such situations. To tackle these situations, team members tend to look for (habitus) appropriate ready-for-use templates and automatically imported reports to support analysis in a timely fashion. By adopting the above mentioned practices, consultants could cope better with the fast-moving business environment. Misalignment arose because the portal was not assumed to support such an organisation-specific approach.
The fifth misalignment theme was related to organisation’s processes of extracting and viewing knowledge items. Given the nature of the work, many users favour searching and grouping results by content type (i.e. brochure, project summaries and references) and languages in Our Knowledge environment as they believe content type is important for refining what they are looking for. For example, the KM roll-out project manager argued that the retrieval process would be more productive and accurate when items are grouped and displayed according to content type. This misalignment could be understood by using Bourdieu’s theory. Indeed, at each point of time during the project, consultants would have a special need for knowledge items because of their positions in the field. If they are related to customer service, what they need to find is customer reference. From a Knowledge-Based View, this practice is related to the speed of pulling the knowledge items for knowledge workers. For instance, a staff working on a Customer Relationship Management project for clients of public transportation in France would need to know about sales of tickets, brochure or case studies on their roles thus lead to their disposition (habitus) to find only things having the appropriate properties and metadata. Hence, the suggested practices may assist consultants in searching for the right items in less time. Misalignment arose because the portal was not assumed to address this organisational-specific need of extracting and view contents.

The sixth misalignment theme was about automatising more the mechanism of creating bidding workspaces. To win bids over competitors, one of the critical tasks is creating bid projects, and thus developing the necessary bidding workspaces in less time. To do so, it is important to be able to ‘reuse’ templates as well as default components of existing workspaces. The context was clarified by the KM programme manager as follows: “Here is an example of UK financial services. OK, I want to create bid workspaces. It may be that I’ve got some local business processes which applied for that particular business unit. So this means [that] all bids must have certain checklists [of what-to-do actions as described by the local business processes] to be followed. In that situation, it’s more convenient for me to set up workspaces with default lists or web-parts or whatever for all bid workspaces which are owned by UK financial services. That is to say, I want to have the ability to save the template for [new bid] workspaces, and to feed that [the templates] into [the local business] processes so one can use that template by default”.

Simply put, such a requested approach could be seen as a way of standardising bid workspaces, which was supposed to be changed in the old version of the portal. Now with the portal’s new version, it is possible for the KM group to do things like default web-parts and default checklists so that the workspace’s configuration can be captured and stored. By employing Bourdieu’s theory, this misalignment can be explained as follows. First, for the position in the field (structures), in some core service sectors of KLN such as FS, IDT and E&U, the pressure for winning bid projects is gradually rising. A fast expanding pool of projects and an ever-increasing demand from clients create pressure on consultants to respond to bids rapidly. Given this context, consultants could quickly create (habitus) bid workspaces using templates without having to manually re-create every aspect of the workspace or default checklists. Such evolving structures would help teams save cost and time, thereby reducing the cost of the bid. Misalignment arose because the portal was not assumed to support consultants to create workspaces in this way.

5 DISCUSSION & CONCLUSION

Misalignment instances found in this case study have shown differences in situated knowledge work context between technology and the organisation. Such misalignments have also illustrated the complexities of the on-going interaction between consultants using the portal and both the immediate
and broader contexts in which these technologies are situated. Specifically, underpinned by Suchman’s (1987) account, we contend that such complexities reflected by the requirements and justifications for changing existing knowledge work structures or embedding emergent ones are contingent upon the immediate and broader contexts rather than the mere intention of the consultants for better knowledge work performance per se. We then argue that it is the intensity of such complexities that may be attributable to the extent, whether it should be superficial or deep, to which consultants wish to change existing knowledge structures or embed emergent ones. Additionally, we believe that the nuances of such complexities may also have certain impact on the consultants’ attitude towards changing or embedding the structures. We have shown how this occurred in the case using Bourdieu’s theory to analyse the surfaced misalignments. For instance, in relation to the misalignment in terms of external workspace access control, the ongoing interaction between consultants and both the immediate context where the clients were operating, and the broader context represented by the company’s global operational environment, created new levels of complexity that in turn opened up the potential for information security breaches. This complexity became more severe in terms of intensity, and more acute in terms of nuance because of the less secure global context and an expanding user (including international clients) base. Under these circumstances, the fact that more and more clients access the external workspaces further threatened the integrity of the company’s knowledge asset. This explains the reasons why some consultants suggested adopting a more subtle controlling mechanism, and why some consultants’ attitudes were rated ‘emergency’ in order to ‘deeply’ embed certain knowledge work structures into the portal with maximum effort and resources. Another illustration is the misalignment related to assisting experts to easily capture contextual organisational metadata. The complexity here had to do with increasing time pressures on consultants causing difficulties for classifying the contents effectively and quickly. This ever-growing pressure was due to the fast expanding pool of international clients and projects which resulted in an increasing volume of contents. Consequently, since this complexity is severe in terms of intensity and nuance, some consultants requested a ‘deep’ change by embedding emergent structures into the portal. Such ongoing interactions between consultants and both the immediate and broader contexts will continually create new complexities because the KM technology in this case study has opened up a new network of relationships across boundaries which eventually make uncertainties and instability become more prevalent (King and Star, 1990). Alternatively, in accordance with Orlikowski’s (2002) account that ‘knowing-in-practice’ is necessarily provisional, not something as ‘given and static’ as people draw upon their physical presence in a social setting, on their cultural background and experience, and on sentient and sensory information (Blackler, 1995; Tyre and von Hippel, 1997), we believe that misalignments of this type will keep arising. More particularly, knowledge work practice does not exist independent of social interaction, and its content does not necessarily mean the same value to all the people getting involved in the designing, developing and adoption processes (Dougherty, 2004). This differentiates KM software packages from the other enterprise systems such as ERP/CRM in that the misalignments relating to KM are perpetual throughout the system life-cycle.

The contributions of this study are threefold. In discussing a number of misalignments, our study strengthens Schultze and Borland’s (2000) account that exploring as well as understanding the unique knowledge work context is critical for the successful adoption and implementation of KM software packages. Besides, in addressing the weaknesses of their work, our results could be useful in terms of understanding the sources of misalignment and further presenting a more diversified, nuanced picture of misalignment in a global operational context comprising a large user base, thereby enriching theory of KM technology adoption and implementation. From an organisational perspective, such an understanding is highly salient as the issues of technology adoption “profoundly affect the manner, quality, and outcomes of organisational realities” (Orlikowski and Scott, 2008: 5). Practically, studies of this nature should assist organisations toward understanding the factors inherent within a successful KM technology implementation project in large or global firms, particularly those rich in high-value

---

4 KLN’s KM group defines four levels of priority to tackle misalignments: emergency, high, medium and low.
text-based knowledge for making decisions like consultancies, healthcare, R&D, defence, financial services and legal organisations.

Results of this study must be interpreted in the context of its limitations. First, given Markus and Tannis’s (2000) four-phase enterprise system lifecycle and the actual implementation, this single case study has only explored the first three phases. Misalignments arising from the onward and upward phase, which continues from normal operation until the system is replaced by an upgrade or a different system, remain undetected. This latter phase is essential for a complete assessment of the misalignments between the technology and this global firm. Second, due to limited organisational resources, we could not interview more users in different countries to gain a better view of situated knowledge work structures, both current and emergent. Future research should look into misalignments identified from a more highly regulated domain such as legal or health-care services, or how such misalignments have influenced organisational responses via organisational adaptation or technology customisation, or the process through which organisational staff arrive at a resolution for the misalignment (e.g. how users identify emergent situated knowledge work structures and then persuade the KM group to accept their requests). Such findings could have some valuable implications for knowledge based structures for innovation (See e.g. Anand et al, 2007).

References


Proceedings ECIS 2009


# SUCCESS FACTOR VALIDATION FOR GLOBAL ERP PROGRAMMES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0098.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Critical Success Factors (CSFs), Partial Least Squares, Structural Equation Modeling, Enterprise resource planning (ERP) (packaged systems)</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
SUCCESS FACTOR VALIDATION FOR GLOBAL ERP PROGRAMMES

Seidel, Gunter, Institute of Information Management, University of St. Gallen, Müller-Friedberg-Strasse 8, CH-9000 St. Gallen, Switzerland, gunter@gunter-seidel.de
Back, Andrea, Institute of Information Management, University of St. Gallen, Müller-Friedberg-Strasse 8, CH-9000 St. Gallen, Switzerland, Andrea.Back@unisg.ch

The success of current global enterprise resource planning (ERP) programmes depends on a wider range of factors than encountered in previous ERP efforts. Therefore, a better understanding of how these factors contribute to the effectiveness of ERP will aid programme development. To this end, the validity of twelve success factors with two to seven associated management actions derived by prior research was verified by a survey evaluated with structural equation modelling (SEM) based on the partial least squares (PLS) algorithm. This study was able to verify all twelve success factors and more than 80% of the associated management actions. The results form a predictive model of success probability of global ERP programmes.

Keywords: CSF, ERP, global, SEM, PLS, predictive model, success factors.
1 INTRODUCTION

The number of companies operating globally is constantly increasing. This expansion applies to large corporations as well as small and medium-size enterprises. In order for these companies to operate globally, they require a global view of processes and their implementation in global enterprise resource planning (ERP) systems. These ERP systems allow to integrate an organization’s information sources and to harmonize its processes across multiple sites and countries. Most ERP systems are based on software packages from companies like SAP or Oracle. ERP implementations are generally cost intensive and have a duration of many months or years.

But, not all global ERP programmes are successfully implemented. Some implementation projects fail in terms of classical project tracking, e.g. slippage of roll-out dates, budget overruns or missed scope objectives, while others fail more severely in their business impact, e.g. intended business benefits are not achieved including process standardization, process automation and asset carrying cost reduction.

A report on ERP implementations in companies with more than $500 million in revenues indicated an average schedule overrun of 230%, an average budget overrun of 178% and an average slide of functional improvements of 59% (Buckhout et al., 1999, p. 61).

To reduce the number of failed ERP programs, their general success factors have been investigated but reports of these are not sufficiently comprehensive for global programmes as multi-national operations generate aspects that need to be addressed for success. To generate a better understanding of how these success factors contribute to the efficiency and effectiveness of ERP and to aid programme development the validity of twelve success factors with two to seven associated management actions derived by prior research was verified by a survey evaluated with structural equation modelling (SEM) based on the partial least squares (PLS) algorithm.

This study was able to verify all twelve success factors and more than 80% of the associated management actions. The results form a predictive model of success probability of global ERP programmes. Use of this model may aid companies in developing effective global ERP programmes.

2 BACKGROUND

Many companies of various sizes must decide the extent and means by which to support worldwide activities through globally-harmonized processes and systems. An example of a typical global ERP implementation based on a case study is extensively described in the literature by Sankar & Rau (2006). A detailed review of whether to implement an ERP system, categorization of the types of benefits to be expected and the guidelines in selection of ERP consultants, software and hardware vendors can be found in Davenport (2000).

One way to face this decision is to use a technical approach based on templates; this approach was defined about a decade ago. Huber et al. (2000) defined a template as "concepts or models for the standardization of processes, functions, and data that could be implemented in a physical (ERP) system (Huber et al., 2000, p. 4)" and further defines a concept for standardization of company-wide ERP systems.

Necessary requirements to implement global ERP systems are investigated in Davidenkoff & Werner (2008). In particular the legal requirements, customs and user requests, but as well challenges in languages, address versions and time zones are reviewed (Davidenkoff & Werner, 2008, p. 37ff).

In general, Critical Success Factors (CSF) for implementing ERP programmes have been broadly analysed in the literature. Early investigation of CSF as Holland & Light (2003) were based on general case studies. Later Nah et al. (2001), Somers & Nelson (2001) and Al-Mashari et al. (2003)
investigated CSF for the different phases of ERP programmes. While recent studies investigated CSF for specific settings as e.g. industry and size (Soja & Put, 2007), no comprehensive investigation of CSF for global ERP programmes was identified from literature.

Aspects of change management are always part of ERP implementations and have been investigated repeatedly. For example, Hossain et al. (2002) propagates the use of the "myth of integration", i.e. the vision of an integrated enterprise as the driver for change (Hossain et al., 2002, p. 17ff). While the CSF list in the literature is extensive, it generally does not consider multi-national aspects of ERP implementation. However, a few papers have been dedicated to this subject. Huang & Palvia (2001) introduce a research framework to compare ERP deployment in developed and developing countries (Huang & Palvia, 2001). Multi-national ERP implementation practices were shown to be affected by national differences, identified as culture and language, government/corporate politics, management style, government regulations, time zone and labour skills (Sheu et al., 2003). The relationship between ERP implementation and a firm’s competitive strategy has also been investigated and national culture and government/corporate policies, in particular, were found to have a significant impact on ERP deployment (Yen & Sheu, 2004).

3 HYPOTHESIS

A series of expert interviews conducted in prior research and an extensive review of CSF literature has led to a definition of success of global ERP programmes and a number of common and site specific success factors (Seidel, 2009). In total 13 interviews were conducted with CIOs and project managers of multi-national ERP programmes defined as covering more than one language area. Content validity was ensured by using a theory based interview guide and inclusion of CSF only if cited without contradiction by multiple interviewees. These success factors and associated management actions and their impact on success as defined below are the hypothesis to be verified in this research.

3.1 Definitions of Success

Success of a global ERP programme is two-fold. On one hand the impact of the resulting Information System on the organisation is of interest. On the other hand the execution of the global ERP programme is of interest as a measure of efficiency. The former is named 'Programme Objectives', the latter 'Project Objectives' in this research.

The programme objectives are either clearly quantifiable 'Savings' in terms of IT costs, process costs or directly associated reductions to specific business expenditures such as stock-carrying costs. On the other hand, a number of indirectly quantifiable long-term business benefits are possible benefits such as risk reduction, auditability, adherence to standards and improved competitiveness. These are company-specific and are grouped under the heading 'Business Improvements'. Therefore the variable 'Programme Objectives' is hypothesized to be defined by 'Business Improvements' and 'Savings'.

The importance of the project objectives varies from programme to programme. Generally an on-time introduction is considered important, as it is highly visible in the organization and directly drives costs. On the other hand, the traditional trade-off between scope and time is always imminent. While in-budget programme completion is desirable, many programmes have multiple budget revisions due to the long timeframe of implementation and the constantly changing conditions during the roll-outs. While in-scope has many company-specific definitions, one common theme was the avoidance of business interruptions during go-live, and functioning maintenance and support. Therefore the variable 'Project Objectives' is defined by 'On-Time', 'In-Budget' and 'In-Scope'.
3.2 Common Success and Endangerment Factors

A number of success and endangerment factors are common to all ERP programmes and influence the programme outcomes of each site. Each factor has a number of management actions associated which drive the respective success factor. All of these success factors and management actions were verified by expert interviews in prior research. The list of these indicators to be verified in this research can be found in Table 1.

<table>
<thead>
<tr>
<th>CSF</th>
<th>Description</th>
<th>Management Actions</th>
</tr>
</thead>
</table>
| Change Management Approach | Company has an effective approach to handle the organizational changes induced by the ERP roll-out. | • Change management is local  
• Change management is formal  
• Vision established for new business models  
• Communication is effective  
• End-to-end view trained  
• Technical and organizational concepts are aligned |
| Management Attention  | All levels of management have been aligned towards the ERP programme.       | • Top management sponsorship established and engaged throughout life cycle  
• Middle management buy-in generated |
| Funding Model         | Funding model chosen supports efficient and effective ERP implementation.  | • Aligned to programme approach  
• Generates incentive for roll-out  
• Generates cost-effective requirements  
• Ensures efficient project operation |
| Human Resources       | Human resources are adequately provided to the ERP programme to fulfil its tasks. | • Top process skills available  
• Intercultural know-how available  
• Joined business and IT, global and local teams  
• Team stability is ensured  
• Team can interact face-to-face  
• Team is full-time  
• External resources are well managed |
| Governance Model      | A stringent governance model is established to manage the ERP programme.     | • Efficient scope change management  
• Governance board to handle mgmt. effectively  
• Programme management established  
• Stable objectives  
• Plans sustain of global ERP |
| Method Selection      | Selection and execution of a method for design, deployment and localization. | • Follows a method consistently  
• Blueprint is comprehensive  
• Method is made company specific |
| Tools                 | Early deployment of suitable tools.                                         | • Selected early  
• Suitable to drive efficiency |
| Technical Factors     | All technical challenges are addressed.                                     | • Technical challenges are addressed  
• Compliance, data conversion, master data, security, unicode, availability, time zones, translation, infrastructure |

Table 1. Common success factors

3.3 Site-specific Success and Endangerment Factors

A number of success and endangerment factors are site-specific. These influence the programme outcomes for each site individually and must be aggregated over all sites to estimate their impact on the overall programme. Each factor has a number of management actions or indicators associated which drive the respective success factor. All of these success factors and management actions /
indicators were verified by expert interviews in prior research. Table 2 summarizes the site-specific factors to be verified in this research.

<table>
<thead>
<tr>
<th>CSF</th>
<th>Description</th>
<th>Management Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market and Business Cultures</td>
<td>Impact of the local market and business culture has been handled adequately.</td>
<td>• Local process requirements addressed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Intercultural work aspect handled</td>
</tr>
<tr>
<td>Unwillingness to Change</td>
<td>Initial level of resistance to a change of the site and the measures addressing it.</td>
<td>• Initial level is low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Handled adequately</td>
</tr>
<tr>
<td>Inability to Change</td>
<td>Limitations in the ability of people to embrace the changes.</td>
<td>• Language abilities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Innovation potential</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Inflexibility</td>
</tr>
<tr>
<td>Necessatory preconditions</td>
<td>Site is suitable for a roll-out.</td>
<td>• Suitable size &amp; business model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Technology life cycle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Resource availability</td>
</tr>
</tbody>
</table>

Table 2. Site-specific success factors

4 RESEARCH APPROACH

The research conducted prior has led to the generation of twelve success factors for global ERP programmes. These factors were derived from the opinions of experts captured in interviews. For each success factor a number of management actions to drive each success factor were identified.

The current research step described here now brings the success factors to another level of certainty and operationalisation. It proves the validity of the success factors as valid drivers of overall ERP programme success as defined in the preceding section based on the actual outcome of a statistically relevant number of ERP programmes, focussing on outcomes rather than opinions. In addition it generates relative weighting of the success factors and the management actions driving them. This will enable a forecast of the ERP programme success based on the implementation of the relevant management actions in early set-up and blueprint phases of global ERP programmes.

4.1 Research Method

The structure to be verified with SEM was generated from the results of the qualitative research conducted prior. This is an acceptable approach according to Skrondal & Rabe-Hesketh (2004): "Instead of defining hypothetical constructs on theoretical grounds, they are sometimes 'derived' from an exploratory analysis (Skrondal & Rabe-Hesketh, 2004, p.5)". Albers (2007) states that "selecting the relevant factors [...] should be done on the basis of expert interviews and a thorough literature review (Albers, 2007, p.9)". While the analysis which has lead to the model is based on expert opinions, this research step focusses on the situation of actual programme outcomes and therefore verifies the model beyond the distortion of opinions.

There will be a number of variables which cannot be directly observed or measured – so called hypothetical constructs of latent variables. This is either because the variable is unobservable by nature (one example is 'Unwillingness to Change' as a collective attitude, which cannot be measured directly) or has not reached the final state which the model intends to predict (one example is 'Programme Objectives' which eventually can be measured but shall be predicted in early stages of the ERP programme). For these latent variables a measurement model has to be specified which defines "the relationship between latent variables and suitable indicators, which allow to measure the latent variables indirectly (Backhaus, 2006, p. 11, translation of the author)".
While the factors related to 'Programme Objective' and 'Project Objectives' are elements which describe the underlying latent variables, the causal relationship is different for the critical success factors. By the design of capturing these factors they are described as elements which can be used to change (i.e. to cause) the latent variables of the twelve success factors. This is often the case in success factor research as was pointed out by Albers & Hildebrandt (2006). Therefore the success factors are formative indicators and the ERP programme outcomes are reflective indicators. This assertion was verified by a list of criteria stipulated by Fassott (2006).

In Herrmann et al. (2006) variance and covariance based SEM approaches are contrasted. While the covariance based approaches result in a model structure to minimize covariance error terms, the variance based approaches aim for an optimal reproduction of the real data structure, i.e. the indicator values. As one of the goals of the SEM based verification is to generate a quantifiable prediction, a variance based approach was chosen in this research.

Based on the software review by Temme et al. (2006) the software SmartPLS version 2.0.M3 (Ringle et al. 2005) was chosen to conduct the analysis. The software is platform independent and without license fees. It supports graphical modelling, the bootstrapping procedure to generate significance measures, and different handling approaches for missing data.

4.2 Data Gathering

The development of the questionnaire was directly based upon the qualitative research conducted in prior research. In particular for each success and endangerment factor a number of management actions driving them were defined (see Table 1 and Table 2) which were asked individually in the questionnaire grouped by the respective success factor. The questionnaire can be found under http://www.gunter-seidel.de/phd/questionnaire_DSAG.pdf. Design guidelines to avoid common method bias as outlined in Podsakoff et al. (2003) were considered in the questionnaire design, but practical limitations of survey administration as 5 point likert scale, anonymous data collection, and non-availability of secondary sources limited their applicability. In particular the anonymity of respondents did not allow to assess whether they possess a complete picture of the EPR programme they assessed. Instrument validation as demanded by Straub (1989) was partially ensured by the development of the questionnaire based on expert interview guides and peer review, but no multitrait-multimethod comparison (comparing multiple measures of the same research subject gathered by different methods) nor pilots of the survey were conducted.

The chosen Likert scale interval data do generally violate the normality assumption. The degree of impact under different estimation methods was investigated by Muthén & Kaplan (1985) with the conclusion that “it is therefore reassuring to find that these normal theory estimators perform quite well even with ordered categorical and moderately skewed/kurtotic variables (Muthén & Kaplan, 1985 p. 187)”. In particular PLS is very robust against non-normal data.

For the sample size the recommendation for PLS based analysis as indicated by Ringle (2004) is the highest number of paths leading to any latent variable times ten (Ringle, 2004, p. 16) – in case of the final model this implies 80 data points. As PLS is considered robust against small sample size (Chin & Newsted, 1999) a number below the theoretical optimum stipulated was deemed acceptable.

The target for the survey was the DSAG (Deutschsprachige SAP Anwender Gruppe), an organisation with 2000 member companies of which more that 400 are subscribed to the globalization group. The survey was administered in a period of three months with multiple reiterations to address potential participants. It resulted in 67 admissible data sets (data sets with incomplete outcome indicators and two test data sets entered by the researcher were excluded). Although the intended sample size had not been reached, the analysis was conducted with the available data in full understanding of the limitations of applicability resulting hereof.

There are a number of missing observations for success indicators in the data set. A visual inspection of the data set resulted in the assertion of data missing at random (MAR). Therefore a number of
approaches could be used to handle the missing data set (Byrne, 2001, p. 289ff) as casewise or listwise deletion or imputation of the missing data via means, regression or pattern matching. The approach of mean replacement was eventually chosen as a method supported by the analysis tool which does not decimate the number of cases further.

The data was analysed regarding *kurtosis and skewness* with the help of SPSS 16. The kurtosis of the indicators is between -1 and +3 with an average of 0; skewness is between -2 and +.5 with an average of -.6. These values are generally not considered extreme skewness or kurtosis, nonetheless it would have distorted the covariance based SEM approach which relies on normal distribution – fortunately the same does not apply for the chosen PLS approach.

*Non-response bias* was analysed by comparing the responses received until one week before end of the initial response period, and further responses received thereafter including based on reiterated requests and extended periods. In particular the design choices of the programmes as centralization, process standardization and adherence to package standards as being the closest to demographic data were compared for these groups. Deviations between both groups were within single digit percentages indicating limited non-response bias. Due to the limited sample size a full PLS model comparison of both groups were not conducted as results were deemed not statistically relevant.

The extent of *common method bias* was assessed using Harman's single-factor test according to Podsakoff *et al.* (2003). All variables of the study were loaded into a principal component factor analysis and the unrotated factor solution examined. Although one factor accounted for 28% of the total variance, it was not concluded that neither "(a) a single factor [did] emerge from the factor analysis [n]or (b) one general factor [did] account for the majority of the covariance among the measures (Podsakoff *et al.*, 2003, p. 889)". Therefore no significant common method bias was corrected for in the remainder of the analysis.

## 5 RESULTS

Henseler *et al.* (2008) provides a general structure for validity analysis. It recommends to check for reflective and formative measurement portions according to different procedures and subsequently for the validity of the structural portion. PLS based analyses does not generate any overall goodness-of-fit measures similar to covariance based SEM outlined in Hu & Bentler (1999).

For the reflective measurement models of the endogenous variables 'Programme Objectives' and 'Project Objectives' the size of the load is one criteria to be used to assess *indicator reliability*. According to Hulland (1999) "items with loadings of 0.7 or more, which implies that there is more shared variance between the construct and its measure than error variance (Hulland, 1999, p.198)" should be accepted. Given the factor loads in the range of 0.708 to 0.921 for the reflective measurement variables the model fulfils this requirement.

*Composite Reliability* (for the reflective measurement models) can be assessed with Dillon-Goldsteins Rho with a suggested value of higher than 0.7 as benchmark for modest composite reliability. It replaces Cronbach's alpha used in regression as "in comparison to Cronbach's alpha, this measure does not assume tau-equivalency among the measure with its assumption that all indicators are equally weighted (Chin, 1998, p. 320)". Both 'Project Objectives' and 'Programme Objectives' do fulfil the requirement with a value of 0.910 and 0.803 respectively.

*Convergence validity* (for the reflective measurement models) can be assessed by a value of the Average Variance Extracted (AVE), which should be higher than 0.5 (Fornell-Larcker-Criteria). Again, 'Project Objectives' and 'Programme Objectives' do fulfil the requirement with a value of 0.772 and 0.675 respectively.
Discriminant validity (for the reflective measurement models) can be shown when the root of Average Variance Extracted (AVE) is larger than the correlation to any other latent variable. In case of the reflective measurement models 'Project Objectives' and 'Programme Objectives' this can be confirmed as indicated in Table 3. On the diagonal of the matrix the square root of AVE for the reflective measurement models are shown, below the diagonal are the correlations. The square root of AVE is larger than the correlation in the respective column and row for both 'Project Objectives' and 'Programme Objectives'.

Table 3. Root of AVE larger than correlation

Table 4. Cross loadings – row with formative indicators were excluded due to space constrains as measure for discriminant validity is only applicable for reflective indicators

Another measure for discriminant validity is the cross loading of indicators assigned to the reflective measurement model variable against any other latent variable. For the indicators 'Business
Improvement' (BI) and 'Savings' (S) as well as for the indicators 'On-Time' (OT), 'In-Scope' (IS), and 'In-Budget' (IB) the highest load is on 'Programme Objectives' and 'Project Objectives' respectively as can be seen in Table 4.

For the formative measurement portion of the success factors the "examinations of correlations or internal consistency have been argued as inappropriate and illogical (Chin, 1998, p. 306)". To rely on the path estimates it is recommended to analyse the data regarding multicollinearity indicated by the variance inflation factor (VIF). As Albers (2007) indicates that "intercorrelated factors do not imply indicators reflecting a construct but are rather the result of applying certain holistic strategies in practice (Albers, 2007, p. 10)". Therefore they are not due to be removed for unidimensionality reasons, but rather as they generate unstable results in the signs and size of the weights. With the help of SPSS 16 a multiple regression was run for each group of indicators belonging to a formative measurement against an arbitrarily chosen indicator of an endogenous variable. The analysis resulted in VIF generally below 2, which does not give raise to serious concerns regarding multicollinearity as level of concerns are stipulated to be above 10 (Henseler et al., 2008, p. 25).

No review of absolute size of loads and t-test were conducted, as these elements were confirmed in the prior qualitative research and only the load onto the objectives were of interest. Albers & Hildebrandt (2006) strongly objects to model changes based on pure statistical criteria, as the causal relationship was established prior and remains even in case of small quantifications. Only loads which signs were in contradiction to the prior established direction of causality were excluded.

Having established the validity of the measures scales the validity of the structural portion to explain the success is of interest next. As for the PLS algorithm no model improvement process is established similar to covariance based SEM the review based on path coefficients was proposed by Henseler et al. (2008): "Structural paths, whose sign is in keeping with a priori postulated algebraic signs, provide

Figure 1. Revised critical success factor model
a partial empirical validation of the theoretically assumed relationships between latent variables. Paths that possess an algebraic sign contrary to expectations do not support the a priori formed hypotheses (Henseler et al., 2008, p. 26f). Therefore from an initial n:m (many-to-many) relationship relating all success factors to both success variables 'Project Objectives' and 'Programme Objectives' all paths with a negative loading were removed to come to the final model outlined in Figure 1. The reported values are the weights for the formative measurement, the loadings for the reflective measurements, and the path coefficients for the inner model.

The measurements to be reviewed follow the logic of multiple regression. First the degree to which the model is able to explain the data set is defined by the $R^2$ value. The value of 0.503 and 0.410 for 'Project Objectives' and 'Programme Objectives' respectively define that about half of the success can be explained by the model, while the other half is driven by factors not captured in the defined success factors. Chin (1998) describes an $R^2$ of .67 as "substantial", and $R^2$ of .35 as "moderate", and an $R^2$ of .19 as "weak" (Chin, 1998, p. 323). Others define an $R^2$ of at least 50% as desirable, as it explains more than it does not explain. While a better $R^2$ would clearly be desirable, already a moderate explanation of the success of global ERP programmes can be considered an addition to the understanding of these programmes.

All prior tests were focussed on the model itself in regard to the data sample. To ensure that the model represents the population requires to assess confidence levels of the relationships. As PLS does not rely on distributional assumptions as covariance based SEM does, no implications can be made based on it. Therefore a distribution has to be generated based on the bootstrap procedure. This procedure generates a number of random samples of same size drawn from the main sample with replacement and calculates the model. The distribution of each path loading can be assessed by the students t-test. According to Albers (2007) "the result of a significance test heavily depends on the number of investigated cases or other non-controlled effects. Therefore, we get richer information if we determine the level of impact that different drivers have on business performance. Insofar we advocate that not significance testing is the main purpose of success factor studies but the determination of the parameter levels (Albers, 2007, p.15)". Therefore the significance levels have been added to Figure 1.

While the figure shows that all indicators to the reflective measurement model are of a confidence of above 99%, some of the indicators to the formative measurements are even below the 85% confidence. Nonetheless as stated by Henseler et al. (2008) "the researcher should keep both significant and insignificant formative indicators in the measurement model as long as this is conceptually justified (Henseler et al., 2008, p. 25)". Given the source of the success factors (expert interviews) none of the factors were removed due to confidence considerations. On the other hand for the paths with lower confidence a substantial relationship remains to be verified. Therefore a review based on a larger data set would be desirable.

In addition based on the weights in Figure 1 a relative impact of each management action to the success factor and a relative impact of each success factor to either 'Project Objectives' or 'Programme Objectives' or both was established. The results can be used for predictive purposes when the implementation or intention of a management action can be gathered. To apply the weights on unstandardised raw scores the effect of standardisation must be reversed. SmartPLS offers the output of index values for the measurement model which are unstandardised and sum to 1 for each CSF (personal communication via SmartPLS forum, Diógenes de Souza Bido, Universidade Presbiteriana Mackenzie, Sao Paulo, Brasil). Building a sum of each index value multiplied by the path coefficient to 'Project Objective' or 'Programme Objectives' leads to an overall impact factor for each management action.

### 6 CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

While the model derived from this research fits the data reasonable well, it should always be noted that any "model - regardless of how well it fits the data - remains only one possible means of its [the data's]
description and explanation (Raykov & Penev, 2001 p. 298). A replicated analysis therefore should always start again with a full many-to-many relationship of all success factors to both 'Programme Objectives' and 'Project Objectives'.

The data set gathered allowed to verify a positive impact of 34 of the initially gathered 41 management actions to drive success of the global ERP programmes (which does not imply the other must be rejected, rather that the limited sample does not allow to verify them). The positive impact of each of the twelve success factors to either 'Project Objectives' or 'Programme Objectives' or both could be verified, although not always generalized to the whole population due to insufficient significance levels. These factors jointly allow to explain about half of the success of global ERP programmes.

Future research to improve the extent of success prediction could take into consideration the lifecycle of the ERP programme, design choices of the programme as mediating factors, a more detailed analysis of the dependent variable of success, or a replication of the survey in a different national culture to compare the impact of management cultures.

To allow prediction of success of global ERP programmes the individual contribution of each management action to either 'Programme Objectives' or 'Project Objectives' or both was established. This will allow to forecast the success probability of a global ERP programme once the implementation of the respective management actions has been decided in an early set-up or blueprint phase of the programme.

References


“AVOIDING MANAGEMENT” OF RESISTANCES DURING IT PRE-IMPLEMENTATION PHASE: A LONGITUDINAL RESEARCH IN A HIGH TECH CORPORATION

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0110.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Conflict management styles, Resistance to change, Enterprise resource planning (ERP) (packaged systems), Implementation</td>
</tr>
</tbody>
</table>
“AVOIDING MANAGEMENT” OF RESISTANCES DURING IT PRE-IMPLEMENTATION PHASE: A LONGITUDINAL RESEARCH IN A HIGH TECH CORPORATION

Meissonier, Régis, CEROM, GSCM - Montpellier Business School, 2300 avenue des Moulins, 34185 Montpellier, France, regis.meissonier@gmail.com

Houzé, Emmanuel, CREGOR, IAE, Montpellier II University, place Eugène Bataillon, 34090 Montpellier, France, emmanuel.houze@univ-montp2.fr

Abstract:

Most of empirical research about users’ resistance toward IT has been conducted after IT been implemented in organizations surveyed. Few longitudinal research have been done about the way individual and group resistances emerge and evolve during prior stages of projects. This focus on pre-implementation phases is all the more important that IS managers need to anticipate potential conflicts and users’ resistances likely to involve project failure. This article delivers the results a two year longitudinal research conducted at Netia corp. (a worldwide leader in video and audio broadcasting) during preliminary phases of its ERP implementation project. As main findings, while conflicts toward IT implemented are often considered as having negative effects and requiring to be actively managed by the hierarchy, the case study delivers an alternative observation: it describes how an affective oriented conflict has been solved while managers adopted an “avoiding management style”. Our observations differ from several prior studies about conflict management styles and support that an avoiding management style can drive team’s members to cope efficiently with conflict situations during IT pre-implantation phase. In conclusion, the article presents research perspectives associated to these results.

Keywords: IT implementation, ERP, user’s resistance, conflict situations.
1 Introduction

Understanding key factors contributing to IT adoption in organisations is a central concern in information system research. Among key factors associated to IT project failures, users' resistance is one of the most salient because related to human resistance to change (Jiang, Muhanna et coll. 2000). Existing literature on IT resistance provides practical knowledge about underlying conflicts types and conflict management styles' performance (Cramton 2001, Montoya-Weiss et coll. 2001, Barki & Hartwick 2001, Markus et coll. 2000, Miranda & Bostrom 1993). However, most of these researches has been empirically conducted after IT been implemented in organizations surveyed and can be considered as observations made on downstream results of upstream resistance process. As a consequence, a large part of resistances are observed as task oriented and related to the non appropriateness of IT users have to cope with. Little empirical investigations were done about the way individual and group resistances emerge and evolve during prior stages of projects (Lapointe & Rivard 2005) while negotiations about IT to implement can raise affective oriented resistances if users perceive threats about their values or power relationships because of organisational changes expected. A focus on pre-implementation phases is all the more important that IS managers need to anticipate potential conflicts and users' resistances likely to involve project failure (Marakas & Hornik 1996, Joshi & Lauer 1998, Robey et coll. 2002, McAfee 2007).

Because enterprise systems are considered as ones of most impacting IT on future actions (Jiunn Chieh Lee & Myers 2004) because of their cross-functional perspective (Markus et coll. 2000) and readiness to change (Kwahk & Jae-Nam Lee 2008), we decided to report resistance evolution toward ERP adoption project during pre-implementation phase. To contribute to this issue, the article is structured as follow. The literature analysis reviews conceptual foundations of resistance, conflict and conflict management styles associated to IT implementation. The case study analysis delivers the results a two year longitudinal research conducted at Netia corp. (a worldwide leader in video and audio broadcasting for TV and Radio channels). Firstly, our observations revealed that task oriented conflicts expressed by users actually hid an affective oriented conflict. These resistances required the abortion of the ERP initially considered for a less impacting application on some specific process changes and on underlying power redistribution across group of employees. Secondly, we observed how this conflict between developers and administration employees switched to a solving solution while managers adopted an avoiding management style. Whereas conflicts toward IT are often considered as requiring to be actively managed by the CEO (Markus et coll. 2000, Barki et coll. 2001), the case study delivers an alternative observation. In conclusion, the article considers users' resistance toward IT as not a systematic negative behaviours aiming project abortion, and invites researchers to explore how task oriented and affective oriented conflicts can turns out to be key processes embedded in information system design.

2 Literature review

In management and organisation theories, the political school of thought developed by famous authors like Mintzberg (1998, 2002) or Crozier (1977), considers strategy formation and implementation shaped by power and political plies. As a consequence, strategical project usually involves shifting coalitions of dominant actors of parochial interests (Jiunn Chieh Lee et coll. 2004). Even if lot of research in IS proposed understanding IT user’s resistance toward a deeper approach (Joshi 1990, Krovi 1993, Joshi et coll. 1998) it seems marginal compared to articles published on the subject. Lapointe & Rivard (2005 p.462) revealed among 43 articles published during last 20 year period about user’s resistance toward IT, only 4 did not settle for considering resistance as a factual characteristic of the context. The majority of these studies treated users’ resistance as a component of an organisational system at individual and group level (Markus et coll. 2000) and only a minor part of these studies were devoted to study the causal conflicts (Jiang, Klein et coll. 2000). While literature stress on “resistance” or “conflict” without making clear differences between both concepts, our analysis, both based on psychology and sociology theories, incites to consider resistance as the behavioural dimension of conflict.

Resistance literature background

User’s resistance is defined as a subjective process psychologically based at individual level (Jermier et coll. 1994). It corresponds to behaviours in reaction to a present or ongoing situation perceived as negative (Ang & Pavri 1994), as inequitable (Joshi 1991), as a threat or as a stressing feeling (Marakas et coll. 1996). According to Joshi (1991) resistance appears when user perceives changes involved by an “unfair” project in regard to his/her personal work or in regard to the group he/she belongs to.
Users can express resistance toward IT with an active form (visible and relatively easy to detect) or in a passive form (hard to detect and difficult to deal with) (Tetlock 1999, Tetlock 2000, Jiang, Klein et coll. 2000). Empirical studies shown that resistance is higher at group level than at individual and organisational levels (Lapointe et coll. 2005). In other words, group of persons (depending on their professional category, professional competencies, age, gender, etc.) represents the more likely unit to develop high resistance toward IT. Indeed, at group level, users’ resistance is often socio-political whereas at the individual level it is more psychological (Markus 1983). Coetsee (1999) identified 4 types of resistance expressions:

- **apathy** corresponds to attitude of disinterest and inaction of a person toward the situation;
- **passive resistance**: a person adopts some behaviours aiming for slowing down changes and keeping the previous system (examples: voluntary delays in task to do, argumentation in favour of so-called advantages of existing rules and processes);
- **active resistance** is considered as a “constructive form” (examples: expression of different points of views, negotiation about a consensus, accommodation);
- **aggressive resistance**: users can resort to threats, blackmails, boycotts and all other actions whom objective is blocking the situation.

According to the author, these forms are not exclusive and should be considered as part of a continuum encompassing on the other extreme user acceptance and involvement.

**Conflict types**

Conflict is defined as a disagreement of persons or groups of persons considering a situation as inconsistent with their own interests (Boulding 1963, Robbins 1974, Putnam & Wilson 1982, Hocker & Wilmot 1985). A conflict can oppose somebody to himself or herself (internal conflict), to other persons, groups of persons or to institutions (Thomas 1992). Several definitions synthesis made in organization theories (Putman & Poole 1987), psychology (Thomas 1992) or information systems (Barki et coll. 2001) considers three properties of interpersonal conflicts: interdependence, interference and disagreement. By itself, each property can not be considered as a sufficient condition. Interpersonal conflicts are more dependant of their overlapping.

- **Interdependence** exists when each party reaches a specific goal, at least because of the actions of the other party. In essence, interdependence is a structural condition for conflicts in a professional context because of respective consequences of the way the other party acts.
- **Interference** is a behavioural condition for conflict and occurs when one or several parties opposes the other party's attainment of its interests, objectives, or goals. Interference thus represents the central behavioural node of any conflict (Barki et coll. 2001 p.198).
- **Disagreement** is a cognitive condition for conflict and correspond to divergence of interpretations toward values, objectives, needs, methods, etc. Disagreement refers to disputant behaviours and is considered as the central process associated to conflict (Wall & Callister 1995).

While first and second properties sounds like relational configuration associated to conflict, the last one deals with upward causes. In professional contexts, these causes can be task (or process) oriented versus affective (or relational) oriented (Deutsch 1969, Pinkley 1990, Jehn 1995, Jehn & Bendersky 2003). Conflicts about tasks are issue oriented and arising from differences between professional missions to be performed, whereas affective conflicts refer to personalized disagreements or individual disaffections. The first ones can be considered as differences of points of view rarely assorted of negative emotions while the second ones can raise frictions and tensions which can affect team performance (Jehn & Mannix 2001). We distinguished 4 different conflict types drawn from task and affective orientations (see Table 1).
## Table 1: Conflict types associated to IT implementation

<table>
<thead>
<tr>
<th>Conflict types</th>
<th>Key authors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Task oriented</strong></td>
<td></td>
</tr>
<tr>
<td>Conflicts about the definition and the execution of tasks that users must fulfil</td>
<td>Robey et al., 2002; Markus &amp; Tanis, 2000</td>
</tr>
<tr>
<td>Conflicts about the new professional skills required</td>
<td>Besson et al. 1999; Markus &amp; Tanis, 2000; Newman &amp; Westrup, 2005</td>
</tr>
<tr>
<td><strong>Affective oriented</strong></td>
<td></td>
</tr>
<tr>
<td>Value conflicts</td>
<td>Besson et al. 1999; Ménard &amp; Bernier, 2004; Kohli &amp; Kettinger, 2004; Leidner &amp; Kayworth, 2006</td>
</tr>
<tr>
<td>Conflicts due to a loss of power</td>
<td>Markus, 1983; Hart &amp; Saunders 1997; Watson et al. 1999; Jasperson et al., 2002; Bancroft-Truner &amp; Morley, 2002</td>
</tr>
</tbody>
</table>

Conflicts about task definition and execution are caused by the way organisational processes have to be adapted or transformed to fit with IT process requirements (for examples: how invoices and orders must be established, new data codification, signature validation process). These conflicts can be “internally initiated” when users compare the way they achieve their tasks and perceive organisational inconsistencies (Besson 1999). They can also be “externally initiated” because of the process constraints imposed by information technology to be implemented. For instance, ERP standard modules represents one of the most well known conflict driver because of new “best practices” imposed to employees without too much consideration of organisation specificities (Davenport 1998, Markus et coll. 2000, Lim et coll. 2005). This type of misalignment with organisation processes (Hsiao-Lan et coll. 2005) is all the more important that problems in MIS are more about the ability of users to understand how they must carry out their new tasks than ability of the firm to manage change (Robey et coll. 2002).

Conflicts about new professional skills deal with competences users must develop in order to be qualified to job transformations involved by IT (Markus et coll. 2000, Besson & Rowe 2001). Accountancy is one of the most salient professional illustrations: before ERP implementation during 90's, an important part of daily work of these employees consisted of collecting, aggregating and synthesising a huge quantity of financial data. Enterprise applications change dramatically their assignments: being no more the ones who collect financial data, they are asked to interpret these information’s ex-post, to make sense and recommendations to top managers (Bernard et coll. 2004).

Value conflicts are psychologically based. They refer to ideology by which some people share beliefs and make sense of their worlds (Trice & Beyer 1993). Firm subunits may have their own subculture varying in their ideological content (Stewart & Gosain 2006). In IS, value conflicts may arise on inconsistency between cultural principles of users or group of users and the perceived underlying strategic objectives assigned to IT implementation (Leidner & Kayworth 2006). Several empirical researches (Besson 1999, Kohli & Kettinger 2004, Ménard & Bernier 2004, Bhattacherjee & Hikmet 2007) revealed how these conflicts raised in the hospital sector. For example in his study, Besson (1999) observed that financial control allowed by the ERP over all the hospital activities was perceived by medical employees as an attempt of a market based activity inconsistent with fundamental principles of health public services. The empirical analysis of Wagner & Newell (2004) revealed complementary observations according to which ERP can be problematic for organisation sub-cultures because mandating one epistemological position through the software design based on “best practices”.

Power conflicts concern the way individual autonomies and capabilities of influence are likely to be redistributed among employees after IT implementation. Research in IS challenged understanding of IT development and implementation deviations by pointing out intricacies due to power influence exerted by actors (Markus 1983, Davis et coll. 1984, Markus & Bjorn-Andersen 1987, Jasperson et coll. 2002, Avgerou & McGrath 2007). On one side, IT can give more power to key users by allowing them to use real time data access functionalities (Davenport 1998). On the other side, IT can reduce the autonomy of employees (Markus 1983). Despite hierarchical monitoring supported by IT, power loses for employees may be caused by more interdependencies with colleagues. For instance, in civil
engineering project management, ERP implementations changed the way main actors (project supervisors, architects, electricians, plumbers, etc.) collaborate (Gilbert & Leclair 2004). Formerly, they did not have to communicate to their colleague the details and calculations on which their analysis and conclusion were based. The integration of processes associated to IT looks like a management of interdependencies (Rockart & Short 1995) by which actor become prescriber of conditions and means of his colleagues. As a consequence the political perspective in terms of power distribution misfit appears to be primarily applicable for cross-functional IS (Markus 1983).

Actually users' resistance forms are not exclusive and can occur simultaneously. However, emotional conflicts are considered as highly contagious (Hatfield et coll. 1993) and likely to overshadow or dominate congruous task oriented conflict (J. Ford et coll. 2008 p.369). Actually, MIS literature based on the interaction theory (Joshi 1992) considered that the fundamental reasons of resistance toward IT systems are not the ones expressed about the system nor persons characteristics, but users' perceived values and social content gain or loss before/after system implementation (Jiang, Klein et coll. 2000, Kendall 1997). Indeed, advocating system inconsistencies or organisational misalignment is probably a more comfortable resistance strategy than the one consisting to express underlying individual socio-political challenges. In this research we assume that users having affective oriented conflicts related to IT project are likely to use a bypassing strategy and to express their resistance only with task oriented conflicts. Following this reasoning, we formulate the following research proposition:

Proposition 1: expressed task oriented conflicts toward IT to be implemented may hidden affective oriented conflict.

Resistance management styles
IT projects can rarely be properly completed without any implication of the CEO. Often, top management ought to appear as “sponsor” of the projects in order to promote their credibility toward employees (Davenport 1998, Markus et coll. 2000). CEO should be able to balance the choice that must be made between satisfaction of individual expectations and the general objectives of IT projects in order to manage efficiently conflict resolutions. There are three common conflict resolution approaches: integrative (solving the problem through collaboration), distributive (solving the problem through assertion), and avoidance (ignoring the problem) (Sillars 1980). The integrative approach aims to identify and achieve outcomes perceived as satisfactory to all team members. These approaches support also previous studies that demonstrated the preference of IT users toward participative resistance management methods in opposition to direct management methods imposed by managers (Robey & Taggart 1981, Ives & Olson 1984). The distributive approach yields outcomes that favour some team members but not others. The avoidance approach consist for managers not intervening in the conflict and relying on the team capability to self resolve the conflict.

Within the conflict domain, many studies have been done to examine the management and resolution of conflicts, identifying a number of conflict management styles and their role in achieving satisfactory outcomes (Barki et coll. 2001, Kankanhalli et coll. 2006). We identified five different management styles using the common conflict resolution approaches: problem-solving, compromising, asserting, accommodating and avoiding (see Table 2).

<table>
<thead>
<tr>
<th>Integrative approaches</th>
<th>Problem solving</th>
<th>Managers identify conflict causes and solve them looking for optimal solutions. Problem-solving occurs when managers try to fully satisfy the concerns of all parties.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Compromising</td>
<td>There is no optimal solution to the conflict; managers try to find a satisfactory solution by splitting the difference where each party win some and lose some.</td>
</tr>
<tr>
<td>Distributive approaches</td>
<td>Asserting</td>
<td>Authoritarian decisions are made and imposed by managers to users. Conflicts are considered as win/lose situation.</td>
</tr>
<tr>
<td></td>
<td>Accommodating</td>
<td>Managers give up their preferences and satisfy users’ claims. Accommodating occurs as managers must adapt, or cooperate with users in an attempt to reduce conflicts.</td>
</tr>
<tr>
<td>Avoiding approach</td>
<td>Managers do not intervene in the conflict and hope for the</td>
<td></td>
</tr>
</tbody>
</table>
situation to resolve by itself. Avoiding occurs when managers refuse to act and to participate in conflict situations.

<table>
<thead>
<tr>
<th>Table 2: Management styles of IT resistances</th>
</tr>
</thead>
</table>

According to Montoya-Weiss et al. (2001) integrative and distributive approaches appear to facilitate team performance whereas the avoidance approach seems to hinder. In their empirical analysis conducted on IS staffs and future users of 162 IS projects, Barki and Hartwick (2001 p.218) observed asserting mode and avoiding management as associated to negative results in terms of interpersonal conflict solving. In others words, the two most opposed styles were considered as inefficient techniques. In the same time, authors considered that negative emotions involved by interpersonal conflicts are not only negative experience, but negatively affect IS project outcome and remain pervasive even when properly resolved (op. p. 220). However, resistance management styles can not be considered as exclusive. MIS literature shows that, depending the project budget, the delays, the evolution of employee perspectives, etc., project managers are likely to change their style several times during the project duration. For instance, Gibson (2004) describes how during an ERP implementation project at Dow Corning corporation, resistance management style evolved from an “improvisation approach” to “big bang” assertions. Then, for a large part, information systems literature incites managers to not remain passive (Leidner et coll. 2006 p.381) and to solve users’ resistance by identifying conflict situations in order to prevent a project or an on going situation to evolve negatively toward IT implementation. On an other side, some empirical studies showed that conflict situations managed by team members were linked to conflict reduction (Kankanhilli et coll. 2006) or team performance improvement (Jehn et coll. 2001). In other words there would have no evidence that depending the context a management style relying on teams self ability to resolve resistance can not not be suitable. However, these observations have been only done on task oriented conflicts and do not permit to expand the corresponding assumption to affective oriented conflicts.

Proposition 2: avoiding management style is not associated to positive results in the case of affective oriented conflict.

This literature review on conflict situations, user’s resistances and management styles toward IT implementation, represents the theoretical background we used to analyze Netia case study. The longitudinal research conducted explores user resistance causes and conflict situations lying behind preliminary phase of ERP implementation while managers decided not to intervene.

3 Case analysis

In general, conflicts in organisations evolve over time which justify the higher adequacy of longitudinal research methodology than the one of static analysis (Jehn et coll. 2001 p.239). This method is often used in IT implementation studies (Molla & Licker 2001) and recommended for Small and Medium Enterprises’ analysis (Chetty 1996). There is also an interest for using a single case study which delivers illustrative stories (Benbasat et coll. 1987). So, we adopted a longitudinal research methodology for Netia from the beginning of 2005 until the end of 2006 using standard techniques of case studies analysis (Miles & Huberman 1984, Eisenhardt 1989, Yin 1994).

Case description

Netia corporation (located near Montpellier, France) is one of the worldwide leaders in broadcasting (40 countries covered). Its customers are TV channels and radios like, BBC, ABC, Rai uno, Canal+, France Télévision, etc. Created in 1993, the company employs an hundred of persons spread over two sites in France and subsidiaries abroad (Amsterdam, Liège, Rome and New York). The firm is an IT service agency dealing with development of audio and video data digital solutions. Besides IT development, Netia offers implementation management services (consulting, process analyses, engineering, training, maintenance and evolution of audio and video data digital solutions.). The information system of Netia has been developed progressively by ad-hoc initiatives. These isolated and independent developments have been involving a lack of data coherence as well as an excessive growth of applications. Consequently, a large part of employees’ tasks was dedicated to re-entering data in order to feed all of the redundant applications implemented to respond to local needs. For example, the management control service developed a set of Excel macros to partially deal with a divided utilisation of SAGE accountancy software. Each process (order forms, delivery forms, etc.) corresponds to a data entry for one or more shared Excel files (on the server there is a file for the order forms, another for the clients, another for prospects, etc.). The information system was structured
around a huge quantity of office files from which data were manually extracted and aggregated into other files to produce performance indicators required by managers.

Thus, a loss in productivity raised because of repeated data entries and redundant procedures. The lack of IS integration was also highlighted by data access problems. For example, the project coordinator did not know the status of the client order in progress; he had to contact directly the logistic service which browsed the SAGE application. Given that, transaction histories were dispersed throughout several isolated applications and purchases' tracking was hazardous to carry out. Customer invoices were not automatically triggered by a delivery note; the logistic staff had to type corresponding informations in a shared Excel file with the account number in order to edit the invoice. Due to these inconsistencies, administrative employees asked for the implementation of an integrated information system to ensure a more coherent and efficient management of the daily tasks.

Our research has been articulated with two phases aiming to identify explicit and tacit causes lying behind the ERP adoption abortion and, later, the implementation of the Genesys software.

**Research design**

To reach an appropriate degree of internal validity, we used the same three sources of evidence as the ones used in multilevel analysis of resistance to IT (Lapointe et coll. 2005): interviews (during the first step); direct observation (during the second step); document analysis, informal discussions and records of events (all along the project). These several data sources allowed us to achieve triangulation in order to ensure satisfactory information interpretations (Yin, 1994).

The aim of the first step analysis (from January 2005 to November 2005) was to identify explicit and tacit causes explaining why the firm had failed at the ERP implementation preliminary phase. This approach was explorative and consistent with thematic analysis (Boyatzis 1998) where codes must be constructed inductively. To carry out the analysis, 8 semi-directive interviews were conducted over 4 months. Even if the overall activity of the firm was highly technological, an understanding of different levels of culture was important to study IT implementation (Leidner et coll. 2006 p.358). It was relevant to analyse how the co-existence of subcultures had influenced conflict situations which involved the IT project abortion. Interviews were realised with key employees of firm departments (see Table 4 in appendix). The interview grid used had been conceived with reference to the risk factor lists of Markus et al. (2000), Akkermans & Van Helden (2002), Besson et al. (1999). The interviews were realised in a one-to-one interaction with an anonymous format response gathering. During the first part, the employees interviewed were asked to select on the grid the factors he/she considered as explaining the rejection of the ERP implementation. In a second part, we asked him/her to explain the causes he perceived as associated to ERP implementation project. Interviews lasted around 90 minutes approximately and were audio-tape recorded in order to avoid potential biases of only one interviewer interpretation. We completed this first step analysis by several formal and informal meetings with key actors of Netia in order to perceive users’ resistances and conflict situations toward the preliminary phases of the ERP implementation.

A second step of analysis (from March 2006 to October 2006) was conducted when, after several invitations to bid, a package editor (Genesys corp.) was asked to present its software. The presentation has been done in front of the Netia employees concerned by the IS implementation (see Table 5 in appendix). We took the advantage of being invited to this meeting to analyse the direct reactions of employees. To avoid any suspicions about our presence, Netia managers presented us as academic researchers interested by IT solutions for firms without any role to play concerning the project. The passive observation method we used was consistent with Yin (1994) who considers this technique like an additional source of data useful to understand the social context of the firm. To control the risk of instrumental biases involved by observational methods (Weick 1968), both authors attended the meeting and aggregated, latter, data collected. The meeting lasted 3 hours and took the form of a presentation and discussion about the software functionalities. Seeing directly on the screen the usability of the product, participants were able to ask questions all along the presentation. This type of interactions allowed us to note verbal and non verbal users’ behaviours.

**4 Results**

**Step 1**

During the first step analysis, computer department employees expressed an aggressive resistance toward the ERP implementation project which was considered as inappropriate to the needs of the organisation. This conflict situation between computer department employees and administration employees was consistent with prior studies which showed that cultural differences within organisations tend to influence contrasted interpretations of IT to be developed (Dubé 1998,
Programmers represent a key competence asset for Netia. In fact, the broadcast software applications developed by the company are in no way standard package applications that can be bought on the market. Consisting of solutions billed for several thousand of euros, these programs ensure storage, management and broadcasting of audio and video programs for TV and radio channels. Therefore, very specific skills are required regarding sound, image, storage (on servers of several terabytes), and data diffusion by hertzian, satellite or GPRS transmissions. The programmers in the company represent a rare workforce on the professional market and this gives them a strong negotiation power towards the hierarchy. Thus, they have gained overtime strong independence in the way they organize their job. “I decide my own objectives!” declared a program coordinator. An administration coordinator described for us the example of holiday management: “The programmers are used to freely organize their work depending on the tasks and on the assignments to be completed. They do not really respect the process for taking holidays. Instead of filling out the holiday sheet and having it validated by managers, the requests (when they are made) usually take the form of an informal conversation”. The implementation of the ERP was perceived by programmers as inconsistent with their ad hoc processes and their autonomy. Considered as a “spy eye”, such ERP system was considered as a threat for their own autonomy.

The top managers avoided any risky decision - in the sense of Cyert and March (1963) - and adopted a passive management – in the sense of Cooke and Lafferty (1987). CEO never interfered in the conflict situations and did not decide to impose this unpopular solution to programmers and preferred to let all employees finding a compromising solution by themselves. An administration coordinator stated: “If we really wanted to impose a standard solution, we could. However, this would mean interfering with the programmers. But they are the makers of the programs sold, so...” Because there has been no concrete or major prejudice due to the unreliability of the existing applications used, managers were not particularly motivated to settle this situation and to take a decision likely to disturb the social climate. “Regarding the successful implementation, the management favours the programmers, only the programmers... The rest, such as improving the organization, is not considered as crucial”. These observations illustrate that IT employees are mainly rewarded for delivering technically sound systems on time and to budget and are not really encouraged to consider organizational issues in IT systems (Hornby et coll. 1992 p.165).

Step 2

Because of the programmers disagreement about the ERP implementation project and the passive attitude of top managers to solve the conflict situation, administrative employees decided to look for less impacting software’s from an organisational point of view.

Among the commercial propositions received, administrative employees of accounting service considered Genesys application as an interesting alternative. Its functionalities covered most salient needs of administration employees: customer and potential customer management, sales and procurement management (quotations, order and invoice tracking), treasury, after-sales management, etc. Then, the application was focused on process management of administration employees without implying cross-functional processes. In other words, the software could not be considered as an integrated information system forcing programmers to cope with badly perceived tasks like reporting their daily work or filling out electronic forms to have holiday demands validated by the managers. Moreover, it was interoperable with SAGE application and was not requiring data migration from the existing database.

During the presentation meeting we attended to, both computer and administration representatives found this new solution satisfactory towards the needs previously expressed (during step 1). The application was supposed to be used only by administration employees. Computer department representatives only made remarks and asked questions about technical specificities of the software. Some allusions to the previous recalcitrant behaviours of developers about the ERP solution were expressed with irony and the general laugh reaction allowed us to observe an alleviation of the initial conflict between administration and computer employees.

Despite the general positive impression about the application, an active form of resistance appeared few minutes later when conversations converged on the required task reconfigurations. For example, because of his frequent travels abroad, the Asia commercial agent of Netia mentioned some practical problems not treated by Genesys functionalities. This employee was used to type a text file within
which he added complementary information and comments about potential customers. Then, he uploaded the file on Netia server in order to make it available to other employees. But the customer management function of Genesys software did not allow joining complementary files like that. So, he firstly considered this as an annoying limit of the application toward his daily activity. Then, the discussion moved on how to overcome this problem until one programmer noted that it was more related to the task definition than the software appropriateness to user needs. Actually, Netia employees were used to include in “transaction” concept all upstream processes to the order (quotations, bargaining, etc.) whereas in Genesys application those tasks were included in an other functionality than the one talked about. Actually resistances expressed during this step were essentially because of ambiguity in professional jargon between Netia and Genesys corporations.

Few minutes later, another active resistance raised while Genesys engineers were presenting the treasury management function. The Finance Director was reluctant to use this function because she explained that it only satisfied a minor part of the activity. According to Netia practices she revealed that on-going payments of invoices sent to customers were included in treasury while not yet cashed. If this practice may sound as inconsistent with accounting classical rules, it sounded consistent with Netia business practices. Indeed, the recovery rate of customer debts is always 100% and paid immediately when the invoice is received. So, any invoice sent to customer are considered as existing cash. However, at the end of the meeting, all employees agreed on the global adequacy of Genesys ERP for Netia needs, only task-related oppositions remained and were resolved. As a result, at the end of 2006, the decision was made to implement Genesys solution and the software was bought. Table 3 presents the evolution that has known the user’s resistance, conflict types, and conflict management styles toward the IT implementation project.

<table>
<thead>
<tr>
<th>Employees concerned</th>
<th>Step 1</th>
<th>Step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer dept. employees</td>
<td>Administrative dept. employees</td>
<td></td>
</tr>
<tr>
<td>User’s Resistance Types</td>
<td>Aggressive</td>
<td>Active</td>
</tr>
<tr>
<td>Conflict Types</td>
<td>Power oriented</td>
<td>Task oriented</td>
</tr>
<tr>
<td>Conflict management styles</td>
<td>Avoiding style by top management</td>
<td>Avoiding style by top management / compromising style by administrative employees</td>
</tr>
</tbody>
</table>

Table 3: resistance and conflict evolution observed

5 Discussion

Our research puts the emphasis on longitudinal research, versus cross-sectional data collection, to analyse the dynamic nature of conflict and user’s resistance during project steps prior to the IT implementation. We believe that researchers should consider these “upstream resistances” like additional influencing factors of IT adoption to explore in order to expand existing theoretical models.

We use this French corporation case study to observe how user’s resistances and conflict situations associated to an IT implementation project evolved over time. Therefore we can not pretend the same generalisation of the result as if we had used several case studies and quantitative analysis. During the first step of analysis, data were mainly collected through interviews which likely induce some interpretative biases on the feelings expressed by interviewees. However, we tried to reduce these biases by interviewing several employees of each department, using a grid to help respondent to identify and formalize the factors he/she considered as explaining the resistance to the ERP implementation, comparing the data collected, adding informal meetings, etc. During the second step of analysis, data were mainly collected through observation techniques during the Genesys presentation to Netia employees. Our presence might have influenced the way persons behaved and participated during the meeting even if we had been presented as having no role to play on the decision process of the IT project. Moreover, one inherent limit of longitudinal research is that the processes observed continue to evolve after the end of the research investigation (Volkoff et coll. 2004 p.302). Further research should be done in order to study other findings in other cultural, structural (large firms), professional and organisational contexts to give a deeper understanding of user’s resistance and conflict situations during a longer longitudinal research which would cover all the IT project life-cycle.
Nevertheless, by exploring conflict evolution during pre-implementation project phase, our results offer additional contributions to IT users’ resistance research. Firstly, task oriented conflict expressed during step 1 hindered affective oriented conflict related to the autonomy treat perceived by developers (our proposition 1 is confirmed). This observation is consistent with previous studies done on value and power conflicts associated to IT implementation (Markus 1983, Hart & C. Saunders 1997, Jasperson et coll. 2002, Kohli et coll. 2004, Leidner et coll. 2006) and is in line with recent investigations of Ford et al. (2008) who observed that emotional conflicts can dominate task conflicts in organisations.

Secondly, resistances moved from an aggressive form (observed during step 1) to a constructive form (observed during step 2) which led to the implementation of an alternative IT solution. The evolution showed that conflicts are not fixed and our results are in line with observations of Jiang et al. (2000 p.32) who observed causes of users’ resistance differed according the IT type to be implemented. Our investigation illustrates how an affective oriented conflict related to IT has been solved during these preliminary phases while top managers adopted an avoiding management style (our proposition 2 is rejected). Netia case study do not support the conclusion of Barki and Hartwick (2001) who observed avoiding management style as associated to negative results in terms of conflict solving.

Because our longitudinal observations delivers the story of one conflict management style we can not assume the effects other conflict management styles would have provided and we can not consider any intrinsic superiority of the avoiding style on other styles. However, our results are in line with the attribution theory (Cramton 2001) and some empirical studies which showed that conflict situations managed by team members were linked to conflict reduction (Kankanhalli et coll. 2006) or team performance improvement (Jehn et coll. 2001). However, these studies have been mainly realised on task oriented conflicts whereas our observations extend the results to affective oriented conflicts which are considered as more difficult situations that managers prefer to avoid to be engaged with (Edmondson & Smith 2006 p.25).

Concerning MIS literature, the article expands the empirical researches which observed the lack of “organizational fit” as a failure cause of ERP implementation (Hong & Kim 2002, Hsiao-Lan et coll. 2005). We could consider our results as a possible extension of these results in the sense we observed the “fit”not limited to the adequacy of IT to business but covering also underlying organisational change consistency with value principles of firm sub-culture units. Indeed, when an organization is composed of several sub-cultures, the use ERP can be problematic because mandating one epistemological position through the software design based on “best practices” (Wagner et coll. 2004).

For IS practitioners, our study suggests a greater attention to issues relating to power, autonomy and professional sub-cultures when implementing IT. The main practical implication of this paper for managers is inciting them not considering task oriented conflicts expressed by users, as sufficient informations to understand whole resistance causes related to IT projects. Discovering and understanding potential underlying affective conflicts about users’ values or power losses turns out to be necessary before deciding the IT to implement.

6 Conclusion

The underlying message of this article is being out of considering users' resistance toward IT as a negative behaviour toward the organisation effectiveness. By considering resistance as dysfunctional conflict, IT project managers can disregard its potential contribution to the change and implementation. In organisations with absolutely no resistances, employees would accept all change projects including those detrimental to the organization effectiveness (J. Ford et coll. 2008). Affective oriented conflicts are not necessarily about present change involved by the project but can be related to unresolved issues form previous changes (Reichers et coll. 1997). So they can be interpreted as appeals for some managerial rectifications, like restoring trust or professional recognition of employees, which should be taken into account in the design the IT to implement. As a consequence, decisions made about the implementation without consensus are likely to involve systems’ usages very different from the ones expected by managers (Soh & Kien Sia 2004). As future investigations, we incite researchers to explore how both task oriented and affective oriented conflicts should be considered as consistent with inscription theory (Orlikowski 1992) and assumed as key processes embedded into IT choices and information system design.
References


Joshi K. (1990), "An Investigation of Equity as a Determinant of User Information Satisfaction", *Decision Sciences*, 21(4), 786-807.


Ménard C. & Bernier C. (2004), "Le cas d'une mise en oeuvre ERP réussie aux centres hospitaliers de LaSalle et de Verdun : comprendre la démarche par laquelle s'installent les nouvelles façons de faire", *Gestion*, 28(4).


Proceedings ECIS 2009
7 Appendix

<table>
<thead>
<tr>
<th>Initials</th>
<th>Department</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>VB</td>
<td>Accounting</td>
<td>Management coordinator</td>
</tr>
<tr>
<td>AG</td>
<td>Computer Dept.</td>
<td>Computer Dept. Coordinator</td>
</tr>
<tr>
<td>PV</td>
<td>Computer Dept.</td>
<td>Software developer</td>
</tr>
<tr>
<td>SR</td>
<td>Accounting</td>
<td>Supplier invoicing</td>
</tr>
<tr>
<td>SB</td>
<td>Accounting</td>
<td>Client invoicing, salaries</td>
</tr>
<tr>
<td>OC</td>
<td>Operations</td>
<td>Project Director</td>
</tr>
<tr>
<td>PD</td>
<td>Logistics</td>
<td>Logistics coordinator</td>
</tr>
<tr>
<td>XZ</td>
<td>Sales</td>
<td>Sales coordinator</td>
</tr>
</tbody>
</table>

Table 4: Interviews realised during step 1

<table>
<thead>
<tr>
<th>Initials</th>
<th>Department</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Presentation meeting during step 2

A project coordinator statement

“The programmers are really expert regarding computer based applications. So, they develop the tools they like without worrying about coherence. Thus we could not impose the development of collaborative systems despite the overwhelming number of meetings!”

A management controller statement

“When they (the programmers) examined the interfaces and the application functions of ERP presented they were systematically pessimistic:”

A programmer statement

“I prefer non proprietary software’s; however interfaces of such applications are ugly!”

Computer department chief

“Administration employees are totally unaware of what they really need, and top managers do not understand ERP implications to decide what should be done. We have already developed several applications which have never been used. That’s out of question to do the same with an ERP.”

Table 6: Most salient statements quoted during step 1

---

Genesys corp. | 1 management engineer  
|              | 1 technology engineer  

Accountancy service | 2 employees  
Finance service | 2 employees  
Computer service | 2 employees  
Customer service | 2 employees  

Table 4: Interviews realised during step 1

Table 5: Presentation meeting during step 2

Table 6: Most salient statements quoted during step 1
BEHAVIORAL ASPECTS IN THE USE OF ERP SYSTEMS: STUDY OF A GLOBAL ORGANIZATION

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0646.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Enterprise resource planning (ERP) (packaged systems), Behavior, Management control, Power</td>
</tr>
</tbody>
</table>
Abstract

Since the 1990’s integrated information systems, better known as Enterprise Resource Planning, really started to take off as a potential solution for the integration of all core business areas in the organizations, allowing more visibility by consolidating all the information in one shared database, in real time. On the other hand, implementing a ERP system causes profound changes in the way employees perform their tasks. Research of the concerning literature suggests that ERP systems increase visibility and control within the company, the better access to information contributes to the empowerment of the employees, in the same way that the procedures of the ERP systems engage employees in self discipline and reflection on work practices. The literature review addresses concepts such as Foucault’s view of disciplinary power, organizational control and empowerment. The study adopted a quantitative methodology through a survey conducted in a multinational organization with 213 respondents, revealed that the visibility offered by the ERP system fosters autonomy making empowerment easier and that the respondents do not feel more controlled by their superiors or peers. In contrast to previous research, which relates reflection to improve working practices, the present study relates reflection with commitment to performing the system’s standardized tasks.

KEY WORDS: Information Technology, ERP Systems (Usage), Control, Empowerment
1. INTRODUCTION

This study seeks to test constructs of the specialized literature and analyze aspects of organizational behavior related to the use of ERP (Enterprise Resource Planning) systems through issues such as empowerment, control, reflection on practice, and conformity with rules. Our primary intent is to consolidate prior knowledge – and add information – on the visibility provided by ERP systems. Our study presumes that multiple, profound organizational changes have arisen in the post-implementation phase of the ERP system life cycle, and presumes that they have a considerable potential to generate employee control and empowerment. Our main points of reference were the concepts of Panoptic Empowerment and Reflective Conformity, proposed by Elmes et al (2005), and Panoptic Control, proposed by Sia et al (2002), in light of Foucault’s concept of disciplinary power and Bentham’s Panopticon.

In recent years, employee empowerment has emerged in parallel with the traditional function of employee control (Drucker, 1988). The post-ERP implementation environment has revealed the considerable impact of these systems on user behavior, both among managers and among their subordinates. This seeming dichotomy has been reinforced by the literature, with authors such as Hanseth et al (2001) claiming that ERP systems are the ideal managerial control technology and Elmes et al (2005) showing that ERP systems are simultaneously able to provide management with ample visibility (and, consequently, greater control) and contribute to employee empowerment by providing access to information. Placing emphasis on the paradoxical view presented in the literature, in which a control technology such as ERP simultaneously facilitates control and fosters empowerment when such control may impose restrictions, and using as a reference the work of Elmes et al (2005), which introduces another, distinct contradictory concept of reflection and conformity, this study seeks to contribute with a greater understanding of these effects on organizational behavior, providing the necessary subsidies for organizations to effectively manage these impacts, with the organizational culture of different countries as its background.

Our study was conducted in a multinational organization in the post-implementation stage of the ERP system life cycle by quantitative field survey. Respondents from eight countries in which ERP had been in use for at least two years were chosen, for a total sample size of 213. The next section provides a theoretical background which represents the most relevant and related studies up to its completion date and presents the proposed research model, hypothesis and the underlying construct. Section three outlines the methodology and the data collection process. The analysis of results is presented in section four and the final section of this paper concludes with the final considerations and the study limitations.

2. THEORETICAL BACKGROUND AND RESEARCH MODEL

In the 1990s, new approaches to management appeared to keep up with the increasingly competitive environment of the corporate world. These new approaches were followed by new technologies, such as ERP systems, considered one of the most important instruments available capable of providing organizations with a competitive edge. ERP tools imposed themselves as a means of maximizing operational efficacy, cutting costs, integrating business processes, and consolidating information into a single database. In addition, there is the ERP system benefit of enabling a transformation from inefficient business processes toward accepted best of practice processes (Gattiker and Goodhue, 2004).

In implementing an ERP system, organizations envision – in addition to intended cost-reduction benefits – an ideal of control over and visibility of their business, which is emphasized by standardized and integrated processes. An integrated system highlights governance capacity, achieved through the integration of data generated by and used in distinct parts of the organization (Hanseth et al, 2001). On the other hand, the implementation of ERP entails profound changes in the manner in which employees carry out their duties, not least because most ERP implementation processes involve
reengineering of business processes, which leads to radical organizational restructuring and changes them, seeking gains from extraction of novel information.

Several studies have sought to investigate individuals’ perceptions on the manner in which work practices have unfolded post-implementation of an ERP system. One such study has found how individual performance can be affected by ERP use, when system users are empowered in high-pressure contexts (sometimes characterized by an overload of work and information) or, conversely, by insufficient provision of relevant information by the system (Elie-Dit-Cosaque et al, 2006).

Some studies have focused on the relation between control and empowerment in the context of ERP system use. An empirical investigation conducted by Psinos et al (2000) suggested that while information systems (including ERP systems) may contribute to empowerment by providing access to information, they can also restrict employees’ freedoms by limiting actions through inflexible processes and increasing performance expectations. Control and empowerment must therefore, be carefully balanced to ensure that the former does not hamper the latter. Hanseth et al (2001) also note that a global organization that implements ERP seeking to strengthen control may in fact obtain the opposite result, that is, less control than it had before.

Our study recognizes the potential of ERP systems to generate control and empowerment, and, particularly, on two prior studies conducted in light of Foucault’s perspective of power and Jeremy Bentham’s *Panopticon*. These concepts are presented in greater detail below:

1) Sia et al (2002) argue that ERP systems create an *information panopticon* through the visibility of information they provide, giving employees decision-making power while simultaneously making them visible to management, thereby leading to employee empowerment; ERP systems thus make both control and empowerment viable. Sia et al call this duality *panoptic control*. Despite finding evidence of its potential in their study, the authors report that management at the studied hospital tended to keep hierarchical structures in place, leading to greater reinforcement of managerial authority than employee empowerment and, consequently, an unbalanced distribution of power with a high level of control and low empowerment. A qualitative methodology was adopted, followed by a quantitative survey whose data collection instrument was validated in a series of interviews.

2) Elmes et al (2005) define empowerment as any increase in employee power, be it through greater formal authority or through maximization of access to useful information, enabling workers – and, consequently, the organization – to achieve institutional objectives with maximum efficiency and efficacy. Like Sia et al (2002), Elmes et al suggest that employees are empowered mainly because they have greater visibility of information, giving them greater control over the factors affecting the manner in which they carry out their functions. From Foucault’s standpoint (1979), they suggest that an organization can empower its workforce without actually transferring power to employees, by considering power to be the “property” of the entire system. The authors call this combined concept of empowerment and multidirectional visibility *panoptic empowerment*, and add to the study the concept of *reflective conformity*. In this concept, they suggest that the rules and procedures of organizational processes in an ERP system lead to greater discipline and induce greater reflection of work and of how the ERP system itself works, making both more effective. The results of the study showed a simultaneous increase in control and empowerment, and also found greater conformity and reflection in business processes. They used an interpretive perspective with “Grounded Theory” methods, using the Glaser (1978) approach, and conducted data collection through observation, formal interviews, and informal conversations in a multinational industrial corporation.

Seeking to answer a series of questions regarding behavioral aspects of the use of ERP systems in global organizations, such as user empowerment, control, reflection on work practices, and conformity with rules – to which this study intends to contribute, be it by consolidating prior knowledge or by obtaining new information – the figure 1 below suggests a structured model built around a few research hypotheses. The research model considers a structure in which the integrated nature of ERP Systems (Davenport, 1998; Al- Mudimigh et al, 2001; Al-Mashari, 2003; Souza and Zwicker, 2003; Volkoff et al, 2005) determines: (a) ample visibility of information, (b) conforming to the rules of ERP
systems embodied in unrestricted way and (c) visibility of employee activities. This causal structure is well discussed in previous literature (Davenport, 1998; Dillard et al., 2005; Elmes et al., 2005; Sia et al., 2002).

The Visibility of Information enhances power of users allowing them to make their job more effectively (empowerment) (O’Leary, 2000; Psinos et al., 2000; Sia et al., 2002; Elmes et al., 2005; Elie-Dit-Cosaque et al., 2006), as well as the integrated business processes of organization embodied in the ERP Systems enable the autonomy improving the group decision making (Saccol et al, 2003). Within the context of a positivist paradigm, we may then formulate and test the following hypotheses: “H1- The Visibility of Information provided by the ERP system fosters empowerment of system user employees” and “H2-The Visibility of Information provided by the ERP system leads to autonomy of user employees”.

The control is enhanced because workers’ performance can be better monitored given the systematic tracking of transaction details and the real time (Sia et al, 2002). Standardized and accessible information makes visible any deviation from the norm (O’Leary, 2000), and as Sia et al. (2002) suggest this knowledge of being visible extends the exercise of self-discipline at the individual level. Thus, we propose the following: “H3- The Visibility of Employee Activities provided by the ERP system makes it easier for the organization to exercise control over user employees”.

Besides the results of Elmes et al’s study (2005) that showed a simultaneous increase in control and empowerment, Saeed et al. (2006) attempted to demonstrate that empowerment and control are intrinsic to each other in the same manner that empowerment supports control on present organizational environment, effective control relies on the empowered employee. So, based on the literature review, we propose the following hypothesis: “H4- Employee control is positively related with empowerment of user employees”.

Dillard et al (2005, p.115) remark that ERP systems are the expression of instrumental rationality revealed in administrative expertise, and “As an instrument of control and domination imposes a mode of organizing and changing social relationships”. Kallinikos (2004, p.10) reinforces this discussion when he observes that “…ERP systems have profound effects on the structuring of work and the forms of human action they enable or constraint.” Eventually, in researches conducted by Saccol et al (2003) in organizations could be noted that the workers with more knowledge on ERP system were more valued by the organization. However, aspects as creativity and innovation in work practices were not used as much they should due to procedures routines. On the other hand, as previously mentioned Elmes et al (2005) suggest that the rules and procedures of organizational processes in an ERP system lead to greater discipline and induce greater reflection of work and of how the ERP system itself works, making them both more effective. Finally, we formulate: “H5- The disciplinary nature of conforming to the rules of ERP systems is positively related with reflection on work practices”.

3. METHODOLOGY

Our study was divided into two phases. The first, exploratory, phase consisted of a review of existing literature on the theme for increasing our understanding of the constructs above. We used several sources, as literature review, interviews, and documents on ERP implementation in the subject organization. Secondly, a quantitative field survey was made with a structured questionnaire as the data collection instrument, seeking to test and verify the inferences and conclusions of the previous authors.
Our subject organization, “Alpha”, is a Belgium-based multinational in the food sector, with 88 offices in 57 countries, 54 production facilities in 40 countries, and with a population of ERP users about 700 employees. The organization adopted the Baan ERP system in 1997, when it had 79 branches. The organization is in the post-implementation phase of the life cycle; the system has been in place at least two years in most branches. It was implemented for two main reasons: software standardization – the company used 19 different software packages at the time – and preparation of 38 branches for the Y2K bug. Analysis of the specific literature shows that most case studies of ERP implementation have focused on SAP systems; few studies of organizations with Baan systems have been conducted.

In light of the difficulties in contacting respondents in all the countries with the implemented Baan system, our criteria for sample definition was to select those countries in which the system was used by 20 or more employees. This narrowed our sample locations to Belgium, Spain, Portugal, Romania, England, Canada, the U.S., and Greece. This type of nonprobability sample, that is, a sample in which the probability of individuals to be chosen for the sample is not known to be equal (Malhotra, 2006) is done at the researcher’s convenience, and is therefore known as a convenience sample. Based on variables found in the literature (Sia et al, 2002; Elmes et al, 2005), we developed a questionnaire to measure the constructs required for hypothesis testing. In order to ensure that items corresponded to theoretical constructs, and, consequently, validate our data collection instrument, we conducted a pilot test (Sprangers and Hoogstraten, 1989) with 14 respondents (managers and non-managerial employees) using the Baan system in a production plant in Athens, Greece.

We used a structured, formal questionnaire with closed-ended questions in pre-established order (Malhotra, 2006). Notably, the questionnaire was self-filling, without respondent’ identification. Marconi and Lakatos (2000) note the advantages of self-filling questionnaires: they allow greater freedom in answering thanks to anonymity and dampen the risk of answer distortion, as there is no researcher influence whatsoever. We used a multiple-item Likert-type scale of one to five, with one (1) being complete disagreement and five (5) being complete agreement. We were concerned that respondents might not fully understand the questions due to the complex nature of the subject and therefore, devised three versions of the questionnaire: one in Portuguese, one in English, and one in Spanish. Questionnaires returned with only one of their two parts filled in were excluded from the sample.

The questionnaire comprised two parts, each of which was saved as a worksheet in an Excel workbook. Part 1 was composed of multiple-choice questions for the following variables: Country, Gender, Age, Department, Organizational Level, Position, Academic Background, and Years of Service. For the second part of the survey, we devised 20 statements based on inferences from Elmes et al (2005) and on the data collection instrument used by Sia et al (2002). These statements were distributed according to certain aspects described in the study hypotheses.

Access to the Baan user population was provided by the company’s Intranet and valid corporate email addresses. Due to security reasons, the company’s IT department would not allow employees access to an electronic form made available in a regular website. Our data collection strategy consisted of making employees aware of the survey through email messages, highlighting the importance of the study and its scholarly nature and purpose. Before mailing these messages, we contacted the company’s general management in each sample country, by telephone, in order to provide clarification on the study’s objectives and foster cooperation to ensure maximum, satisfactory results.

Data collection began on March 30, 2008, and concluded on May 30, 2008. Over the course of these two months, we received 240 responses; 213 of these questionnaires were considered valid (response rate, 89%). The 27 remaining questionnaires (11% of total) were excluded from the sample due to inconsistencies in data or incomplete response (one of the two parts of the questionnaire missing). Response was considered incomplete if less than 90% of the second part of the questionnaire had been filled out.

4. ANALYSIS OF RESULTS
4.1. Data Analysis Procedures and Sample Characteristics

Data analysis was carried out with the Microsoft Excel for Windows spreadsheet software and the SPSS for Windows statistical package. We sought to describe respondents’ demographic characteristics and compare their responses to analyze any possible patterns: high levels of agreement or disagreement with certain statements and grouping of statements related to empowerment, control, self-discipline/conformity and reflection on practice factors. Data were analyzed through univariate and multivariate procedures (Hair et al, 2005). To validate our research model, we used descriptive sample analysis, factor analysis, Pearson correlation analysis, and univariate analysis of variance (ANOVA). We verified three hypotheses through descriptive analysis, by measuring mean agreement with the questionnaire’s statements. We then tested two factor association measurement hypotheses through Pearson’s correlation coefficient, seeking to measure the degree and magnitude of the relationship. ANOVA was used to detect any statistically significant differences between the perceptions of different groups, as defined by respondents’ demographic data and average factor scores. For both ANOVA and correlation analysis, we created a score for each construct to be tested, determined by averaging the construct’s component items.

The sample was distributed as follows: 81 respondents in Belgium, 29 in Spain, 28 in Romania, 26 in the United States, 18 in Greece, 14 in Portugal, and 11 in the United Kingdom. The majority of respondents (54.1%) were male. Respondents were concentrated (38.2%) in the 30–39 age bracket; 50% of respondents were 40 years old or older and 51.4% were college graduates or had a graduate education. As for organizational aspects, 58.7% of employees were non-managers; interestingly, 55.6% of managers who responded were stationed at the company’s headquarters in Belgium. Allocation of respondents was broken down as follows: 27.3% worked in the Sales areas and 28.3% in Operations/Production.

4.2. Analysis of Respondent Perceptions

Descriptive statistics and the number of valid cases (n) were calculated for each variable (see Table below), as well as percentage agreement rates (PAR), calculated by the formula \((\text{mean} - 1)/4 * 100\), showing responsibility, commitment, reflection on practices translating into good work, and interest in obtaining greater knowledge of the ERP system. Low levels of agreement were found with the following statements: 4, with a PAR of 41.86% and a 2.67 average; 9 with a PAR of 44.81% and a 2.79 average; 12, with a PAR of 45.42% and a 2.82 average; and 20, with a PAR of 43.66% and a 2.75 average. The first two statements concern employee control through peer and supervisor visualization. For both ANOVA and correlation analysis, we created a score for each construct to be tested.
4.3. Factor Analysis

Tests showed that factor analysis could be an appropriate method for analysis of correlation matrices, by means of the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO 0.846) and Bartlett’s test of sphericity. Factors are determined by their factor loadings, which must be significant; their values are obtained by rotating the loading matrix with Varimax rotation, a method that seeks to minimize the number of variables with large loadings, simplifying interpretation (Malhotra, 2006). Measurement scales are analyzed through communalities and Cronbach’s alpha coefficient, based on the tenet that, if all items in a measure are drawn from the domain of a single construct, responses to these items should be highly correlated (Churchill, 1999). The lowest acceptable threshold for Cronbach’s alpha coefficient is usually set between 0.6 and 0.7 (Hair et al. 2005). The measure of communalities, which refers to the variance shared by the original variable and all other variables, was included in the analysis.

Validation of the constructs mentioned in prior studies was carried out through factor analysis, in which factor loadings showed the significance of responses, which are able to express the behavioral impacts of the research model as perceived by individuals in the sample, with loadings distributed among the variables of four factors which we found and named Empowerment, Control, Reflection, and Conformity to Rules (see Table 2 bellow). We had total explained variance of approximately 60% (within acceptable limits) and an eigenvalue criteria of >1, as recommended for definition of factors. We considered the values closest to 1 (in modulo), which indicate strong correlation between the statement and the factor it is supposed to represent. We did so also with the Conformity to Rules construct, which did not show satisfactory reliability as measured by Cronbach’s alpha, but it showed high communality (>0.50) and we therefore decided to keep it.

### Table 1

*Questionnaire (Part 2) and Descriptive Statistics*

#### 4.3. Factor Analysis

Tests showed that factor analysis could be an appropriate method for analysis of correlation matrices, by means of the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO 0.846) and Bartlett’s test of sphericity. Factors are determined by their factor loadings, which must be significant; their values are obtained by rotating the loading matrix with Varimax rotation, a method that seeks to minimize the number of variables with large loadings, simplifying interpretation (Malhotra, 2006). Measurement scales are analyzed through communalities and Cronbach’s alpha coefficient, based on the tenet that, if all items in a measure are drawn from the domain of a single construct, responses to these items should be highly correlated (Churchill, 1999). The lowest acceptable threshold for Cronbach’s alpha coefficient is usually set between 0.6 and 0.7 (Hair et al. 2005). The measure of communalities, which refers to the variance shared by the original variable and all other variables, was included in the analysis.

Validation of the constructs mentioned in prior studies was carried out through factor analysis, in which factor loadings showed the significance of responses, which are able to express the behavioral impacts of the research model as perceived by individuals in the sample, with loadings distributed among the variables of four factors which we found and named Empowerment, Control, Reflection, and Conformity to Rules (see Table 2 bellow). We had total explained variance of approximately 60% (within acceptable limits) and an eigenvalue criteria of >1, as recommended for definition of factors. We considered the values closest to 1 (in modulo), which indicate strong correlation between the statement and the factor it is supposed to represent. We did so also with the Conformity to Rules construct, which did not show satisfactory reliability as measured by Cronbach’s alpha, but it showed high communality (>0.50) and we therefore decided to keep it.

### Table 1

<table>
<thead>
<tr>
<th>Measurement Item</th>
<th>Factor Loadings</th>
<th>Communality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1: Empowerment</td>
<td>Alpha de Cronbach: 0.795</td>
<td></td>
</tr>
<tr>
<td>% Variance Explained in Varimax Rotation: 15%</td>
<td>Eigenvalue: 1.904</td>
<td></td>
</tr>
<tr>
<td>1) With the use of the Baan System, I depend less on someone for additional information to execute my job.</td>
<td>0.834</td>
<td>0.644</td>
</tr>
<tr>
<td>2) Through the Baan system, I have access to historical information making my job easier.</td>
<td>0.781</td>
<td>0.634</td>
</tr>
<tr>
<td>3) In my job, the Baan System makes information from other functional areas more visible to me improving my decision making.</td>
<td>0.774</td>
<td>0.589</td>
</tr>
<tr>
<td>4) I have more control over my tasks with the Baan system.</td>
<td>0.774</td>
<td>0.589</td>
</tr>
<tr>
<td>Factor 2: Reflection</td>
<td>Alpha de Cronbach: 0.749</td>
<td></td>
</tr>
<tr>
<td>% Variance Explained in Varimax Rotation: 15%</td>
<td>Eigenvalue: 1.160</td>
<td></td>
</tr>
<tr>
<td>5) The Baan System increases my commitment with others departments.</td>
<td>0.551</td>
<td>0.473</td>
</tr>
<tr>
<td>6) I think that the rules and procedures of the Baan System make me engage in greater self-discipline.</td>
<td>0.551</td>
<td>0.473</td>
</tr>
<tr>
<td>7) By using the Baan System, I understand that my acts have impacts on other departments / other people.</td>
<td>0.575</td>
<td>0.478</td>
</tr>
<tr>
<td>8) If I don’t perform my job well in the Baan System, it is likely that other processes depending on my inputs will be affected.</td>
<td>0.575</td>
<td>0.478</td>
</tr>
<tr>
<td>9) The Baan System increases my reflection on improvements of my work practices.</td>
<td>0.575</td>
<td>0.478</td>
</tr>
<tr>
<td>10) I am interested in acquiring more knowledge concerning the Baan system in order to improve my work practices.</td>
<td>0.575</td>
<td>0.478</td>
</tr>
<tr>
<td>Factor 3: Control</td>
<td>Alpha de Cronbach: 0.806</td>
<td></td>
</tr>
<tr>
<td>% Variance Explained in Varimax Rotation: 15%</td>
<td>Eigenvalue: 1.585</td>
<td></td>
</tr>
<tr>
<td>11) Through the Baan system, it is very easy for my supervisor to find out my mistakes.</td>
<td>0.731</td>
<td>0.671</td>
</tr>
<tr>
<td>12) Through the Baan System, it is very easy for users from other departments to see my job performance.</td>
<td>0.731</td>
<td>0.671</td>
</tr>
<tr>
<td>13) The Baan System increases the control of my supervisor over my job performance.</td>
<td>0.581</td>
<td>0.485</td>
</tr>
<tr>
<td>Factor 4: Conformity to Rules</td>
<td>Alpha de Cronbach: 0.734</td>
<td></td>
</tr>
<tr>
<td>% Variance Explained in Varimax Rotation: 15%</td>
<td>Eigenvalue: 1.003</td>
<td></td>
</tr>
<tr>
<td>14) I have to perform my job on the most efficient way, besides the Baan system I use other parallel (or alternative) procedures</td>
<td>0.734</td>
<td>0.485</td>
</tr>
</tbody>
</table>

### 4.4. Univariate Analysis of Variance (ANOVA)

Seeking a more in-depth analysis of the factors that mirror impact on behavior in the organizational environment, we used sample segmentation, analyzing possible relationships between perceptions and...
the demographic variables that make up the first part of the questionnaire. We employed one-way univariate analysis of variance (ANOVA), a statistical test used to compare means in two or more populations; in this particular case, it was used to detect whether there were any statistically significant differences between the perceptions of different populations. The null hypothesis is that samples arise from populations with equal means for a dependent variable (Hair et al, 2005).

ANOVA uses the F-test to prove significant variance between mean group scores. In this case, groups may be countries, age brackets, or departments within the organization. However, according to Glass and Hopkins (1996, p. 405 apud Vidal, Zwicker & Souza, 2003), when used to compare groups with different variances, ANOVA may not produce accurate results. In such cases, one may use the Brown-Forsythe test, which does not presume homogeneity of variance and corrects the results of the F-test ($F^*$). Although univariate ANOVA allows us to reject the null hypothesis that group means are equal, it does not pinpoint the location of significant differences. In-depth investigation of a specific group means that may be of interest requires post hoc tests such as Tukey's honestly significant difference (HSD) (Hair et al, 2005).

ANOVA revealed no significant differences between countries in the Empowerment and Reflection factors, whose mean values ranged from 3 to 4. There were significant differences in the Control factor between Belgium and Romania, between Romania and Greece, and between Romania and the UK (see Table). Belgium data (81 respondents, 2.51 mean) suggests significant disagreement regarding control issues addressed by the visibility provided by ERP systems, while Greece and Romania show higher agreement (means 3.26 and 3.52 respectively) and a smaller number of respondents. Likewise, Romanian respondents (mean agreement, 3.52) disagreed markedly with those in the UK (mean, 2.36).

Our analysis of Organizational Level groups found no statistical differences between them in the Empowerment, Reflection, and Conformity to Rules factors, while there was a significant difference in the Control factor, as shown by a striking difference in mean agreement: the mean agreement (or, rather, disagreement) of managers was 2.44, while that of respondents in non-managerial positions (the majority of the sample) reached 3.12. There was a total of 125 respondents in the sample. Post-hoc testing with Tukey’s honestly significant difference could not be performed, as there were less than three groups to compare.

Among Level of Education groups, ANOVA found no statisticant difference in the Empowerment, Reflection, and Conformity to Rules factors. Significant differences were found in the Control factor between respondents with secondary-level education and those holding master’s degrees, and between those with secondary-level education and those holding doctorates (see Table). The mean agreement of respondents with a secondary-level education (3.07) – respondents who usually hold non-managerial positions and make up the majority of the sample – corroborates the notion that respondents without a graduate education feel that ERP increases the control their superiors hold over them. Respondents with master’s or doctorate degrees showed significantly lower mean agreement with this statement (2.55 and 1.44 respectively), meaning they do not feel controlled, watched, or monitored; this leads us to infer that graduate-level respondents hold managerial positions.

<table>
<thead>
<tr>
<th>Sites</th>
<th>ANOVA</th>
<th>Brown-Forsythe</th>
<th>Sites vs Rom</th>
<th>Tukey HSD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F$</td>
<td>Sig</td>
<td>Mean Diff</td>
<td>Sig</td>
</tr>
<tr>
<td></td>
<td>$F^*$</td>
<td>Sig</td>
<td>Mean Diff</td>
<td>Sig</td>
</tr>
<tr>
<td>Control</td>
<td>5.03</td>
<td>0.00</td>
<td>4.26</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>1.0716 vs 1.0716</td>
<td>0.00*</td>
<td>-0.7530 vs 0.7530</td>
<td>0.03*</td>
</tr>
<tr>
<td></td>
<td>1.160 vs -1.160</td>
<td>0.01*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>ANOVA</td>
<td>Brown-Forsythe</td>
<td>Sec vs Master</td>
<td>Tukey HSD</td>
</tr>
<tr>
<td></td>
<td>$F$</td>
<td>Sig</td>
<td>Sec vs Master</td>
<td>Sec vs Doct.</td>
</tr>
<tr>
<td></td>
<td>$F^*$</td>
<td>Sig</td>
<td>Mean</td>
<td>Sig</td>
</tr>
<tr>
<td>Control</td>
<td>4.50</td>
<td>0.00</td>
<td>5.94</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>0.53 vs -0.53</td>
<td>0.019*</td>
<td>-1.63 vs -1.63</td>
<td>0.05*</td>
</tr>
</tbody>
</table>

* denotes significant difference at a sig. Level of 5 %

Table 3 Factor ANOVA, sample broken down by country and Level of Education
ANOVA of the Control factor again showed significant inter-group differences when respondents were divided into groups according to the department in which they worked, while, once more, there were no significant differences regarding agreement with the Empowerment, Reflection, and Conformity to Rules factors (see Tables below). Disagreement with the Control statements was most pronounced among employees of the company’s Lab and R&D Department (2.07), compared to the Financial (3.3), Sales (5.7), Accounting (1.5), and Operations/Production (6.2) departments, which showed mean agreement values of 3.18, 2.98, 3.58, and 2.81 respectively. This may be explained by the fact that use of the ERP system is more widespread in these departments, which may lead to increased recognition of the activity and performance visibility afforded by ERP among the employees of these departments. The difference in agreement may also be explained by the nature of each department; the Marketing department (mean agreement, 2.38, 8 respondents), for instance, differs markedly from the Accounting department (mean, 3.58, 15 respondents), whose activities feature a high level of monitoring and control. The difference between Accounting and IT (mean agreement, 2.50, 10 respondents) is likely because the IT department does not use the ERP system and therefore does not recognize the monitoring and vigilance imposed by it.

4.5. Hypothesis Testing

After the reliability of scales has been assessed, the next stage is hypothesis testing. We will test hypotheses H1, H2, and H3 by using the mean scores of significantly loaded explanatory variables, and in hypotheses H4 and H5, we will analyze the intensity of association between factor variables with Pearson’s correlation coefficient. To carry out this test, we created a score for each construct from the mean of its component items.

Hypotheses one and two, “H1 – The Visibility of Information provided by the ERP system fosters empowerment of system user employees” and “H2 – The Visibility of Information provided by the ERP system leads to autonomy of user employees”, concerning the visibility of information provided by the ERP system and its role in facilitating empowerment and fostering autonomy, were proven by verifying mean respondent agreement with questionnaire statements that explain the Empowerment factor. Respondents agreed somewhat (average agreement 3.6) with the assertion that ERP provides visibility of information, thereby fostering autonomy and facilitating a decision making. With access to a greater volume of information comes increased employee power, as employee decisions are now supported by a more secure and reliable information base. According to an interviewee in the company’s Greece branch, adoption of the ERP system was followed by delegation of greater responsibility and autonomy to operational areas, which then showed greater and more significant changes. Prior to ERP implementation, all transactions were performed by employees in the Financial department; according to the interviewee, the ERP system encouraged the (logical) conclusion that each department should carry out its own transactions. ANOVA revealed significant differences in the average agreement with factors related to these hypotheses between age brackets: respondents aged 50–59 or >60 expressed greater perceived recognition of the Empowerment factor than did employees in the 30–39 age range.

The third hypothesis, “H3 – The Visibility of Employee Activities provided by the ERP system makes it easier for the organization to exercise control over user employees”, verifies whether the visibility of actions provided by the ERP system makes it easier for the organization to exercise control over employees who use the system. Respondents disagreed slightly (average agreement, 2.7) with the assertions that express the visibility of actions provided by the ERP system to employees’ peers and superiors. We were unable to confirm the effects of this concept as presented in the studies of authors mentioned previously in this article. Agreement with the two component variables of the Control factor with greater explanatory power (0.812 and 0.859) ranged from neutrality to disagreement, (41.86% and 44.81%, averages 2.67 and 2.79 respectively), that is, respondents did not express recognition of increased control over their work performance, whether exercised by their superiors or by their peers. Respondents were neutral to the assertion that ERP makes it easier for superiors to spot errors (mean agreement, 3.03). From these results, we may infer that, although ERP systems make employees’ actions more visible to their superiors and peers, employees did not feel more controlled,
monitored, or watched, as Sia et al (2002)’s Panoptic Control proposed they would. H3 is therefore not supported.

ANOVA revealed significant differences in the mean agreement of respondents with the Control factor. Managers and respondents holding master’s and doctoral degrees disagreed more with the Control statements than did respondents with a secondary-level education; respondents working in Belgium disagreed more than did respondents in Romania and Greece, who showed the highest levels of agreement with the Control assertions; likewise, respondents in the UK disagreed more with these assertions than did those in Romania. Respondents in the Lab/R&D department disagreed most with the Control factor, as in comparison to those in the Financial, Sales, Accounting, and Operations/Production departments. These figures may be explained by the fact that ERP use in the Lab/R&D department is less prevalent than in other departments, which suggests that R&D employees underrecognize the visibility of actions and performance provided by the ERP system.

Although there was no express recognition that the increased visibility of actions provided by the ERP system facilitates control, our fourth hypothesis, “H4 – Employee control is positively related with empowerment of user employees”, verifies whether employee control is positively correlated with empowerment of user employees. Our finding of a positive correlation between control and empowerment corroborates the Panoptic Empowerment concept defined by Elmes et al (2005), that is, the association of these two constructs arises with the visibility of information provided by the ERP system, which leads to greater autonomy and decision-making power for employees in operational positions, consequently giving them a greater sense of responsibility and fostering a concern with their actions within the ERP system. Pearson’s coefficient was found to be 0.374 at a significance level of 0.01.

Our final hypothesis, “H5 – The disciplinary nature of conforming to the rules of ERP systems is positively related with reflection on work practices” tests whether the disciplinary nature of conforming with ERP rules is positively correlated with reflection on work practices. The concept of Reflective Conformity identified by Elmes et al (2005) emerged in our hypothesis testing, revealing a correlation between Reflection and Conformity to Rules – however, the correlation was negative. H5 is thus rejected. This may be explained by the fact that the Conformity to Rules construct sought to measure conformity that does not derive from hierarchical power, but rather from the inherent disciplinary power of ERP system procedures. Conformity was measured by two variables: one referred to the use of parallel resources in addition to the ERP system in order to perform tasks more efficiently, that is, non-conformity, and the other stressed conformity as performance of activities in accordance with the system’s demands, without questioning its procedures. Our analysis revealed that respondents do not conform to the rules. Agreement with the statement that referred to the use of parallel resources was above average (PAR, 63.85%; mean agreement, 3.55), while agreement with “not questioning the system’s procedures” ranged from neutrality to disagreement (mean agreement, 2.75; PAR, 43.66%). This shows a latent concern with doing a good job, even if it means not using the ERP system.

Regarding Reflection, Elmes et al (2005) captured this concept as being allied with conformity to the inherent rules and procedures of the ERP system, that is, the limits imposed by ERP systems are explained by Elmes et al (2005) from the standpoint of Foucault’s “regime of truth” – the power of discourse that values disciplinary action. According to Elmes et al (2005), increased reflection on practices arises from the search for solutions to problems related to the use of ERP systems, or rather, the restrictions imposed by ERP systems would trigger reflection on how to improve one’s practices at work. This Reflection construct featured the highest levels of agreement, as shown by the following variables: recognition that the respondent’s actions may have an impact on other areas of the company had the highest PAR overall (76.17%) and a mean agreement of 4.05; user commitment to performing well within the system, PAR 74.41% and mean agreement 3.98; and interest in learning more about the ERP system in order to perfect work practices, PAR 72.65% and mean agreement 3.91. When asked whether ERP increased their reflection on how to improve work practices. However, respondents were neutral (PAR, 51.53%; mean agreement, 3.06).
This result showed that ERP system users are greatly concerned with their responsibilities and their commitment, and their Reflection on work practices is translated into a concern with doing a good job; however, their vision strayed from that proposed by Elmes et al (2005), according to which the ERP system would foster reflection on improving one’s work practices in addition to promoting conformity to the system’s rules. According to our results, it would be more realistic to characterize the Reflection construct as fostering increased user responsibility, so that users may correctly perform activities that were made highly standardized by the implementation of ERP.

5. FINAL CONSIDERATIONS AND LIMITATIONS

Interpretation of the results of our field survey led to the following conclusions:

- Although ERP provides greater visibility of employees’ actions, respondents did not report feeling more tightly controlled or monitored by their superiors or peers. This contrasts with the results reported by Sia et al (2002), who recorded the Panoptic Control mechanism, and Elmes et al (2005), who claimed that employees know they are being watched and that access recorded by the ERP system correlates with disciplinary power, in accordance with Foucault’s perspective. We must note that ANOVA showed significant differences in the average perceptions of manager-level respondents, with greater disagreement regarding the Control factor; this finding is coherent with the fact that employees in managerial positions do not feel the effects of vertically exercised disciplinary power.

- As for Panoptic Empowerment, verified by Elmes et al (2005) but not corroborated by Sia et al (2002), our results showed that the real-time, multidirectional visibility of information provided by the ERP system within Foucault’s perspective fosters autonomy and increased decision-making power. This helps empowerment, regardless of the manner in which hierarchical power is transferred. None the less, the concept of Panoptic Empowerment – which Elmes et al (2005) stress is a contradiction, as it productively correlates high levels of control with high levels of empowerment – was not corroborated by high levels of control as reported by our respondents.

- The Reflective Conformity concept identified by Elmes et al (2005) correlates Conformity to Rules – brought about not by hierarchical control, but rather by the inherent disciplinary power of ERP procedures – with increased Reflection on work practices; that is, the limits imposed by these systems are explained by Elmes et al (2005) in light of Foucault’s “regime of truth” – the power of discourse that values disciplinary action. According to Elmes et al (2005), increased reflection stems from the search to solve problems related to use of the ERP system, or rather, the restrictions of ERP systems would lead employees to reflect on their work practices. Our results, however, showed a negative correlation. The present study has shown that users of an ERP system show great concern for their responsibilities and commitment, and their Reflection on their practices translates into a concern for doing a good job. Our view strays somewhat from that of Elmes et al (2005), however, according to which ERP systems would promote reflection on work practices in addition to obedience to the rules of the system. A more appropriate characterization of this construct would be that reflection fosters increased user responsibility, to ensure that users carry out their activities (which become highly standardized upon implementation of ERP) correctly.

Some limitations should be considered in this study. First, our results do not permit a generalization of findings for the user population of systems ERP, getting restricted to the sample. The second limitation is concerning the problems of scale of Conformity to Rules. The reliability of Cronbach’s alpha coefficient indicates that subsequent analysis should be seen with some reservations. The results of this factor suggest the need for improvement in the scale, probably through a search of the theoretical framework for new items that would assist in adequately explaining its meaning. The impossibility of identifying a relation between cause and effect by correlation analysis is the third limitation. However, as the literature points to the existence of such variables, the final analysis considers this relation as proposed in the literature. Future research can go deeper into behavioral...
aspects, especially the control factor, taking into account the findings identified here as well as cultural dimensions of the countries.

References


ANTECEDENTS AND DRIVERS OF IT-BUSINESS STRATEGIC ALIGNMENT: EMPIRICAL VALIDATION OF A THEORETICAL MODEL

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0048.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Strategic Alignment, IT Strategy, IT/IS management, IT Business Alignment / Value</td>
</tr>
</tbody>
</table>
ANTECEDENTS AND DRIVERS OF IT-BUSINESS STRATEGIC ALIGNMENT: EMPIRICAL VALIDATION OF A THEORETICAL MODEL

Yayla, Ali, Binghamton University, PO Box 6000, Binghamton, NY 13902, USA, ayayla@binghamton.edu

Hu, Qing, Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431, USA, qhu@fau.edu

Abstract

Aligning information technology (IT) strategy with business strategy has been one of the top concerns of practitioners and scholars for decades. Despite the documented positive effects of strategic alignment on organizational success, only a few organizations consider themselves in alignment. Although numerous studies exist about how to accomplish IT-business alignment, empirical studies based on strong theories have been rare in the literature. This study attempts to fulfill this gap by proposing and empirically validating a comprehensive strategic alignment model. Drawing on prior literature, we identified five antecedents of alignment: centralization, formalization, shared domain knowledge, successful IT history and relationship management. We further hypothesized that the effects of these antecedents are mediated by two drivers of alignment, which are conceptualized as the level of connection of IT and business planning and the level of communication between IT and business managers. Using survey data and structural equation modeling methodology, we show that both drivers had significant effects on alignment, and the effect of connection is about twice that of communications. Our findings also confirm the effects of all antecedents except centralization. Overall, the main contribution of this study is the development and empirical validation of a comprehensive strategic alignment model, providing a more ample prescriptive insight for managing IT-business strategic alignment.

Keywords: IT-business alignment, strategic alignment, antecedents of alignment, drivers of alignment
1 INTRODUCTION

Aligning information technology (IT) strategy with business strategy enables an organization to maximize the impact of its IT investments, achieve harmony between IT and business (Byrd et al. 2006, Henderson & Venkatraman 1993, Nelson & Cooprider 1996, Tan & Gallupe 2006), and increase the organization’s competitive advantage, profit margins, and growth (Alter 2005, Byrd et al. 2006, Pearlman & Baker 2005). During the past two decades, executives have consistently identified achieving alignment as one of the top key issues in IS management (Brancheau & Wetherbe 1987, Luftman et al. 2006, Niederman et al. 1991, Rosa 1998). Alignment between business and IT has been an enduring subject for IS scholars as well (Chan et al. 1997, Henderson & Venkatraman 1993, Reich & Benbasat 1996, Sabherwal & Chan 2001). In fact, IT-business alignment is one of the early research streams in the IS literature (King 1978, McLean & Soden 1977), yet it is gaining even more attention in recent years as organizations rely more on information technology for competitiveness and survival in the global economy (Chan et al. 2006, Hu & Huang 2006, Luftman et al. 2006).

However, given the importance and potential benefits of alignment, the number of organizations that successfully align their IT strategy with business strategy is shown to be considerably low. A study by Rosa (1998) has revealed that only eight percent of IT and business managers viewed themselves very effective in aligning IS and corporate goals. Similarly, Luftman et al. (1999) noted that only half of more than one thousand executives that participated in their survey reported that their organizations have achieved some degree of alignment. Naturally, questions like “why haven’t we mastered alignment?” have been raised by scholars (Chan 2002).

One of the strategies to improve alignment suggested in the literature is to identify what managerial practices improve the probability of alignment (Chan 2002). Finding these factors has captured the attention of many scholars. These scholars conceptually agreed upon a number of antecedents, yet empirical validation and support for the arguments raised in these studies are lacking, and the few quantitative studies have mixed results (Chan et al. 2006, Reich & Benbasat 2000, Sabherwal & Kirs 1994). Thus, our knowledge regarding the process that leads to alignment is largely limited at conceptual and intuitive level. There is a clear need for both theory and practice for a comprehensive empirical study based on the recent advances in the alignment knowledge and statistical techniques. This study attempts to fulfill this need. The main objective of this study is to integrate and extend the literature by investigating and statistically testing a theoretical IT-business strategic alignment model.

2 LITERATURE REVIEW

Achieving IT-business strategic alignment within the organization has been considered as one of the key issues in IS management for the past decades (Brancheau & Wetherbe 1987, Henderson & Venkatraman 1993, King 1978, Niederman et al. 1991). Aligning IT and business strategies enables organizations to use their IT resources to support their business strategy, thus leading to higher levels of organizational success. Several studies in the IS literature have focused on this performance effect of IT-business strategic alignment. This body of literature has argued conceptually and found limited empirical support for the enhancing effect of alignment on organizational performance (Chan et al. 1997, Chan et al. 2006, Henderson & Venkatraman 1993, Kearns & Sabherwal 2007, Palmer & Marcus 2000, Sabherwal & Chan 2001, Sabherwal & Kirs 1994, Teo & King 1999).

Another stream of research has investigated the antecedents of IT-business alignment to understand the alignment process, and also to provide prescriptive guidance on how to achieve alignment. Recent literature has in general found that the most important antecedents of strategic alignment are shared domain knowledge of business and IT managers, previous success of IT units, connections of business and IT planning, and communications between IT and business executives (Chan et al. 2006, Hu & Huang 2006, Reich & Benbasat 1996, Reich & Benbasat 2000, Sabherwal & Kirs 1994). Although
studies have conceptually agreed on these antecedents, they are largely at conceptual or qualitative level. The number of empirical studies on strategic alignment is limited and the findings across the studies are inconsistent and have three main shortcomings. First, the data for most of the empirical studies were collected in the mid 1990s (e.g., Chan et al. 1997, Chan et al. 2006, Sabherwal & Kirs, 1994). The role of IT in today’s organizations is considerably different, and therefore, their values in providing insights to the strategic IT-business alignment process in today’s business environment are limited. Second, only a few studies have utilized the benefits of sophisticated statistical methodologies, such as structural equation modeling (SEM), to test the proposed research models. The extant empirical studies mostly utilized multiple regression, ANOVA, and t-tests, which could severely limit their ability to uncover complex interactive relationships among the key alignment constructs. Finally, no empirical studies as we know of have tested the comprehensive models that integrate alignment constructs proposed in multiple studies.

In summary, the current state of the strategic alignment literature calls for a comprehensive empirical study based on the recent advances in the alignment theory and statistical techniques. In the next section we present our research model and develop our hypotheses based on the extant literature.

3 RESEARCH MODEL AND DEVELOPMENT OF HYPOTHESES

After integrating the studies in the literature, we included IT unit structure, shared domain knowledge, successful IT history and relationship management into the model as the most significant antecedents of alignment. We defined strategic alignment as the fit between IS strategy and business strategy of organizations. Thus, when organizations achieve high degrees of strategic alignment, their IS strategy (e.g., IT for efficiency) would support their business strategy (e.g., operating efficiency). Parallel to Reich and Benbasat (2000), we hypothesized that the immediate drivers of strategic alignment are the current business practices and they mediate the effects of antecedents of alignment (Figure 1).

![Figure 1. Proposed research model for IT-business strategic alignment.](image-url)

**Connection:** Connection captures the level of integration of IS planning and business planning. When IT units do not refer to business objectives, their contribution to the organization can be limited (Pearlman & Baker 2005). Higher levels of integration provide an effective way for chief information officers (CIO) to identify top management objectives (Lederer & Mendelow 1987) and leads to better understanding of the role of the IS function, more contribution of IS to the organization, less output and process related problems, and consequently to a better organizational performance (Pearlman & Baker 2005, Teo & King 1996, Teo & King 1999). Parallel to that, Teo and King (1997) reported that managers choose to increase the level of planning integration when their organizations need to use IS to support business strategy and to align IS objectives with business objectives, as well as when they realize the importance of IT for their organization.

Although there is qualitative support for the positive effect of planning integration on strategic alignment (Hu & Huang 2006, Reich & Benbasat 2000), the empirical support is somewhat mixed with studies reporting insignificant (Chan et al. 2006, Sabherwal & Kirs 1994) as well as significant (Newkirk & Lederer 2006a, 2006b) effect of planning integration on strategic alignment. Overall, literature seems to suggest that successful strategic planning leads to alignment of business and IS strategies, better analyses of internal operations, more cooperation between organizational groups and
IS group, improvement in capabilities of IS planning process, and achieving competitive advantage (Earl 1993, Segars & Grover 1998). Building on these theoretical arguments, we posit that;

Hypothesis 1: The level of connection between IT and business planning processes will positively influence the level of strategic alignment.

Communication: Communication process involves sharing and exchanging information between parties for the purpose of coordination and mutual understanding (Bacharach & Aiken 1977, Johnson & Lederer 2005). When business executives meet with IS managers frequently and discuss where the business is heading or how to resolve some of the emerging issues, it is more likely for them to converge on how IT can help the organization to achieve its goals (Johnson & Lederer 2005, Pearlman & Baker 2005). This convergence leads to mutual understanding of the organization’s business and IT functions and the strategic role of IT in the organization, resulting in collective actions to use IT for competitive advantage (Johnson & Lederer 2005). Overall, communication between IT and business executives is considered as an important enabler for strategic alignment (Alter 2005, Earl & Feeny 1994, Hu & Huang 2006, Johnson & Lederer 2005, Reich & Benbasat 2000, Tan & Gallupe, 2006). Thus, we posit that;

Hypothesis 2: The level of communication between IT and business managers will positively influence the level of strategic alignment.

IT Unit Structure: The structure of organizations has a significant influence on the information flow and human interactions through channeling collaboration, specifying modes of coordination, and prescribing levels of formality (Miller 1987). Considering its potential effect on alignment, the IT unit structure aims to capture the degree of centralization of IT decision making and formalization of IT activities.

Centralization: Centralization represents “the degree to which the right to make decisions and evaluate activities is concentrated” (Fredrickson 1986, p.282). In centralized governance modes, IT activities are coordinated at the corporate level (Sambamurthy & Zmud 1999), and therefore, organizations may require less effort for alignment (Chan et al. 2006). In contrast, in decentralized modes, divisional managers assume authority (Sambamurthy & Zmud, 1999), and they are more likely to focus on their own objectives rather than the central objective. When organizations aim to leverage their IT for a more strategic role or when they are dissatisfied with the level of alignment, one of the first tasks executives undertake is to centralize the decision making (Brown & Magill 1994, Rothfeder 2005).

Centralization has an influence on the communication levels as well. Although centralized mode does not foster interactive and participative decision making across hierarchical levels in organizations (Ranganathan & Sethi 2002), this does not necessarily affect the communication level within the top management. Hage et al. (1971) argued that power and status between job occupants are important inhibitors of communication. As the social status increases in organizations, free flow of information decreases (Barnard 1964). However, since in centralized structure there is less power and status discrepancies among decision makers (e.g., top management team, all C-level executives, etc.), centralization enables better communication among decision makers. Overall, the literature suggests that centralized IT governance is indeed an enabler of IT-business strategic alignment. However, we posit that the positive effects of centralization are mediated by the drivers of alignment. Thus,

Hypothesis 3.1a: The level of IS centralization will positively influence the level of connection between IT and business planning.

Hypothesis 3.1b: The level of IS centralization will positively influence the level of communication between IT and business managers.

Formalization: Formalization indicates “the extent to which an organization uses rules and procedures to prescribe behavior” (Fredrickson 1986, p.283). It facilitates planning process (Miller 1987) and use of formal planning procedures in organizations (Pyburn 1991). Especially in complex organizations, formalization enables the clarification of business objectives, thus increase the effectiveness of IS
planning process in terms of reaching to a consensus on the role of IT (Earl 1993, Lederer & Sethi 1988, Pyburn 1991). Formalized IS planning also ensures that overall IS goals are consistent with business goals of the organization (Lederer & Mendelow 1986).

Top management views identifying corporate strategic direction and business plans as one of the most difficult aspects of strategic IS planning (Lederer & Mendelow 1986). Lack of formal organizational strategic plans creates difficulties for CIOs to identify business objectives as well (Lederer & Mendelow 1987). Therefore, the absence of formal business planning could severely damage the IS planning process (Lederer & Mendelow 1989, Lederer & Sethi 1988, McLean & Soden 1977). Overall, formalization promotes better task coordination through frequent communication and integration of planning, and in turn, increases the quality of strategic IS planning and enhances the IT management and decision making process (Bai & Lee 2003, Ranganathan & Sethi 2002). Thus, we posit:

**Hypothesis 3.2a:** The level of IS formalization will positively influence the level of connection between IT and business planning.

**Hypothesis 3.2b:** The level of IS formalization will positively influence the level of communication between IT and business managers.

**Shared Domain Knowledge:** Shared domain knowledge captures the IT knowledge of business managers and the business knowledge of IT managers (Ranganathan & Sethi 2002). The lack of shared domain knowledge is considered as an inhibitor of communication and strategic IS planning (Lederer & Mendelow 1987, Feeny et al. 1992) since it increases the information asymmetry and results in inaccurate interpretation of messages, ultimately leading to intergroup conflict (Nelson & Cooprider 1996). CIO’s business knowledge enhances formal and informal interactions of CIO with top management and increases the assimilation of IT in organizations (Armstrong & Sambamurthy 1999). Moreover, business competency of IS managers has a significant effect on determining the extent of IT-business planning integration (Teo & King 1997). Kears and Sabherwal (2007) also argued that when top managers possess knowledge of IT, the opportunities are created for business and IT managers to participate in each other’s planning process.

Shared domain knowledge is considered as an important component of strategic alignment (Reich & Benbasat 1996, Tan & Gallupe 2006, Teo & Ang 1999). In their respective case studies, both Reich and Benbasat (2000) and Hu and Huang (2006) reported the positive influence of shared domain knowledge on the communications between IT and business executives and connections between IT and business plans. Overall, the literature converges on the positive effects of shared domain knowledge on IT-business alignment through enhancing current business practices. Thus, we posit:

**Hypothesis 4a:** The level of shared domain knowledge of IS and business managers will positively influence the level of connection between IT and business planning.

**Hypothesis 4b:** The level of shared domain knowledge of IS and business managers will positively influence their level of communication.

**Successful IT History:** Successful history of an IT unit gives credibility to the IT unit and creates favorable perceptions of IT in top management (Chan et al. 2006, Earl & Feeny 1994, Hu & Huang 2006, Reich & Benbasat 2000, Rockart et al. 1996). Also, it is an important determinant of the participation of business managers in the planning process (Pearlman & Baker 2005, Teo & Ang 1999). The confidence of top management in the IT department and IT department’s efficient and reliable services are found to be important critical success factors for aligning IS plans with business plans (Teo & Ang 1999, Luftman et al. 1999). Teo and Ang (1999) postulated that confidence of top management in IT increases their commitment to the strategic use of IT, making them more likely to allocate appropriate resource for planning and development of IT applications. On the other hand, lack of IS management credibility discourages top executives from communicating their needs and problems (Pearlman & Baker 2005), and more importantly, from communicating their goals, objectives and plans (Lederer & Mendelow 1989), thus inhibiting strategic alignment (Luftman &
Brier 1999). Clearly, successful IT history itself won’t impact IT-business alignment directly. However, the increased confidence of top management and the higher credibility of the IT unit are likely to enable IT unit to participate effectively in the strategic planning process and communicate effectively with business managers. Based on these arguments, we posit that:

**Hypothesis 5a:** The level of successful IT history will positively influence the level of connection between IT and business planning.

**Hypothesis 5b:** The level of successful IT history will positively influence the level of communication between IT and business managers.

**Relationship Management:** Relationship management captures the extent to which IS and business managers invest time and effort in managing relationship between each other. A close relationship between IS and business managers enables them to work together to understand business and technological requirements (Jones et al. 1995, Rockart et al. 1996, Watson 1990). Ongoing personal relationships facilitate the parties to engage in knowledge creation and their availability for knowledge exchange (Hatzakis et al. 2005, Nahapiet & Ghoshal 1998). Having a good relationship between CIO and CEO is considered as an enabler of IS-business planning integration (Feeny et al. 1992). Bai and Lee (2003) noted that the CEO/CIO relationship can be crucial for the alignment and the quality of strategic IS planning. This relationship, in turn, enhances the ability of IT to add value to the organization (Earl & Feeny 1994) and ensures the successful integration of business and IT strategies (Rockart et al. 1996).

Moreover, several scholars suggested that stronger relationship between business and IT managers would improve their communication level (Coughlan et al. 2005, Hu & Huang 2006). Rockart et al. (1996) noted that only through the established relationship between IT personnel and line managers can the necessary communication occur. The informal relationship networks are also an important component of alignment (Chan 2002). The development of relationships makes the alignment concept more tangible through the understanding of existing communication channels and networked relations in organizations (Coughlan et al. 2005). Overall, IT/business relationship is considered as an important enabler of strategic alignment (Luftman & Brier 1999, Tan & Gallupe 2006). Thus, we posit that the active relationship management by IT and business managers enhances the connections between IT and business planning, and improves their communications, or;

**Hypothesis 6a:** The level of relationship management will positively influence the level of connection between IT and business planning.

**Hypothesis 6b:** The level of relationship management will positively influence the level of communication between IT and business managers.

### 4 RESEARCH DESIGN

Based on the theoretical alignment model and a thorough review of the literature, we developed a survey instrument to collect data for validating the constructs in the model and testing the research hypotheses (Table 1). A seven-point Likert scale was used for all the measurement items in the survey.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Sample Questionnaire Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategic</td>
<td>The IT plan contains detailed action plans/strategies that support organization’s business</td>
</tr>
<tr>
<td>Alignment</td>
<td>objectives and strategies.</td>
</tr>
<tr>
<td>Connection</td>
<td>Business and IT executives both participate in the strategic planning process.</td>
</tr>
<tr>
<td>Communication</td>
<td>Business and IT executives have frequent, direct, formal communications.</td>
</tr>
<tr>
<td>Centralization</td>
<td>Issues related to major IT investments have to be referred to business executives at the top</td>
</tr>
<tr>
<td></td>
<td>management.</td>
</tr>
<tr>
<td>Formalization</td>
<td>Operating rules/procedures play important roles in how issues are handled in the IT unit.</td>
</tr>
</tbody>
</table>
IT executives have a good understanding of the organization’s business environment.

The IT unit has been considered as credible.

Business and IT executives make an effort to maintain a better relationship with each other.

Table 1. Constructs and Sample Measurement Items Used in the Survey

Following a pilot test with EMBA students enrolled in a large public university to validate the survey instrument, we collected the data from organizations in Turkey primarily because of one of the authors’ connection to the Turkish industry organizations. The survey instrument was translated into Turkish and then back-translated for discrepancies with the original instrument in English. Considering that constructs such as strategy, technology and structural characteristics are considered as having low cultural dependency and are invariant in different cultural setting (Samiee & Athanassiou 1998), with respect to this study, we believe that potential threats to conceptual equivalence was minimal.

An invitation for participation to the survey website was distributed through e-mail to 440 organizations in the directories provided by Istanbul Chamber of Industry. Also, one of the authors personally contacted the executives and business managers of 120 organizations. A total of 560 invitations were distributed. The target population included business administrators with complete knowledge of IT and business strategizing processes. At the end of the process, there were 169 usable questionnaires. The average revenue of the firms in the sample was about $1.3 billion. Manufacturing firms represents one third of the sample. This is followed by Wholesale/Retail firms (20%) and Technology/Telecommunication (10%) firms.

We used partial least square (PLS) analysis to analyze the survey data and test the research hypotheses. After verifying that the missing data were not systematic, multiple regression method was used for data imputation. Item loadings and average variance extracted (AVE) values were examined to assess convergent validity. Item loadings range between 0.68-0.91 (p < 0.01), and AVE scores range between 0.61 and 0.78. These results showed that the constructs demonstrate convergent validity. We assessed discriminant validity by examining the AVE scores and the cross-loadings of the items. The loadings demonstrated that the items loaded higher for their corresponding constructs than for other constructs, and the items loaded higher for their corresponding construct than other items, thus providing adequate statistical support for discriminant validity. The Cronbach alpha values of the constructs range between 0.79 and .93, suggesting adequate level of construct reliability (Table 2). Lastly, we utilized Harman’s one-factor test and common method latent variable test to measure the common method bias, and found that the effect of this bias is insignificant.

Table 2. Descriptive Statistics and Psychometric Characteristics of the Constructs

Notes: (1) ** p< 0.01; (2) Diagonal elements are square root of AVE; (3) α: Cronbach’s alpha
5 HYPOTHESIS TESTING AND RESULTS

SmartPLS software (Ringle et al. 2005) was used to conduct PLS analysis (Figure 2). Initially, organizational size was included in the structural model as a control variable, however, considering its insignificant effect, we excluded it from the model for parsimony. The results of the analysis provided support for the positive effect of connection ($b = 0.50, p < 0.01$) (H1) and communication ($b = 0.24, p < 0.01$) (H2) on alignment. The $R^2$ value of .46 demonstrates that the model explains a considerable amount of variance in strategic alignment.

We found partial support for H3. On one hand, the effect of centralization on connection and communication was insignificant. On the other hand, formalization had significant positive effect on both connection ($b = 0.20, p < 0.01$) and communication ($b = 0.13, p < 0.05$). The results also provided support for the positive effect of shared domain knowledge of IT and business executives on connection ($b = 0.28, p < 0.01$) and communication ($b = 0.18, p < 0.01$) (H4). Similarly, we found support for our arguments that when both business and IT executives foster good relationship, this, in turn, increases the connection of IT and business planning ($b = 0.20, p < 0.01$) as well as the communication between IT and business executives ($b = 0.19; p < 0.01$) (H6). Lastly, prior success of IT unit had significant effect only on the communication of the executives ($b = 0.26, p < 0.01$), thus providing partial support for H5. Both connection and communication have $R^2$ value of .31, indicating that the antecedents explain a good amount of variance in the drivers of alignment.

![Figure 2. Results of the PLS analysis](image)

Overall, we found strong support for our arguments regarding the positive effect of drivers of alignment on strategic alignment. Also, most of the antecedents had significant positive effect on the drivers of alignment as expected.

6 DISCUSSION

Studies in the IS planning and strategic alignment literature underline the importance of IT and business planning connection and communications between IT and business executives (Hu & Huang 2006, Reich & Benbasat 2000, Teo & King 1996). Building on these arguments, we hypothesized that the effects of antecedents are mediated by these current business practices. The statistically significant positive associations provided empirical support for these arguments. When the effects of these drivers are compared, the results indicate that the positive effect of connection is about the twice that of communication. This suggests that integrating the planning process is more essential than the level of communication between the executives in terms of achieving and sustaining strategic alignment. This finding is important in a sense that although previous studies postulated the effects of connection and communication, due to the qualitative nature of these studies, a comparison of their effects on alignment was not possible. This empirical study not only provides support for the qualitative studies, but also extends our understanding of how the underlying mechanisms of strategic alignment work. One possible explanation of this finding is that planning process is more formal and results in written
documentation, thus, enforcing a stronger form of alignment compared to communication which is bound to individuals and embedded in relationships. However, this finding does not in any means diminish the importance of communication between IT and business managers.

We identified five antecedents of strategic alignment based on literature review and validated their roles in alignment based on the data. First, we examined two characteristics of unit structure; centralization and formalization. In centralized structures, IT activities are supposedly coordinated at the corporate level, requiring less effort for alignment. Our results unexpectedly showed insignificant effects of centralization on drivers of alignment. Although studies in the alignment literature found similar results (Chan et al. 2006), more research is needed to fully understand the underlying reasons of this insignificant effect. On the other hand, the results provided support for positive effect of formalization as expected. Highly formalized organizations have clearly defined job descriptions and standardized policies and procedures, and they make use of task forces and committees more often. These characteristics are beneficiary for IT executives since unclear business goals and objectives present challenges to them. It is considerably easier and more effective for IT executives to integrate IT strategy into business strategy when the business strategy is formally outlined in the organization. Similarly, as a result of formalization, it is likely that executives would participate in more committees or task forces which would increase direct interactions and information sharing, which, in turn, would increase the level of communications and the number of channels used for communications.

Shared domain knowledge is one of the widely studied antecedents of strategic alignment. It refers to business and IT executives’ knowledge and understanding of each other’s environments. This understanding also brings the appreciation to each other’s accomplishments. If executives do not have shared domain knowledge, communications among them will be ineffective, and this will consequently inhibit the level of communications. In other words, without shared domain knowledge, effective communications shouldn’t be expected. On the other hand, when executives have shared domain knowledge, this creates more opportunities to achieve integration during the strategic planning process. Our results clearly support these argument by indicating that an increase in shared domain knowledge leads to better integration of IT and business plans and more frequent communication between IT and business executives.

Similar to shared domain knowledge, successful IT unit history is another commonly recognized antecedent. When IT units are reliable, credible, and deliver their commitments on time, this creates a positive impression on business executives. Our findings showed that successful IT history has a positive effect on the communications between IT and business managers. This outcome is only logical because business executives would like to consult some of their ideas with or query certain solutions from IT executives only if they consider the IT unit as credible, reliable, and successful. In an organization where the IT unit cannot deliver its promises, business executives wouldn’t take the time and effort to communicate with IT executives and the role of IT would certainly be marginalized.

Our findings also show that the significant effect of relationship management on both drivers of alignment is approximately the same, lending empirical support to the largely qualitative arguments in the literature. Having good relations between CEO and CIO is considered crucial for the quality of IS planning (Bai & Lee 2003) and it is also identified as an important enabler of strategic alignment (Feeny et al. 1992). When IT and business executives invest their time and give effort to have a good relation, this in turn would increase the level of communications between them. Furthermore, they would value more each other’s inputs during the planning process, and consider each other’s priorities, goals, and objectives as their own, resulting in higher levels of planning integration.

7 CONCLUSION AND CONTRIBUTIONS

Strategic alignment between IT and business has been one of the top concerns of practitioners and scholars. However, despite its documented positive effects on organizational success, only few organizations consider themselves in alignment (Luftman et al. 1999, Rosa 1998). We attempted to
fulfill the need for a comprehensive alignment model that not only integrates and extends the alignment literature, but also provides prescriptive insights to practitioners. Given the limited empirical findings in the literature, our first aim was to provide a more comprehensive framework that is empirically tested with survey data, and find support to the largely qualitative framework of IT-business alignment theory. Moreover, we chose to use SEM to test the proposed model to overcome the limitations of previously utilized techniques (i.e., ANOVA, t-tests). Also, by collecting the data from Turkey, we added a different dimension to the alignment literature, which mostly utilized data from US and Canada. Another contribution to the literature is the direct measurement of centralization in organizations. In the literature, organizational size was used as a proxy to centralization, assuming that small organizations are centralized and large organizations are decentralized.

The findings of this study present a more complete view of the strategic alignment process, thus providing a finer prescriptive guidance to executives for achieving and sustaining the alignment. First, with regards to communication and connection, business managers may reconsider their strategizing process and keep planning integration as one of their primary goals and seek opportunities to increase their communication level with IT managers. Second, given the positive effect of formalization on strategic alignment, executives can choose to increase the level of formalization in their organizations through different strategies including clearly outlining the responsibilities of IT and business units in terms of integration and implementing policies and procedures regarding decision making processes. Third, considering the importance of shared domain knowledge, business and IT executives should seek opportunities to increase their knowledge in each other’s domains by attending internal and external trainings and seminars. Forth, they should invest in time and effort to maintain a good relationship amongst each other. Lastly, IT executives should be more proactive in terms of increasing the visibility of their unit’s success and making sure that they deliver on their promises.

There are some notable limitations in this study that deserve attention. For example, although we based the causalities in our research model on the findings of the strategic alignment literature, this does not preclude potential recursive relations between some of the constructs. In addition, due to the characteristics of the data collection process, we weren’t able to calculate the non-response bias. One venue for future research is to investigate the proposed model with regards to the type of business strategies an organization is pursuing. Such an investigation would reveal how the effects of antecedents vary by strategy type. Also, this study was built on the common proposition that IT strategy should support business strategy. However, other schools of thought (e.g., synchronization or convergence of IT and business strategies) are also suggested in the literature (Sambamurthy 2008). Future research can investigate the model with respect to these considerably new approaches to strategic alignment.

References
Chan, Y.E. (2002). Why haven't we mastered alignment? The importance of the informal organization structure. MIS Quarterly Executive, 1(2), 97-112.


MAJOR ISSUES IN SISP: INSIGHTS INTO THE MAIN REASON OF SISP FAILURE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0059.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Information Strategy, IT Strategy, Organizational learning, User participation</td>
</tr>
</tbody>
</table>
MAJOR ISSUES IN SISP: INSIGHTS INTO THE MAIN REASON OF SISP FAILURE

Pita, Zijad, RMIT University, Building 108, Level 17, 239 Bourke Street, GPO Box 2476V, Melbourne, VIC 3001, Australia, zijad.pita@rmit.edu.au
Cheong, France, RMIT University, Building 108, Level 17, 239 Bourke Street, GPO Box 2476V, Melbourne, VIC 3001, Australia, france.cheong@rmit.edu.au
Corbitt, Brian, RMIT University, Building 108, Level 17, 239 Bourke Street, GPO Box 2476V, Melbourne, VIC 3001, Australia, brian.corbitt@rmit.edu.au

Abstract

The ‘SISP community’ is an evolving environment, which generally learns from the past. However, the lesson that top management commitment is a key determinant of SISP success has not been learnt. Regardless of the industry type or size, a lack of real commitment from senior management is still the main reason for the SISP formulation and implementation failure. The study introduces the SISP Stakeholders’ Designation construct to analyze commitment and participation of the available SISP resources in the light of maximizing success of SISP. Also the association between organisational learning and management commitment to SISP is examined. A postal survey of top management from 260 Australian companies revealed that obtaining high-level stakeholder engagement is critical to SISP success. The study presents optimal roles and level of engagement for the SISP Stakeholders’ Designation. It was found that if every managerial stakeholder’s designation is committed to SISP in all its phases it may result in wasting of valuable time and resources. The findings point to SISP learning reviews as a significant antecedent for managerial commitment to SISP.

Keywords: SISP success, top management commitment and participation, organisational learning.

1 INTRODUCTION

Organisations respond towards the pronounced need for strategic planning of IT/IS resources when they are critically dependent on IT/IS (Palvia & Palvia 2003). The need for Strategic Information Systems Planning (SISP) is present in small and large organisations (Porter 1998). It is nearly impossible to find an IS strategist who did not raise management commitment to SISP as an issue. The most general lesson to be learned from the more successful SISP cases is that when SISP has management commitment, success is almost guaranteed (Kearns 2006). Despite that, there is equal evidence that a top management support for SISP process in many firms is only declarative, and hence that is one of the reasons for SISP failure (Hartono 2003). Not believing in SISP ability to create a competitive advantage, drive revenue and innovation, could be a reason for top management lacking of real action. Or perhaps, excessive organizational commitment to SISP is detrimental (Basu et al. 2002). In any case, at present, business profitability is noticeably in decline, and accountability is a major requirement for SISP practitioners (Pisello 2001). The successful CIO can be a strategic business partner rather than an infrastructure provider and if the CIO holds strategic responsibilities he/she may significantly contribute to the organisation’s expansion and growth (Dearstyne 2004). Otherwise, IT practitioners could be the obstacles to implementing business solutions.

Little empirical research exists that examines involvement and commitment of SISP stakeholder’s like top managers, senior business, IT and user managers, IT computing personnel and users. In addition, because of very conceptual nature of the SISP studies, the relations among SISP stakeholders remains hidden on the variable level. Previous research (Basu et al. 2002) suggested further investigation in
this area, in particular, senior management involvement and the impact of additional matters like organizational learning on SISP success. This study is an empirical work conducted in that direction using data obtained from a large scale survey of 2000 Australian companies. Thus, close examination of the role of SISP stakeholders can provide valuable insights not available elsewhere.

The purpose of this paper is: (1) to confirm that traditionally the lack of commitment from top management (Lederer & Sethi 1988, Ward & Griffiths 1998, Teo & Ang 2001) is still the main reason for SISP failure (2) to examine the relationships between SISP stakeholder’s participation/commitment and SISP (3) to examine the association between learning reviews experience and top management commitment to SISP.

This paper proceeds as follows. After an overview of SISP problems, some key stands from SISP literature are presented to be able to form hypotheses and a context within which the survey results may be interpreted. The methodology section follows to explain the collection and analysis of data. Finally, a summary of the survey results and the possibilities for further research is presented.

2 THEORETICAL BACKGROUND

SISP is needed to produce a strategic plan that addresses the future needs for IT/IS in accordance with the business objectives in formal or less formal ways (Hackos 1997). However, over half of SISP plans formulated are never implemented, or fail to achieve their goals and objectives (Flavel & Williams 1996). The reasons for SISP failure remain fairly constant over the years. A summary of SISP problems (Yeo 2002, Cerpa & Verner 1998, Lederer & Sethi 1992, Ward & Griffiths 1998, Wilcocks 2000) has been compiled in Table 1 (problems are not ranked). The reasons for SISP failure very well correspond to the top five management concerns: IT and business alignment, attracting, developing, and retaining IT professionals, security and privacy, IT strategic planning, business process reengineering (Luftman 2006).

<table>
<thead>
<tr>
<th>SISP Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of management commitment and involvement</td>
</tr>
<tr>
<td>Misalignment of SISP and business goals</td>
</tr>
<tr>
<td>Problems with resources (recruiting &amp; education)</td>
</tr>
<tr>
<td>Security and privacy issues</td>
</tr>
<tr>
<td>Quality of the plan inadequate</td>
</tr>
</tbody>
</table>

Table 1. The Key SISP Barriers.

It is not only difficult to convince management to implement SISP, but is also difficult to convince them even to fund the initial SISP study (Lederer & Sethy 1999). The reason may be that top management may not understand the plan or they are not confident in IT’s ability to carry out SISP. Higher levels of managerial IT knowledge have positively influenced the extent of IT use, but top managers do not view information as a business resource to be managed for long-term benefit (Ward et al. 1998). They only appreciate it critically, when they cannot get what they need. As a result, the work is often delegated to ‘experts’ and thus managers are increasing the risk of losing control over IT applications. The applications turn out to be independent of the strategic context of the organization as a whole. Localized justification of investments can produce benefits that are actually counterproductive when overall business goals are considered (Ward et al. 1998). If control is lost, in particular, if the control of IS/IT investments is left to individuals or departments often striving to achieve incompatible objectives through IS/IT, the outcome can be disastrous (Broadbent & Weill 1997). Thus, top management involvement in IS/IT planning is crucial.

When the CEO and other top managers understand the value and potential of IS/IT assets for competitive advantage they provide support for SISP as allocation of quality resources. Commitment of the CEO to SISP would provide better knowledge about organisational objectives and can lead to a greater alignment between IS and organisational strategies (Kearns 2006) and consequently to greater
SISP success. While this alignment is necessary, it may not be sufficient to support the ‘simple’ explanation that the real planning or implementation problems are usually due to a lack of executive commitment. The reason is because it has been suggested that excessive top management commitment to SISP impedes SISP as it promotes excessive planning and provoke excessive bureaucracy (Basu et al. 2002). Therefore, these conflicting results are a great deal of confusion and the reason worth investigation of the optimal levels of top management commitment to SISP. This study acknowledges that SISP, as a highly political process, has complex interrelations between stakeholders of which a detailed investigation is a study per se. The aim of this research is to assess the stakeholders’ phenomena that are open to direct observation. Management commitment is one of them and the study in the light of overall SISP success tests the following hypothesis:

H1 SISP is more successful if: a) the CIO is committed to SISP from start to finish; b) the CEO is committed to SISP only at the start; c) the senior business management is committed to SISP from start to finish; d) the middle business management is committed to SISP from start to finish; e) the IS management is committed to SISP at implementation phase

In the terms of taking responsibilities for setting up the IT strategic directions, top management involvement is seen as unavoidable (Basu et al. 2002). However, IT professionals and business managers find it difficult to communicate and focus on improving their communications often caused more problems than it addressed (Luftman 2005). Very often they have a very different perspective on organisational values and how best to implement SISP. Implementation of SISP is not a trivial process and letting IT managers lead the SISP projects is seen as a potential obstacle to successful SISP implementation. The importance of the managerial role of IT managers is not in question, but IT has no ability to address the policy, procedures and organizational issues critical for successful business (Brown 1992). On the other hand, IT professionals are the only ones who really understand the IT process analysis. Thus, the question of what would be appropriate role of SISP participants can be raised. This study investigates the current stakeholders’ roles in Australian SISP practice and tests the following conclusions:

H2 SISP is more successful if: a) the CIO has a champion role; b) the CEO has a sponsor role; c) senior business management have a supporting role; d) IS management have an active leadership role

Successful SISP seems to require users and managers working in partnership with the IS function (Luftman 2005). Integration of inputs from all stakeholders should be an ongoing activity as their creative energy can help in generating more innovative strategies. Therefore, a participative SISP should enrich the SISP content (Ismail et al. 1996) and consequently will have a positive effect on overall SISP success. However, Ismail et al. (1996) also pointed to potential obstacles to stakeholders’ involvement. Many participants may slow down the planning process or have some other negative influence. Also, an obstacle to a participative SISP can be a managerial misalignment. A typical organisation has many types of managers who are suppose to collaborate but they often compete with each other (Goranson 2003). When they collaborate, i.e. if SISP is initiated by top management and IS management coalition, it would significantly add value to the SISP process (Spremic & Strugar 2002). In addition, when top management rather than IS management initiates the SISP study, more environmental assessment is done (Chi et al. 2005). Spremic et al. (2002) found that the main initiators of SISP in Croatian organisations are IS management (33%) and top management (27.5%). They relate the high rate of SISP failure due to the lack of line management initiating SISP. Flynn et al. (1993) found that more often SISP was initiated by the IT department than management (22% against 55%). Full integration or ‘partnership’ (Luftman 2005) of IS and business is seen as an imperative to avoid the barriers of a managerial misalignment, therefore it is hypothesised that:

H3 If SISP is initiated by business management and IS management coalition, it will be more successful.

The quality of people (intellectual capital) involved in the SISP process in terms of the ability to think innovatively, having adequate skills, knowledge, and experience is a key contributor to SISP. Appropriate business management knowledge of IT is significant for ensuring their commitment to
IT/IS projects (Kearns & Lederer 2000) and vice versa, IT management need to possess business acumen to improve strategic IT alignment with business (Teo & King 1997b). Learning from past experience and sharing knowledge is an important method of directing future SISP activities. Regular change reviews and learning reviews will enable better communication, knowledge sharing and improve likelihood of SISP success (Pai 2006). This discussion provides the basis for:

H4 Learning reviews positively influence management commitments towards SISP.

3 SISP SUCCESS (THE LACK OF FAILURE) AND CONCEPT OF SISP STAKEHOLDERS’ DESIGNATION

Multi-dimensional, multi-item measures of SISP success were proposed in the SISP literature. However, these dimensions of SISP success were examined and it was found that two dimensions: improvements in SISP capability and fulfilment of SISP objectives were not fully supported because of an overlap of these two concepts, suggesting the use of either of these dimensions (Warr 2006). As organisations do not have a single predominant SISP objective, SISP success as a composite of related perceptual items would rely on their average value and should be treated with caution. Thus, Basu et al. (2002) suggested the use of a single item measure for dependent variables. This study measured SISP success explicitly and operationalised it as a single item which measured the respondent’s perception of the overall SISP success. It is acknowledged that this measure can be biased, but some other researches (Pyburn 1983, Hartono et al. 2003) also used single items. SISP success is measured by directly asking the executives (business 31%, IT 69%) questions about their perception of SISP achieving its objectives. Other studies confirmed an insignificant error between the use of a single overall measure of success from the respondent and a multi-item measure (Warr 2006).

The extent to which powerful organisational actors are involved and committed to SISP will increase the success of SISP. Many studies (Ismail et al. 1996, Ruohonen 1991) categorise and characterise stakeholder groups for SISP. Generally, organisational stakeholders include business owners, data owners, developers and technical operational staff. SISP participants can be broadly characterised as managerial, IT non-managerial and other stakeholders like vendors and users. This study identifies managerial SISP stakeholders’ designations as: CEO, CIO, senior business management, middle business management and IS management group. Also, stakeholders’ designation is assessed through participation of other available resources, such as consultant, systems analysts, vendors, users and computer operations personnel. Participants are differentiated by the roles they have in the SISP process. For example, managerial members involved in SISP can have different roles, but top executives are supposed to have a champion role (Basu et al. 2002). Other roles are identified as supporting (assisting in all SISP activities), sponsorship (committing capital, personal time as well as resources for SISP, and making decisions with or without approval from higher-levels), and active leadership (guidance and empowering SISP team to higher levels of performance).

4 RESEARCH METHOD

This study uses only subset of data collected from an Australian-wide survey for a large study on SISP assessment and measurement. Data collection was completed by the end of 2003. The use of older data to mirror the current happenings ‘is not uncommon in SISP research because they tend to be relatively time invariable even though IT changes rapidly’ (Newkirk & Lederer 2006). To ensure generalizability of our findings, the questionnaires were administered to CIOs and senior IT executives in various industry sectors. Targeted key informants were IT executives, as they are usually the most involved in SISP (Segars & Grover 1998). To assess the validity of the answers, questions regarding respondent’s SISP involvement and experience in SISP were asked. A pilot survey was used to test the data collection method. The content of the questionnaire was discussed and refined through consultation with three prominent SISP practitioners being IT managers from three different industry types of organisations. A pilot survey
was completed with five organisations. The targeted companies were from government and private sectors, and they were small, medium and large in size. From a population of 2000 questionnaires sent, a reasonable number of questionnaires (260) with complete data were received. 86 surveys were received from companies that do not perform SISP but they supplied valuable data which is used to analyse the characteristics of organisations that do not perform SISP. This response rate (17.3%) is considered high as the chosen method of collecting data usually has a low response rate (Kress 1988).

Different statistical techniques were used to analyze the data. To calculate the strength of a relationship, Spearman's correlation coefficient, rho (\( r \)) was used for ordinal variables (Cavana et al. 2001) where data were not normally distributed. By convention, values of 0.15 or higher for rho are accepted as an indication of the strong relationships. Structural Equation Modelling (SEM) was used to confirm theoretical constructs that cannot be observed directly by explaining how the measured and latent variables are related to one another. Fit indexes, recommended by Byrne (2001), are reported but not discussed here as they are widely described in the statistical literature. Reliability of scales was checked for inter-correlation among their items using principal component factor analysis (PCA). The results of reliability tests are reported in Appendix A (Table 0.1).

5 DATA ANALYSIS AND DISCUSSION

Received responses indicated that about 24% of Australian organisations do not perform SISP at all. This study found that organisations that attempt IS planning in Australia in about 17% cases perform SISP regularly, 15% of them were developing SISP at the time the survey was conducted, 38% have irregular SISP, while about 31% of organisations have some form of IS planning.

Having 92.0% of respondents as active contributors to SISP, with 81.6% of them having more than 11 years of industrial experience, 89.6% having more than 6 years experience in the IS area, gave full credibility and confidence to the survey answers. The majority of respondents came from the manufacturing (18.4%), followed by public administration (11.5%), and banking (financial) services (9.2%). Industry types are grouped according to the size of the company: small (annual turnover < $10 million), medium (turnover between $10 and $500 million) and large for turnover of more than $500 million. Thus, respondents were: 5.7% from small, 62.1% from medium and 32.2% from large companies. About half of the surveyed companies have 100 to 1000 employees and 43.7% of companies have less than 10 IS employees. Only 36.8% of companies have more than 50 IS employees where about half of these companies employ more than 200 IS employees. The literature review revealed that SISP increases directly with the total number of employees (Groznik & Kovacic 2000, Teo et al. 1997a). The existence of relationships (Chi-Square=31.722, df=6, \( p<0.05 \)) was found in the case of the regular SISP planning. Thus, organisation size in terms of number of employees is a significant antecedent for conducting SISP.

Participants positions were: CEO (4.60%), CIO (33.33%), Information Systems Manager (35.63%), Divisional Manager (3.45%), Accounting Manager (3.45%), Financial Controller (4.60%) and General manager finance & administration (4.94%).

5.1 Key Reasons for SISP Formulation and Implementation Failure in Australia

The key reasons for the failure of SISP in Australian organizations are shown in Table 2. The rank position was calculated by ordering the mean values obtained as responses to the questionnaire in regard to the importance of the reason for SISP formulation and implementation failure (from (1) no importance to (5) crucial). The reasons for the SISP implementation and formulation do not differ significantly. The main reason for the SISP formulation and implementation failure in Australian organizations is the lack of commitment from senior management and the third and fourth failure reasons are also related to management (the ‘lack of senior management involvement’ and ‘IS
management is not part of the corporate planning process’). This is very much in line with the recent SISP literature (Palvia et al. 2003).

Budget limitation, as the second key reason for SISP formulation failure, at first was somewhat unexpected, but after finding that responses of business executives on their thoughts about IS/IT adding value to the business were only 32% positive, this became more clear. Misalignment with business objectives is the number one ranked management concerns (Luftman 2006) but this study found it to be the fifth reason for SISP formulation failure and as the second reason for SISP implementation failure. This may sound unexpected but on the contrary, when the common rationale for failures is known and widely publicized as IT - business misalignment, SISP actors focused their activities to avoid those barriers, which led to the reduction of SISP failures caused by misalignment at least in the plan formulation phase.

<table>
<thead>
<tr>
<th>Rank by Means</th>
<th>Key Reason for SISP Formulation Failure</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Key Reason for SISP Implementation Failure</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lack of commitment from senior management</td>
<td>1.99</td>
<td>1.54</td>
<td>Lack of commitment from senior management</td>
<td>1.78</td>
<td>1.42</td>
</tr>
<tr>
<td>2</td>
<td>Budget limitation</td>
<td>1.87</td>
<td>1.29</td>
<td>Lack of alignment with business objectives</td>
<td>1.67</td>
<td>1.35</td>
</tr>
<tr>
<td>3</td>
<td>Lack of senior management involvement</td>
<td>1.85</td>
<td>1.44</td>
<td>IS management is not part of the corporate planning process</td>
<td>1.66</td>
<td>1.34</td>
</tr>
<tr>
<td>4</td>
<td>IS management is not part of the corporate planning process</td>
<td>1.76</td>
<td>1.41</td>
<td>Lack of senior management involvement</td>
<td>1.64</td>
<td>1.27</td>
</tr>
<tr>
<td>5</td>
<td>Lack of alignment with business objectives</td>
<td>1.70</td>
<td>1.30</td>
<td>Budget limitation</td>
<td>1.63</td>
<td>1.20</td>
</tr>
<tr>
<td>6</td>
<td>Inadequate framework used for setting IT investment priorities</td>
<td>1.69</td>
<td>1.23</td>
<td>No adequate knowledge and expertise</td>
<td>1.62</td>
<td>1.19</td>
</tr>
<tr>
<td>7</td>
<td>Inappropriate planning horizons</td>
<td>1.66</td>
<td>1.14</td>
<td>Inadequate framework used for setting IT investment priorities</td>
<td>1.60</td>
<td>1.25</td>
</tr>
<tr>
<td>8</td>
<td>No learning from past experience</td>
<td>1.63</td>
<td>1.15</td>
<td>No motivation for the initialisation of SISP reviews</td>
<td>1.60</td>
<td>1.21</td>
</tr>
<tr>
<td>9</td>
<td>No adequate knowledge and expertise</td>
<td>1.61</td>
<td>1.21</td>
<td>Inappropriate planning horizons</td>
<td>1.55</td>
<td>1.12</td>
</tr>
<tr>
<td>10</td>
<td>No motivation for the initialisation of SISP reviews</td>
<td>1.60</td>
<td>1.20</td>
<td>Failure to consider the external business environment</td>
<td>1.53</td>
<td>1.16</td>
</tr>
</tbody>
</table>

Table 2 Key Reasons for the SISP Formulation and Implementation Failure.

Also, ‘problem with resources’, a highly ranked management concern is not highly ranked as a reason for SISP formulation failure – it occupies the ninth position. A possible explanation is that the normal SISP evolution process added to the knowledge and experience of those involved and the popularity of IT attracted many students over last decades. These are now manifesting in terms of higher availability of IT resources. These factors elevated the problem, but the lack of adequate expertise is still considered as one of the ten most important reasons for SISP failure.

The study tested the relationship between the reason for SISP failure and the company size and found no relationship (r <0.15, p>0.05). SISP failures occur in any size of organization. This finding reinforces the result of the study of Flynn and Goleniewska (1993).

5.2 SISP Stakeholders’ Designation

The SISP Stakeholders’ Designation construct is assessed through the analysis of participation and commitment of the available SISP resources in the light of enhancing SISP success. The survey results suggest that managerial commitment in Australian SISP practice is not so pronounced. About 35% of organizational management has no commitment towards SISP, about 63% is committed only at the...
start of the SISP process, and about 13.3% of management is committed only during the implementation phase. As an average, 46.7% of the managerial structure is committed from start to finish of the SISP process. The highest percentage is related to the commitment of CIOs. They are dedicated to SISP from start to finish in 65.52% of cases. Also, the survey results showed that IT required skills are generally available at an average rate of 82%.

The Stakeholder’s Designation construct is operationalised by four latent factors (Figure 1). The PCA demonstrates that the original segregation between SISP participation and commitment is not fully supported. There is some overlapping between those two factors. Generally, management participation is considered as a commitment to SISP; while for stakeholders, such as users, vendors, consultants, etc., participation is distinguished from commitment. Certainly, the participation of users cannot be taken as their commitment to SISP. It is important to point out that CEOs’ contribution to SISP emerged as a latent factor on its own, indicating the importance of CEO involvement in SISP. The structural path analysis confirms that the most important factor to SISP is the commitment of organizational management and the least important factor is who the SISP initiator is. Using SEM, the goodness-of-fit statistics shown in Table 3 confirm that this model fits the data well.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>DF</th>
<th>P</th>
<th>$\chi^2$/DF</th>
<th>RMR</th>
<th>GFI</th>
<th>AGFI</th>
<th>NFI</th>
<th>RMSEA</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>SISP –Stakeholder’s Designation</td>
<td>100.64</td>
<td>52</td>
<td>0.000</td>
<td>1.93</td>
<td>0.04</td>
<td>0.93</td>
<td>0.87</td>
<td>0.91</td>
<td>0.07</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Table 3  SISP Stakeholder’s Designation Model Fit Summary

Figure 1  Stakeholder’s Designation: Measurement and Structural Model.

5.3 Commitment to SISP

Hypothesis H1 is tested in two steps. First, a nonparametric correlation test is used to generally confirm the existence of the relationships between top management commitment and SISP success. Secondly, the mean values are compared to discover SISP maximums for various levels of commitment. Correlation statistics ($r=0.55$, $p<0.001$) generally support H1. The SISP success means for various levels of commitment are graphically presented in Figure 2. H1 (a), (d), (e) are fully supported. If senior business management is committed to SISP from start to finish, the success of SISP may not be greater than if their commitment is only at the implementation of SISP, hence H1 (c) is not supported. Figure 2 also shows that CEO’s commitment is more important in the implementation phase than at the start of the SISP process, thus H1 (b) is not supported. Hypothesis H1 can be interpreted that while top management commitment to SISP is generally positively related to SISP, it could be wasting of time and resources if every managerial shareholder’s designation is
committed to SISP in all its phases. Also, obtaining a management support in the implementation phase is more important than at the start of SISP.

Gottschalk (1999) found a relatively lack of importance of management support for the implementation of SISP, which is quite the opposite of the findings in this study. This study can offer several explanations for the present results. Firstly, there is no guarantee that a good plan will be adequately translated into action plans (Hartono et al. 2003, Teo & Ang, 2001). Furthermore, high SISP failure rates and the promoted importance of SISP success may influence greater support from management during implementation to ensure success of SISP. Finally, only implemented SISP can be (more or less) successful, but non-implemented SISP plan (regardless of its quality) is 100% failure.

![Figure 2: SISP Success and Managerial Commitment Relationships (categorical)](image)

Basu et al. (2002) found that organisational commitment predicted SISP success in an inverted-U relationship, i.e. as organisational commitment increases, SISP success increases until it reaches a maximum; as organisational commitment continues to increase, success decreases. A close lookup of the variables in a quadratic regression model (the square of independent variables used) revealed that in this study only senior business management commitment predict SISP success with an inverted-U shape relationship (Table 4). This could be because they increase the level of bureaucracy, which impede SISP. Therefore, strong devotion from the CIO and CEO does not impede SISP. Maximum SISP success is achieved if the CIO is committed from start to finish of SISP.

<table>
<thead>
<tr>
<th>Inverted U-model</th>
<th>R</th>
<th>Adjusted R Square</th>
<th>S.E.</th>
<th>Beta</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEO commitment</td>
<td>0.43</td>
<td>0.18</td>
<td>0.57</td>
<td>-1.13</td>
<td>-1.84</td>
<td>0.07</td>
</tr>
<tr>
<td>CIO commitment</td>
<td>0.51</td>
<td>0.26</td>
<td>0.55</td>
<td>-1.16</td>
<td>-0.88</td>
<td>0.38</td>
</tr>
<tr>
<td>Senior business mgmnt commitmt</td>
<td>0.42</td>
<td>0.17</td>
<td>0.58</td>
<td>-2.03</td>
<td>-4.08</td>
<td>0.000*</td>
</tr>
<tr>
<td>IS management commitment</td>
<td>0.51</td>
<td>0.25</td>
<td>0.45</td>
<td>0.59</td>
<td>0.55</td>
<td></td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.01 level (2-tailed).

Table 4: Regressions between SISP Success and Managerial Commitment.

Managerial commitment increases as the size of the company increases. This relation is statistically significant (r=.34, p<0.01). When this relationship is inspected at an item level it was found to be a non significant coefficient for the CEOs. This is possibly because this designation does not commonly exist in small and medium organisations. There is no similar study to compare this result as the SISP literature mainly investigates the SISP phenomenon in large companies.

5.4 SISP Participation

The survey data shows that the highest SISP participants are CIOs (mean 3.14, std. dev. 1.49) and IS managers (mean 2.95, std. dev. 1.32), which is widely supported in the SISP literature (Teo et al.
Champion, sponsor and active leadership roles are the most pronounced roles for CIOs. The CEO has mainly more a sponsor and than a champion role. The data shows that other IS managers mostly have a champion or an active leadership role while senior business managers have a sponsor and an active leadership role most of the time.

It was found that only management participation is significantly correlated to SISP success ($r = .55$, $p < 0.01$). The influence of IT personnel as systems analysts, developers, programmers and other computer operations personnel on SISP success is not support. Consultants did not influence SISP success; perhaps they had not carried the significant SISP roles. This scale was assessed at an item level and a weak but significant relationship for users participation ($r = .183$, $p < 0.05$) was found. This is in agreement with the findings which suggested the involvement of users in SISP (Palanisamy 2005), and that computer experts should not handle SISP (Lederer et al. 1999). Still other stakeholders, not explicitly targeted by the study, could be influencing SISP success ($r = .18$, $p < 0.05$). Finding that SISP participants are mainly from senior management does not diminish the findings of Ismail et al. (1996), it simply could mean that wider SISP participation needs to be promoted in Australia. Also, the highest SISP success mean value is associated with the CIOs and IS managers’ participation in SISP.

The correlation statistics generally support H2 ($r = .55$, $p < 0.01$). The mean values depicted graphically (Figure 3) partially support H2: H2(a) and H2(b) are supported as the highest mean for SISP success is achieved if the CIO and CEO have a champion and a sponsor role respectively. H2(c) is rejected as the highest mean for SISP success is achieved if a champion is selected from senior business management. H2(d) is also rejected on a ground that IS management is more successful in a sponsorship role rather than in an active leadership role. These results are somewhat unexpected. They show that if the champion is elected from IS management, SISP may not be so successful, which could not be said for senior business management. On the contrary, if the champion is elected either from senior business management or if that role is fulfilled by the CIO, the SISP success rate may be equal.

![Figure 3](image-url)  

**Figure 3**  
**SISP Success & Main SISP Participants and their Roles (means.)**

SISP management participation is statistically significantly correlated ($r = .25$, $p < 0.05$) with company size. Large organisations are more associated with SISP success in comparison with small or medium sized companies. As SISP success is influenced by participation, small and medium sized companies should promote more participative SISP.

**5.5 SISP Initiators**

This study found that top management initiates SISP in 41.81% of cases (CIO - 28.25% and CEO - 13.56%). Senior business management more often initiated SISP than IS management and in their coalition, SISP was initiated only in 8.47% cases. The result that senior business management more
often initiated SISP (18.1%) than IS management (15.8%) is not consistent with the earlier findings of similar studies. IS management is expected to play a leading role in SISP, as confirmed by the survey results of Teo et al. (1997a). Perhaps, a high percentage of SISP initiation by CIOs’ is a compensation for the somewhat lower than expected percentage of SISP initiation by IS management. This research fails to support hypothesis H3 i.e. SISP is more successful if it is initiated by a senior business manager and an IS management coalition. The correlation shows positive and significant relationships only between the CIOs (rho =0.22, p<.005) and CEOs (rho =.28, p<.001) as initiators and SISP success. The underlying data show that SISP success can be as much as doubled if it is initiated by CIOs (mean .57 std. err. 0.038) rather than CEOs (mean 0.28, std. err. 0.034).

5.6 Organisational Learning

The sharing of experience and knowledge from SISP learning reviews is positively and strongly correlated (r=.79, p<.01) with commitments of management designations and consequently has an impact on SISP success. This statistic confirmed H4. Thus, when attendees of reviews have an opportunity to share learning about SISP they become more compelled to the process they better understand and they are capable to influence. Possibly, their communication and cooperation improve which reaffirm their commitment.

The scale mean value (2.62 on a 1 to 5 point Likert scale, std. dev. 1.43) shows that Australian organisations do not perform the learning meetings a great deal. Only about 36% of the population positively answered the question on the learning review. This could be probably one of the main reasons for the lack of managerial commitment to SISP. Consequently, if an organisation pays more attention to organisational learning, in particular to the sharing of experience on SISP, more commitment from managerial structure could be expected. Therefore, this study adds further validation to the work of Pai (2006), where the relationship between knowledge sharing and SISP in Taiwan was investigated.

6 CONCLUSION AND FURTHER RESEARCH

This study confirmed that the lack of managerial commitment and involvement are still the main reason for SISP failures. To make SISP deep and far reaching, the planning modules such as a commitment plan should be part of the SISP study. In a time of limited resources, eliminating unnecessary duplication of efforts and delegating clearly defined and most suitable roles to SISP stakeholder’s will lead to more successful SISP. While top management commitment to SISP is necessary in all SISP phases, obtaining their support in the implementation phase is more important than at the start of SISP. Practitioners may find it of interest that SISP success can be as much as doubled if it is initiated by CIOs rather than CEOs. In that case, CIOs would most likely have a champion role, they would more actively participate in business planning, gain better understanding of business needs and more successfully promote strategic aspects of SISP. The findings also point to lack of SISP learning reviews which are a significant antecedent for managerial commitment to SISP.

Future research may involve multiple respondents from the same organization to ascertain top management’s perception of SISP. This study used general description of SISP variables. Perhaps, the use of more specific scales could reveal different results. If future research is associated with specific types of organisations, their results could be extended into fine granularity and greater comprehensiveness. SISP is one of the processes within the overall IT governance and understanding influence and causality of other factors on managerial commitment to SISP would lessen the barriers to SISP success.

APPENDIX A

<table>
<thead>
<tr>
<th>Description</th>
<th>No of Items</th>
<th>Alpha</th>
<th>KMO</th>
<th>Total variance</th>
<th>Scale/Av. Rating</th>
<th>Note</th>
</tr>
</thead>
</table>

Proceedings ECIS 2009
<table>
<thead>
<tr>
<th>Scale</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>% Valid</th>
<th>Range</th>
<th>Items dropped</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q24 SISP participants</td>
<td>10</td>
<td>.83</td>
<td>.78</td>
<td>70.2%</td>
<td>1 to 5/2</td>
<td>Items dropped: Q24.5, Q24.10</td>
<td>F24_1(4), F24_2(3), F24_3(3)</td>
</tr>
<tr>
<td>Q25 SISP initiators</td>
<td>9</td>
<td>.67</td>
<td>.55</td>
<td>61.7%</td>
<td>0 to 1/1.19</td>
<td>Items dropped: Q25.4, Q25.6</td>
<td>Q25_f1(4), Q25_f2(3), Q25_f3(2)</td>
</tr>
<tr>
<td>Q26 Commitment to SISP</td>
<td>5</td>
<td>.86</td>
<td>.81</td>
<td>77.3%</td>
<td>1 to 4/1.72</td>
<td>Factor: F26_1</td>
<td></td>
</tr>
<tr>
<td>Q30 Learning Reviews</td>
<td>4</td>
<td>.97</td>
<td>.86</td>
<td>92.34%</td>
<td>1 to 5/2.62</td>
<td>Factor: F30_1</td>
<td></td>
</tr>
<tr>
<td>Q38 SISP formulation failure</td>
<td>21</td>
<td>.983</td>
<td>.88</td>
<td>75.3%</td>
<td>1 to 5/1.61</td>
<td>Factors: F38_1</td>
<td></td>
</tr>
<tr>
<td>Q39 SISP implementation failure</td>
<td>19</td>
<td>.986</td>
<td>.79</td>
<td>79.9%</td>
<td>1 to 5/1.52</td>
<td>Items dropped: Q39.8, 9, 19, 20</td>
<td>F39_1</td>
</tr>
<tr>
<td>Q42 Overall SISP success</td>
<td>5</td>
<td>.872</td>
<td>.78</td>
<td>66.9%</td>
<td>1 to 5/3.32</td>
<td>Factor: F42_1</td>
<td></td>
</tr>
</tbody>
</table>

Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (cut off point >0.5)

Table 0.1  Reliability and PCA Summary for Scales Used for Statistical Analysis

References


THE ROLE OF STRATEGY IN THE EVOLUTION AND INNOVATION OF INFORMATION SYSTEMS: A SIMULATION EXPERIMENT

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0734.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Strategy, IT Strategy, Organisational Change, Simulation</td>
</tr>
</tbody>
</table>
THE ROLE OF STRATEGY IN THE EVOLUTION AND INNOVATION OF INFORMATION SYSTEMS:  
A SIMULATION EXPERIMENT

Kevin P. Gallagher, Business Informatics Department, College of Informatics, Northern Kentucky University, Nunn Dr. 233, Highland Heights, Kentucky, 45226, kevin.gallagher@nku.edu

Bryan Hosack, School of Information Technology, Illinois State University, Campus Box 5150, Normal, Illinois, 61790-5150, bhosack@ilstu.edu

Abstract

Information Systems (IS) have come to play a larger role in how organizations function, including consideration in strategic planning. Responding to competitive change or enacting a new strategy often requires modifying or enhancing an organization’s IS. In this paper, the authors develop a process model of evolutionary change to address the role that IS plays in organizational strategy. The model is based in part on Miles’ strategic typology of Defenders, Prospectors and Analyzers. The model examines how strategic approaches influence attention to environmental change, investments in IS capabilities, and evolution of information systems.

The behavior observed in the simulations corresponds with the theoretical hypothesis developed from the model. The conservative defender organization experienced few discontinuous changes in the simulated 20 years. The innovative prospector organization exhibited extensive change, while the balanced approach of the analyzer organization fell between the two extremes. This study has taken steps to create simulations that approximate the behaviors predicted by a theoretical process model. The success of the simulation modeling in this study builds a foundation for future research.

Keywords: Strategy, IS Strategy, Organizational Change, Simulation.
1. INTRODUCTION

Responding to changes in the competitive environment, be it to take advantage of new opportunities or to respond to competitive threats, remains an important concern for organizations and the management of their information systems (IS). Such environmental changes may necessitate modifications or enhancements to an organization’s IS in order to facilitate changes to products, services and the work processes that support them. Possessing or lacking the IS capacity and capabilities to undertake such changes may hold important implications for an organization’s performance.

The occurrences of change in the environment are not completely independent events. To an organization, change events may will hold greater or lesser relevance depending on the importance that the events and adaptability hold for an organization’s strategy. Understanding the implications of different strategies, the occurrence of environmental change and how those events effect the evolution of systems can provide important contributions to theory and practice. Yet, little research has explored the connections and interrelations of organizational strategy, environmental volatility and the evolution of IS.

This research explores these relationships by examining how opportunities and demands for change in organizations affect the evolution of information systems. Our approach is to model patterns of environmental change and the evolution of IS capabilities in organizations by conducting controlled experiments with a computer simulation, an approach identified as one that can aid in the study of IS strategy (Sabherwal et al. 2001). We believe this theoretical model and simulation establishes a foundation for continued research in this area.

2. THEORETICAL FOUNDATIONS AND RESEARCH MODEL

A fundamental concern in strategic management is the relationship an organization has with its environment. In approaching the evolution of an organization and the interaction between strategy, environment and information systems, we subscribe to the strategic choice perspective as outlined by Miles, et al (1978). As stated, “organizational behavior is only partially preordained by environmental conditions”. Management determines target customers and markets, in conjunction with investments in technical and administrative functions (Miles, et al. 1978).

Management also makes choices about investments in resources. From a resource-based view, Grant conceptualizes strategy as “the match an organization makes between its internal sources and skills…and the opportunities and risks created by its external environment” (Grant 1991, p. 114). Investments in capabilities play a role in how organizations perceive their ability to respond to opportunities and demands. As a result, there is an interrelationship between the occurrence of environmental change, and the evolving design of products and services and the perceived need to undertake incremental or discontinuous changes in an IS.

2.1 Strategic Choice and Adaptation

Undertaking an evolutionary perspective, we are interested in what Miles, et al called the “adaptive cycle”. In this cycle, organizations seek to solve several interrelated problems; entrepreneurial, engineering and administrative. entrepreneurial activity is realized through the enactment of an organization’s market position and product and service designs (Venkatraman 1994). The engineering problem is realized in their enactment of technology, such as information systems. The administrative problem has two conflicting issues; how to create structures and processes that enable the current strategy, while also enabling future innovation. We approach the administrative choices as realized investments in the ability to deliver both current and future capabilities.

The challenge for organizations is in “not allowing the systems to become so ingrained that future
innovation activities are jeopardized” (Miles, et al. 1978). We view IS as a key strategic resource for organizations trying to enact changes in product, service and market approach. However, the structures and processes of an IS designed to serve today’s environment can enable, but sometimes constrain future capacity and capabilities for future innovation. Miles et al. asked how do organization’s move through this adaptive cycle? We ask how do information systems evolve as these adaptive cycles unfold?

Miles et al. (1978) proposed a typology of three strategic types that characterize organizational approaches to dealing with environmental change, uncertainty and adaptation—defender, prospector, and analyzer. Defenders are concerned with maintaining stability and efficiency in the design of their products and services, systems, processes and technologies that support them. On the other end of the continuum is the prospector strategy, in which flexibility and change are the driving forces. Prospectors explore new markets, product and services continually, much the same as the innovating firms explored in the work of Nelson and Winter (1982). This typology was applied to IS strategy by Sabherwal and Chan (2001), as well as others who have built on their research.

Lying between these two are analyzers, who try to undertake the difficult balance of maintaining flexibility while being efficient. Analyzers behave in similar ways as imitating firms conceptualized in the work of Nelson and Winter (1982). Analyzers wish to reap the benefits of both Defenders and Prospectors by remaining adaptable and controlling costs. A fourth possibility is the Reactor. The reactor is characterized by Miles et al. as a non-strategy and will therefore not be included in this study (see also, Sabherwal and Chan 2001).

These three approaches to strategy influence how organizations respond to environmental change. For this paper, we define environmental change events by their frequency, amplitude and instability, as defined by Wholey and Brittain (1989). We draw on work by Child (1972) and Hannan and Freeman (1977) to develop these constructs. Frequency is of course the number of environmental changes identified over a set period of time. Amplitude is the magnitude or size of changes experienced by organizations. Instability is the randomness resulting from uncertainty and inability to predict future environmental events.

2.2 Information Systems and Strategic Adaptation

Advances in information technology increasingly offer organizations the opportunity to adopt or create new innovative products, work processes and market strategies (Sambamurthy and Zmud 2000). An evolving information system allows organizations to extend their processing capacity and the capabilities to support business processes. By extending and enhancing their existing information system, organizations can expand the system’s range and reach (Keen 1991), while also developing new capabilities (Copeland and McKenney 1988).

Effectively leveraging information technologies can enable organizations to sense and respond to opportunities and demands in their competitive environments (Sambamurthy, et al. 2003). These IS capabilities are formulated in both technologies and human resources and provide the ability to employ resources in ways that enable delivery of new products and services (Bharadwaj 2000).

This approach to strategy is supported by the resource-based view of the firm (Penrose 1959, Grant 1991, and Bharadwaj 2000). Bharadwaj (2000) discusses the role of IT infrastructure and IT human resources. In this paper, capacity is defined as the resources, technology and man-power that allow organizations to cope with change. Bharadwaj (2000) also discusses the intangible IT resources that a firm holds, which are collectively grouped as capabilities in this study. These assets can include processes, routines, and regular and predictable behavior patterns of a firm (Nelson and Winter 1982). These assets are improved over time through organizational learning and experience (Cohen and Levinthal 1990).

At the heart of Miles’ et al. work is the importance of innovation. Innovations are viewed in this study as realized changes to the design of products or services that take advantage of opportunities or respond to
competitive demands in the organization’s environment. IS can enhance an organization’s ability to respond to these demands, adapting its product and service offerings by, for example, altering their product scope and market reach.

The scope of the organization’s product is defined for this study as the number of components in the product. If the components are closely interrelated, then any change in the design of one component will require corresponding design changes to others. This association increases the rate of change in a system, as adoption of one innovation creates ripple effects throughout the overall IS design. Reach is the number of markets in which a firm competes. Reach is determined, for example, by geographic or channel expansion. Extending reach, an organization encounters a greater number and variation in competitors, regulations, and customer needs.

The dynamics of the adaptive cycle and turbulence in the environment increases with the growth of scope and reach. So relevant changes are expected to occur more frequently, especially those having a greater amplitude and instability when moving along the continuum from defender to prospector. As the number of changes increases, it raises the demands placed on an organization’s IS department. Thus, the organization’s choices when solving “engineering” and “administrative” problems hold implications for its ability to adapt an IS in a way that enables it to realize its chosen strategy.

2.3 Operationalizing Change in Information Systems

In the management of IS, maintaining or developing flexibility relies on the development of certain technical and organizational capabilities. Capacity is supported in information system design through the concept of scalability, which allows for growth in the number of users and the volume of data processed, but also to expand the type and volume of data collected and stored. For human resources, capacity is concerned with the presence of slack resources, which facilitate the ability to undertake change and provides adequate time for planning, monitoring and learning from experience. Adequate levels of staffing in the organization influence the ability to support the adoption of innovations.

Important to the development of capabilities is organizational learning. Learning is realized and accumulated over time and influences the organization’s tendencies toward remaining innovative and maintaining its ability to implement change (Copeland and McKeeney 1988). Capabilities for change are enhanced as the organization experiences events, especially those offering greater development of experiences and knowledge (Cohen and Leventhal 1990).

Alternatively, inertia constrains an organization’s capacity and capabilities for change. Inertia builds through a series of design decisions (Gersick 1991), escalating commitment to decisions over time (Huff and Huff 2000), and exploiting current knowledge and capabilities (March 1990). Depending on various factors regarding an information system, such as its age (Swanson 1994) and design (Broadbent, et al., 1999), technology may enable or constrain the implementation of change in information systems.

The inability of an organization to respond to desirable opportunities or competitive threats in the environment due to inertial constraints will increase stress in the organization. Stress is realized in the desire for new approaches to resolving existing problems (Huff and Huff 1991). This often translates into a radical break from existing structures and practices (Tushman and Romanelli 1985).

Inertia and stress are not mirror images of the same concept, but complimentary (Huff and Huff 1991). As inertia rises in a system, stress also rises, but the rate at which stress increases is also dictated by the demands placed on the capacity and capabilities of IT to implement change. If stress exceeds the organization’s threshold of acceptability, then the organization will begin to look beyond existing solutions and begin searching for alternatives (Tushman and Romanelli 1985). Until the tolerance for stress is surpassed, then the cumulative effects of inertia will continue to constrain the organization’s ability to implement change, potentially stifling innovation (Huff and Huff 2000).
Outcomes that result from the tension between inertia and stress are patterns of incremental and radical change that punctuate long periods of equilibrium with short periods of radical change (Huff and Huff 2000). Such patterns are consistent across theories of scientific advancement, organizational change and information systems (Kuhn 1996, Tushman and Romanelli 1985, Gersick 1991, Orlikowski 1993). Changes in this study are classified as small adaptations that amount to the continuation of the status quo; enhancements to the system that extend its reach and range and discontinuities that reengineer or replace significant parts of the system.

2.4 An Evolutionary Process Theory of IT Innovation

Simulations were constructed based on the research model proposed below in Figure 1. The model includes relationships found in the literature review above. The behavior of the model also simulates the findings of historical case studies conducted in the U.S. property and casualty insurance industry (Gallagher 2008, Gallagher and Worrell 2008, Gallagher 2002). The studies documented the occurrence of discontinuous changes in these organizations over a thirty years period. The organization’s strategy and attention to changes in the competitive environment resulted in varying occurrences of discontinuous change events, i.e. the reengineering or replacement of IS.

The model is based on the theoretical relationships among organizational strategies (defender, analyzer, prospector), their implications for managing the competitive environment (product scope, market reach), and occurrence and characteristics of change events (frequency, amplitude, instability). The ability to implement change is dependent on an organization’s IS capacity and capabilities. When the ability to respond to change is constrained by increasing levels of inertia, the organization’s stress will rise. Ultimately, the inability to respond to the environment leads to occurrences of discontinuous change.

In the research model, changes in the environment are encountered as opportunities and demands for change that ultimately result in the evolution of the system. The outcomes of these events are either a continuation of the status quo (no change adopted) or the enactment of change events. The events are adopted as either incremental or discontinuous changes, based on the events and the current state of the organization.
system. The incremental or discontinuous nature of the change affects the evolutionary patterns of change in a system.

The organization’s strategy (using Miles, et al. 1978 Defender, Analyzer, or Prospector) acts as the directional forces in the model that influence the trajectory of events. Of specific importance are the strategy’s influence on scope and reach of the organization’s products and services and the resulting demands for change in an IS.

3. GENERAL SIMULATION MODEL

The simulation experiments conducted in this study were discrete and event-based. Four simulation experiments were conducted. Each simulation model in the experiment ran for 20 years (after an initialization period of two years). Arena simulation software was used to create the models. Since the models presented here are intended as a baseline for future studies, other simulation approaches maybe consider as model complexity increases.

As the simulated organizations encounter opportunities and demands for change in the competitive environment, characterized by their frequency, amplitude, and instability, the model describes how these variables influence the evolution of an IS and ultimately an organizations ability to implement change. The ability to change is enabled by having both sufficient capacity and capabilities. Over time, as existing processes are stabilized through status quo changes (i.e., changes within the capacity and capabilities of the organization) and adopted incremental changes (i.e., changes within either the capacity or capabilities of the organization, but not both). Thus, each event has the ability to increase the level of inertia in an IS due to the passage of time, increased size and complexity. The challenge is to then alleviate the stress that builds as the ability respond to environmental changes decreases. For example, lacking adequate capacity for adopting an incremental change, a major enhancement to the system may be necessary, requiring a discontinuous change.

The first variable associated with a change is frequency, which represents the stability or turbulence of the competitive, regulatory and/or technical environment in which the organization competes. Frequency is influenced by strategy and determines the number of changes that ultimately enter the simulation. As changes enter the simulation, they are assigned, amplitude and instability values. Based on these values, a change will be considered as a potential enhancement for the IS. If a change enters the system, it is then evaluated, based on the amplitude, which represents the time, effort and talent required to implement a change into the existing systems. Amplitude is compared with organizational capability to handle the change. The other variable that is considered is the instability of a change, representing the difference from the existing information system’s technology, the capabilities of the IT personnel and the design of the products and services it supports. The consideration of these variables is to determine if there is enough capacity and capability to handle the change. For each simulation, inertia acts as a multiplier of amplitude and instability when comparing these values with capability and capacity, respectively. When the current capacity and the capability are sufficient, the change is made and inertia increases based on a portion of the amplitude associated with the change. In this situation, only half of the amplitude is added to the current level of inertia for status quo changes.

When technical capacity are inadequate at the time a change is encountered, but there is enough knowledge capability to handle the change, an incremental change occurs increasing inertia at the full amplitude associated with that change. The logic behind this process is indicative of a situation where excess capacity is not available, yet the acquisition of additional resources can supplement the adoption of the intended change. Capacity and capability are both fixed at 0.75 for the purpose of our simulations. Future studies will examine varying capacity and capabilities. The function for the increase of inertia with regard to capacity and capability is shown in Equation 1.
The implementation of status quo and incremental changes will increase the levels of inertia in a system as enhancements to the design increases the size and complexity of the system. The inability to respond to changes as a result of constraints on knowledge capabilities or a combination of inadequate capacity and capabilities creates stress in the organization. Stress is realized in the desire for new and innovative approaches to the resolution of problems and cumulatively contributes to the organization’s consideration of radical change. In this case, radical changes translate into the redesign or replacement of some or all of the existing system. Once a new system is implemented and the organization’s capacity for change and the capabilities to adopt future changes are available, stress is reduced. As inertia builds in the simulation due to status quo and incremental changes, stress begins to build with each change that is not within the capabilities of the organization at a rate of half the amplitude associated with a change. If both capacity and capability are inadequate to deal with a change, stress increases by the full amplitude associated with the change. When discontinuities occur the stress variable is reset. Stress increases under these conditions, as shown in Equation 2:

\[
G(S) = \begin{cases} 
\frac{S + A}{2}, & \text{if } I \times A \leq Cp \text{ and } I \times Is \leq Cb \\
S + A, & \text{if } I \times A > Cp \text{ and } I \times Is > Cb \\
0, & \text{otherwise}
\end{cases}
\]

Equation 2: The Growth of Stress

Each simulation experiment builds on this simplified model. Three levels of complexity are added to provide a higher-level organizational realism. Each experiment is outlined below. First, the implementation of strategy in the simulation is discussed. Next scope and reach are added to the strategy models. In the third set, capacity and capability are added to the model. Finally, all variables are combined into an integrated model that includes strategy, scope, reach, capacity, and capability.

3.1 Experiment 1: Strategy and the Environmental Conditions

Strategy determines how the organization sets performance objectives and the demands placed on its systems to adopt changes over time. First, goals may focus on growth, stability or both. As a result, demand for change differs across organizations or within any one organization over time. For instance, growth may expand the geographic reach of the firm, resulting in increased opportunities and demands for innovation as more customers, competitors and regulators must be considered. This places increasing demands on the capacity and capabilities for change in the IS. Stability on the other hand, represents a focus on using IS to increase efficiency, which in turn leads to increased profitability in the short run.

Three simulations were conducted—one for each proactive strategy as outlined in the typology offered by Miles et al. (1978). The Defender approaches environmental conditions with the goal to minimize the impact of frequency so that fewer discontinuous changes occur. The Prospector, on the other hand, encounters change events of larger size having a much higher frequency, amplitude, and instability associated with change than the Defender. The Analyzers try to balance efficiency with opportunity in
evolving the system, falling somewhere between Defenders and Prospectors. In the first simulation, change events occur with a probability of 25% for Defenders, 50% for Analyzers, and 75% for Prospectors.

Within the first set of experiments, the organization’s IS capability and capacity for change are fixed at the same predetermined level for each strategy. Prospectors carry a lower level of inertia, due to the desire for rapid change increasing. Defenders carry the highest levels of inertia because stability is valued in the quest for efficiency. In contrast, Prospectors have a much lower threshold for stress than Defenders due to the perceived need for change. Again, Analyzers fall somewhere in between trying to integrate both the Defenders and Prospectors goals. The threshold for each strategy was set to allow for adequate variation between strategies based on pilot studies. The first experiment tests the following hypothesis:

**H1:** Prospectors will experience a greater number of discontinuities than Analyzers, which in turn will experience a greater number of discontinuities than Defenders.

### 3.2 Experiment 2: Strategy, Scope and Reach

The second set of simulations introduces the effect of scope and reach associated with changes in product and service designs. Reach defines the size of the geographic market, while scope deals with the range of customers served in a market. Scope and reach will increase the frequency, magnitude and instability of the demands for change. Specifically, scope of the product will increase the frequency and magnitude of these events as it extends the number of potential customers within a population by expanding the range of possible contingencies necessary to compete across customer segments. Conversely, reach increases the frequency and instability of changes entering the system. For example, if a new set of features are added to an existing set of product dimensions, it increases the complexity of the overall design. Within the simulation, the set of environmental conditions (frequency, amplitude, and instability) increase over time as scope and reach is increased, resulting in more frequent and varied change events entering the system.

In this experiment, the Defender will focus on stability and efficiency, minimizing the scope of products offered and the reach into different markets, thus having a narrower scope to its products and more limited in changes to market reach. The Prospector on the other end of the continuum will always be scanning for new markets to enter and products to offer, so it will experience greater frequency of change events. The Analyzer balances scope and reach to achieve efficiencies, without passing up profitable opportunities. The effect of scope and reach are hypothesized as follows, based on an organizations strategy:

**H2:** Increases in scope and reach of products will result in more discontinuous enhancements to the system for all strategies when compared to Experiment 1.

### 3.3 Experiment 3: Strategy, Capacity and Capability

Until now capacity and capability have been held constant across strategy. However, it is unrealistic to expect organizations not to invest in capacity or develop greater capabilities as they learn from their experiences in adopting change. Investments and system feedback, therefore, are incorporated in our model. Each strategy is assumed to improve both capacity and capability, but each strategy increases at a rate of change consistent with its intended goal. As incremental changes are adopted (versus status quo changes) capacity increases to accommodate the change. This expanded capacity can then be exploited to meet future demand for change. Defenders focus more heavily on capacity investments to increase the efficiency of the system. Prospectors, in contrast, are less concerned with capacity in favor of increasing capabilities and Analyzers strike a balance between the two ends. A feedback loop for learning occurs to reduce stress as the organization handles status quo change events increasing the organizations confidence as shown in Equation 3:
$H(S) = \begin{cases} 
S - \frac{A - IS}{4}, & \text{if } I \times A \leq CP \text{ and } I \times IS \leq CB \text{ and } S - \frac{A - IS}{4} \geq 0 \\
0, & \text{otherwise}
\end{cases}$

Equation 3: The Decrease of Stress from Increased Knowledge Capabilities

Learning stabilizes each strategy by increasing capacity and capabilities levels to a point where Prospectors increasingly perform as Defenders in our earlier simulation experiment, experiencing a lower number of discontinuities due to decreased levels of cumulative stress. The effect of organizational learning is hypothesized as follows, based on an the organization’s strategy:

**H3:** Increases in capacities and capabilities of products will result in fewer discontinuous enhancements to the system for all strategies when compared to Experiment 1.

### 3.4 Experiment 4: An Integrated Simulation Model

Finally, a set of models were tested that incorporates each of the variables discussed above. Strategy provides the foundation for the behavior of scope and reach, as well as investments in capacities and capabilities. The expected result is a balanced model where the increase in discontinuities resulting from scope and reach will be offset to a degree as investments are made to capacity and capability. The result is a decreasing number of discontinuities. The expected behavior is predicted by the following hypotheses:

**H4:** Increases in capacity and capability from learning offsets the increases in product scope and reach, causing the number of discontinuities to be equal to or less than those seen in Experiment 1.

### 4. Simulation Results

Figure 2 provides an example of how the three strategies behaved in Experiment 1 accumulating inertia (black line), and stress (gray line), over a ten year period, or half of the simulated time. With each discontinuous change inertia and stress are reset. The level of inertia, as seen in the Figure, is higher for the Defender and lowest for the Prospector. For the Defender, stress builds slowly and the organization operates under stress for longer periods of time before reaching the threshold for change.
Table 1 contains the number of status quo, incremental, and discontinuities experienced for each strategy in the four experiments.

In comparing the three strategies, Defender has the fewest discontinuities and the highest level of inertia (see Figure 2). The Prospector has a much higher number of discontinuities (80) than the Defender or Analyzer and the lowest level of inertia. The findings support Hypothesis 1.

For Experiment 2, the addition of scope and reach to the first set of experimental models indicates an increase in discontinuities for all three strategies indicating support for Hypothesis 2. For example, the Defender experienced three discontinuities versus the two seen in Experiment 1.

The effect of learning exhibits a decrease in number of discontinuities over time for each strategy. Thus, Hypothesis 3 holds true for the third experiment. There is also a related trend of increasing inertia, due to higher levels of capabilities to handle change (Figure 1).

The final simulation experiment examines the full model behavior. Each model had fewer discontinuities as predicted by Hypothesis 4. Inertia in each run is higher than previous models, indicating that a slight interaction effect between scope and reach variables and the capacity and capability variables may exist.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Strategy</th>
<th>Status Quo Changes</th>
<th>Incremental Changes</th>
<th>Discontinuous Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td>Defender</td>
<td>108</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Analyzer</td>
<td>151</td>
<td>13</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Prospector</td>
<td>208</td>
<td>15</td>
<td>80</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>Defender</td>
<td>115</td>
<td>18</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Analyzer</td>
<td>203</td>
<td>44</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Prospector</td>
<td>222</td>
<td>22</td>
<td>127</td>
</tr>
<tr>
<td>Experiment 3</td>
<td>Defender</td>
<td>90</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Analyzer</td>
<td>166</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Prospector</td>
<td>269</td>
<td>13</td>
<td>26</td>
</tr>
<tr>
<td>Experiment 4</td>
<td>Defender</td>
<td>91</td>
<td>52</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Analyzer</td>
<td>240</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Prospector</td>
<td>326</td>
<td>18</td>
<td>27</td>
</tr>
</tbody>
</table>

Table 1: Change by Strategy for Each Experiment

5. DISCUSSION

The behavior observed in the simulations corresponds with the behavior observed in the historical case studies referenced above. The Defender in the final model had one discontinuous change in 20 years, which is closely approximated by the Defender strategy in our simulations. The Prospectors and Analyzers in this simulation study also mirrored the behavior of the organizations observed in that study.

The simulations themselves behaved as intended. The first set of models set the baseline by which to compare the other models. The addition of scope and reach creates more discontinuities due to the increase of magnitude and instability associated with the changes, thus all strategies behave more as Prospectors as time increases due to the increase in discontinuities. The addition of organizational learning as a stress reducer creates a set of models where strategies behave more and more like Defenders over time with fewer discontinuities. Finally, the last set of models exhibits the expected cancellation of increasing inertia and stress from the effect of scope and reach due to the ability of an organization to handle more inertia through learning. For Experiment 4, the simulation behavior is dictated primarily by strategy in the same manner as the first experiment. The findings for this set of experiments has illustrated
that the behaviors predicted by the research model can be simulated. The findings provide the foundation for this ongoing research project. The findings contribute to the MIS field by providing a set of simulations that model IS strategy, an area of research that is of continuing interest to the field (Sabherwal, et al. 2001).

6. LIMITATIONS AND CONCLUSION

The experiments conducted in this research were designed to model and operationalize the process model proposed above. The study examined the role of strategy in the adoption of changes to IS within the organization in response to changes in product scope and reach in addition to increases in capacity and capability. The simulations successfully illustrated the behavior predicted by the model.

There are associated limitations with the study. The first limitation is in using simulations to approximate organizational behavior. While the simulations are complex, they do not account for all possible variables seen in the adoption of IS change. Yet, limiting the number of variables in the model is necessary to make the simulation models parsimonious and testable. Secondly, sensitivity analysis is not included here. Sensitivity analysis is necessary to determine the robustness of the models with regard to extreme cases of behavior. Archival industry data can be used to set parameters and compare behavior across an industry as a step toward validating the process model. Once behavior is consistent within one industry, the model can be tested and the predictive behavior can be assessed. Finally, testing the scope of the model across industries will be an important step in examining the explanatory value of the process model.

This study has taken steps to create simulations that approximate the behaviors predicted by a process model. The next step is to integrate this research with historical data to further test the validity of the model. The final result will provide insight into how strategy, knowledge capabilities, and IT capacities can improve organizational performance. The success of this study will allow the research to proceed to the next step, the calculation of model parameters from a subset of the historical data collected via case studies. Simulation performance can then be compared with the remaining case data to validate the model. Once validated, the simulations will provide a useful tool to model changes in strategy and provide insight into when organizations should adopt a particular strategy.

References


Hannan, M.T., and Freeman, J. “Structural Inertia and Organizational Change,” *American Sociological Review*, (49), 1984, 149-164.


EMPIRICAL COMPARISON OF METHODS FOR INFORMATION SYSTEMS DEVELOPMENT ACCORDING TO SOA

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0393.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Service oriented architecture (SOA), Experiment, Information Systems Development (ISD), Information systems analysis and design</td>
</tr>
</tbody>
</table>
EMPIRICAL COMPARISON OF METHODS FOR INFORMATION SYSTEMS DEVELOPMENT ACCORDING TO SOA

Offermann, Philipp, Deutsche Telekom Laboratories, Ernst-Reuter-Platz 7, 10587 Berlin, Germany, philipp.offermann@telekom.de

Bub, Udo, Deutsche Telekom Laboratories, Ernst-Reuter-Platz 7, 10587 Berlin, Germany, udo.bub@telekom.de

Abstract

While service-oriented architecture (SOA) as an architectural principle for information systems is gaining momentum in research and industry, the field of methods for designing information systems according to SOA is still poorly developed. However, the implementation of SOA design principles, e.g. service reusability, business alignment and autonomy, demands a methodical approach. In order to overcome the shortcomings of current methods, we have developed the SOA method (SOAM). The method is based on existing methods for SOA. Activities are specified along with roles, techniques, modelling notations and a meta-model. A tool supports all necessary modelling notations as well as the generation of XSD, WSDL and WS-BPEL from the models. The newly developed method has been compared to other methods using a laboratory experiment with students. Different methods have been used on different company scenarios; the results have been recorded using a questionnaire. Results show that according to the evaluated criteria, SOAM together with IBM’s method gets the best scores. With respect to the alignment of the software architecture with business processes, one of the primary goals of SOA, SOAM received a better rating than IBM’s method.

Keywords: Method construction, service-oriented architecture, SOA, design science, software engineering
1 INTRODUCTION

The service-oriented architecture (SOA) provides concepts for designing and implementing information systems. Its widespread adoption has been made possible by generally accepted technical standards such as WSDL (Web Service Description Language), SOAP (formerly Simple Object Access Protocol, now proper name), UDDI (Universal Description, Discovery and Integration) and WS-BPEL (Business Process Execution Language). While technical standards have reached an acceptable level and agreement on SOA design principles is growing, discussions on how the software architecture according to SOA is to be defined and how the architecture can be designed continue. When planning to introduce an SOA, a methodical approach is of great importance if business-aligned, reusable services are to be designed and implemented.

We have developed the SOA method (SOAM) based on existing methods for SOA. The method should overcome weaknesses in existing methods; its applicability should be proven in practice. The method that we developed specifies activities, necessary roles and modelling notations and is supported by a tool. We have evaluated the method in different companies using action research. It was possible to show that the method is highly usable and facilitates the design of an SOA system that adheres to the SOA design principles. In this article, the results of a laboratory experiment comparing SOAM to existing methods are presented.

First, service-oriented architecture is introduced. This is the foundation on which methods for SOA are based. Then the weaknesses of existing methods are presented. Our SOA method is explained in detail in this article. Finally, the results of the laboratory experiment are described.

2 SERVICE-ORIENTED ARCHITECTURE

Service-oriented architecture (SOA) combines elements of software architecture and enterprise architecture. It is based on the interaction with autonomous and interoperable services that offer reusable business functionality via standardised interfaces. Services can exist on all layers of an application system (business process, presentation, business logic, data management). They may be composed of services from lower layers, wrap parts of legacy application systems or be implemented from scratch.

The underlying principle for SOA is the service. A service is a software component that can be accessed using commonly known communication technologies. In most cases, Web service technologies are used to implement SOA software. Service types can be deduced from application systems layers:

- Business process service: a service that orchestrates other services according to a business process. Implemented e.g. by using WS-BPEL.
- Presentation service: a service that provides a user interface to perform a user interaction. Data processing is not done automatically, but by the user.
- Business logic service: business logic like calculations, data verifications, transformations etc. that make business sense. Business process activities on the finest level of granularity are usually supported by this type of service.
- Data management service: data management for business entities (data object types). Usually, operations like create, read, update and delete (CRUD) are offered.

The aim of SOA is to create reusable, flexible and business-aligned IT systems. Especially the alignment of software systems with business processes is a step that has not been achieved yet. In order to facilitate the alignment, methods have to take into account business requirements for the design of technical artefacts (Erl 2007). Legner and Heutschi (2007) have summarised nine
publications on design principles. They identified four classes of design principles: interface orientation, interoperability, autonomy/modularity and business suitability.

3 EXISTING METHODS

Methods (sometimes also called methodologies) describe a way to transform an initial state into a target state. (Cronholm and Ågerfalk 1999; Wynekoop and Russo 1997) Activities explain what has to be done. They may be structured hierarchically, e.g. in phases, activities, tasks and steps. The sequential order of the activities is the process model. For each activity, executing roles should be specified. Activities produce results, but may also use existing results as inputs. Results may be linked, e.g. lanes in a process model to organisation units on an organisational chart. Techniques support the generation of results and, therefore, are used by activities that need to create such results. Finally, a meta-model for the results can be specified for clarity and consistency.

The construction of SOAM is based on several published methods. All of these methods have some major or minor weaknesses. SOAM is designed to overcome these.

• Erl published a very comprehensive book on SOA (Erl 2005, pp. 366-370). Within the book, a method for SOA system development is described. The method combines a top-down with a bottom-up-approach and positions an agile procedure as a compromise. On a coarse level, the method contains the steps “service-oriented analysis”, “service-oriented design”, “service development”, “service testing”, “service deployment” and “service administration”. Unfortunately, Erl’s method does not sufficiently take into account legacy systems. Legacy systems are only used as a source for requirements, not as elements for system integration. Additionally, roles are not specified consistently.

• The method of Papazoglou and van den Heuvel (Papazoglou and Heuvel 2006) encompasses the phases from planning to execution, but only analysis and design are specified in detail. During planning, it is decided if a Greenfield-approach is taken or if analysis should proceed top-down, bottom-up or out-of-the-middle. Interestingly, processes are not used for analysis but are rather treated like services. The main problem of the method is the confusing documentation and missing proves for its practical relevance. There is no integrated example; the application of the method seems difficult. No application in practice has been published yet.

• The method of OASIS (OASIS 2005) has a somewhat different focus compared to the other methods. It only defines the specification of Web services. Analysis of business processes and legacy systems and the identification of services is not part of the method. Therefore, business processes are not represented in IT systems and legacy systems are not integrated.

• The method published by IBM (Arsanjani 2004; Endrei, Ang, Arsanjani, Chua, Comte, Krogdahl, Luo and Newling 2004) is very well specified. It consists of the phases “domain decomposition”, “goal-service model creation”, “subsystem analysis”, “service allocation”, “component analysis”, “structure components and services using patterns” and “technology realization mapping”. As can already be seen from the phases, services are identified from domains and goal and not from business processes. Business processes only play a subordinate role. We are conducting a laboratory experiment to determine which approach gives better results.

• The method of Jones und Morris (Jones and Morris 2005) is very pragmatic. It explicitly starts with a domain analysis and identifies services based on a domain decomposition. Unfortunately, the method ends after the identification of services. Further analysis and modelling is not provided. Therefore, the process of designing SOA software is not completely specified.

• Erradi et al. define a framework for SOA development (Erradi, Anand and Kulkarni 2006). The framework consists of the phases “information elicitation”, “service identification”, “service definition”, “service realization” and “roadmap and blanning”. Unfortunately, the process description is rather confusing. Roles and modelling notations are not specified. A case study is only described very briefly, many aspects remain unclear.
• Marks and Bell describe services identification, analysis and design (Marks and Bell 2006). However, no process model is defined. The description is narrative, mixing different aspects of a method. As the description is not systematic and is lacking a consistent chronological order, it is difficult to evaluate and compare the method.

4 THE SOA-METHOD

We developed the SOA method (SOAM) based on the existing methods and their shortcomings. It has a high integrity and consistency regarding the constituent elements and supports the architecture realms “workflow management”, “application architecture” and “enterprise application integration”. It is vendor-independent and explicitly states the architecture goals, which is not the case with any other method. The six phases of SOAM contain all necessary activities, various activities containing several steps. Every activity is specified with executing roles, input and output artefacts and supporting tools. All necessary modelling notations are supported by a tool. The SOAM tool can generate XML schemata (XSD), WSDL files and WS-BPEL process descriptions directly from the graphical models. Action research on the method has been published in (Offermann 2008).

![Figure 1. Phases of SOAM](image)

The sequence of phases can be seen in figure 1. The method uses the top-down approach and the bottom-up approach in parallel. The company requirements are analysed following the top-down approach. Required service operations are discovered based on this. Following the bottom-up approach, legacy systems are identified and analysed regarding data and/or functionality that can be wrapped. Top-down requirements and bottom-up findings are then consolidated. Finally, services are designed and service properties ensured. Processes are prepared for execution. In the remainder of the section the method is explained along these phases.

4.1 Company Analysis

Company analysis focuses on business strategies, business processes and functional domains to identify company requirements. The guideline is that business processes have to be supported by services, but reusability can only be ensured by functional specialisation. All activities in this phase are performed by a business analyst.

In a first, optional step, a business motivation model is created according to the OMG specification (OMG 2006a). Based on the company strategies, business processes required to support the company’s strategies can be identified. Next, business processes or parts of business processes that should be supported by an SOA system have to be identified. Among the criteria for deciding which processes to choose are the frequency of process changes, the current level of integration of legacy systems along the process and the degree to which the process can be predetermined and is based on division of labour (Heutschi 2007).
The processes are modelled down to a level of granularity where each process activity is a single step from the business perspective. A good indicator for this is a clear, non-divisible transformation on a data object. Each task is classified according to whether it can be executed automatically (class “business service”), whether it requires an interaction between a user and the software system (class “user interaction”) or whether the activity is purely manual (class “manual”). The modelling notation supported by the SOAM tool is the Business Process Modeling Notation (BPMN) (OMG 2006b). The data flow should be annotated as far as possible in the process models. Each data object (e.g. “customer”, “order”) is created in a data model (see below). If a domain data model exists in a company, it should be used for annotating the process data flow. Figure 2 shows a business process on the finest granularity level modelled in BPMN.

Parallel to business process modelling, a layered model of functional domains (FD) is created. The top levels of the model can be guided by a functionally organised organisational chart. On lower levels, data objects used in an FD can be used as a guideline to identify refined functional domains. The modelling notation used for the FD model is based on UML Use Case diagrams. Each activity in business process models is linked to a functional domain. If there are more than seven activities assigned to an FD, the FD has to be refined and the activities reassigned. Business process and FD
modelling are iterative. The link between activities and functional domains is later used to ensure the reusability of services. Figure 3 shows a model of functional domains.

4.2 Service Operation Discovery

After having analysed business processes and functional domains, service operations that support the process activities can be discovered. This is done by FD. The activities in this phase are performed by a software architect. For each FD, all activities assigned are analysed per class. For service tasks, service operations that support the activities are modelled in a service model. For user tasks, the user interface (UI) is specified. If two tasks are similar, a service operation that supports both activities can be defined. Each individual service operation and UI is assigned to a functional domain, usually to the same unit as the activity the operation supports. As there is no standardised modelling notation for services, we have developed a notation of our own (see figure 6).

Until now, the service operations and UI discovered only support business process activities. This facilitates process flexibility. Reusable services could also exist on a more technical level. Therefore, service operations and UI use cases are modelled. These are use cases as seen from an operation or UI point of view (e.g. “get clients” and “store order” for an order creation dialog). The use cases are again assigned to FD and classified (e.g. “business logic”, “data management”). The UML Use Case Diagram is used as a modelling notation. The use cases are then ordered by FD and class. If there are similar or equal use cases, a service operation can be defined to support them. When implementing the original service or UI, the service operations derived by the use cases can be called.

4.3 Legacy Systems Analysis

Usually, legacy systems have to be integrated in the SOA software systems. Therefore, in parallel to the company analysis, legacy systems that are used to support the processes are analysed. First, an IT analyst has to identify the relevant systems. Then, these systems are analysed with regard to functionality and/or data that can be accessed by a wrapping service. Possibilities are existing APIs, accessible databases, software components that can be wrapped or even command line invocation. A
modelling notation for legacy systems has been developed to support this activity. Figure 4 shows a legacy systems model.

Every possibility found is then modelled as a service with service operations. Operations are assigned to an FD. Also, used data objects are modelled in the data model. Each object in the data model is assigned to an FD. In the SOAM tool, the UML profile for Core Components (UN/CEFACT 2006a; UN/CEFACT 2006b) is used as a modelling notation.

4.4 Consolidation

After analysing top-down requirements and identifying existing systems bottom-up, consolidation is necessary. It is performed by the software architect and done for data object types and for service operations. For data objects, redundancies have to be identified and merged (e.g. address with country vs. address without country). This can be done by FD, if functional domains have been annotated to data object types. When merging, data object types from legacy systems have precedence because the systems already exist. Only if the discrepancy with top-down requirements is too big can a change in a legacy system be considered. In case two object types cannot be merged, a mediator has to be created. Figure 5 shows the model of data objects types.

![Data object types from Vattenfall Europe](image)

Service operation consolidation is also done by FD and guided by class. Operations that are similar (e.g. “get customer data” and “get customer address”) are merged. Thereby, reusability is assured. Again, operations from legacy systems have precedence. A change in a legacy system can be considered if the discrepancy between requirements and the existing system is too big. When the operations are consolidated, data input and data output are added for each operation based on the consolidated data model.

4.5 Service Design

Once a uniform model of service operations has been established, services can be designed. Usually, a service consists of several operations. Therefore, service operations have to be grouped by service. The software architect uses rules specific to each service class. Operations of class “business logic” are grouped by business logic. Each operation of class “user interaction” is usually represented by a
single service. Data management operations are grouped by business entity. Again, only operations from the same FD are considered for grouping. Therefore, there will only be very few services per FD.

Once services are designed, the design principles that could not be guaranteed during earlier steps have to be enforced. For each service, the software architect has to consider if it is stateless, if rollback operations exist to implement transactional behaviour and if services are loosely coupled. To manipulate the service, the software architect can use operations such as “unification”, “intersection”, “decomposition”, “subset” and “subtraction” (Marks et al. 2006, pp. 111-117). If services become too big during this process, they have to be split up and rerun through the steps above. Figure 6 shows a simple service model. Finally, services are implemented and tested. For these steps, known software engineering methods can be used.

![Figure 6. Services as modelled in the SOAM tool](image)

4.6 Process Preparation

After the services have been defined, the processes have to be prepared for execution by the developers. Service operations have to be assigned to each activity that is supported by a service. Exception handling that has not been modelled before has to be included in the process models. The data flow has to be verified to ensure executability. If necessary, data mediators have to be built. Finally, the tool translates the BPMN models into WS-BPEL. The user interfaces that have been designed earlier have to be included in the process specification. Usually, these steps depend heavily on the middleware infrastructure.

5 EMPIRICAL METHOD COMPARISON

A laboratory experiment was carried out to compare the newly developed SOAM with SOMAM, MSA, SDS, WSIM and SODDM. The comparison was done by using the methods on different company scenarios, rating the experiences of the application using a questionnaire and carrying out a
subsequent statistical analysis of the answers. A company scenario consists of models of a company that are relevant for the application of the methods. The models represent e.g. business processes, organisational charts, data object types and legacy systems. By using predefined company scenarios, it was possible to apply different methods under the same conditions and thereby to retrieve comparable results. This is usually not possible in a company setting. Often, companies are not prepared to provides the resources necessary to apply several methods. Also, because of a possible time shift of method application, general conditions (e.g. legacy systems, staff responsible) might change. On the other side, practical relevance of a laboratory experiment might be limited. A practical application of SOAM is presented in (Offermann 2008).

The first part of the questionnaire is based on the theoretical model for validating information systems design methods by Moody (2003; Moody, Sindre, Brasethvik and Sølvberg 2002). Using a questionnaire, “perceived ease of use”, “perceived usefulness” and “intention to use” can be evaluated. Because all methods evaluated were concerned with SOA, for the usefulness, in addition to general questions, the quality of the realisation of SOA design principles and the achievement of goals of enterprise application system design were evaluated. In the second part of the questionnaire, completeness, consistency and applicability of the method’s components (Greiffenberg 2004) were evaluated.

When designing the questionnaire, 5-point Likert scales were used to simplify the analysis. As the survey was conducted in German, English translations are given here. For the first part of the questionnaire, a scale of accordance strongly agree (5), agree (4), undecided (3), disagree (2), strongly disagree (1) was used. For the second part, a scale of intensity extremely (5), a great deal (4), moderately (3), a little bit (2), not at all (1) was used. The German version of the scales used corresponds to the scales used by Bortz and Döring (2002). They are equidistant as far as possible. The numerical value used for the statistical analysis is given in brackets.

The survey was conducted amongst graduate students of computer sciences and industrial engineering in the time from October 2007 till February 2008 who, alone or in pairs, used different methods on a scenario. According to (King and He 2006), results from students of relevant subjects are comparable to results from professionals. Because the documentations were sufficiently well specified, the methods SOAM, SOMAM, MSA, SDS, WSIM and SODDM were used. To reduce falsification due to the properties of a specific scenario, every team used a different scenario. The scenarios used were “order processing of a mail order snowboard store”, “material management and provisioning of an automotive manufacturer”, “conduction of an acquisition event of a consulting company”, “order processing in industry” and “address and bus bar management of a power utility company”. The student were handed all available written documentation for each method. No questions related to methods were answered.

There were 43 usable responses to the questionnaire. The focus of the method comparison was, in accordance with a prior theoretical comparison, on SOAM and SOMAM. These methods were used most often. For analysing the first part of the questionnaire, first of all the validity is checked using factor analysis. The analysis of the questions concerning usability shows that one component can explain 70% of the variance. This is also the only component with an eigenvalue bigger than 1. For the questions concerning usefulness, only 62% of the variance can be explained by one component. By leaving out questions concerning “clearly specified service contracts” and “abstraction from implementation”, the value is increased to 66% so that only one component has an eigenvalue bigger than 1. For the questions concerning the intention of use, one component explains 94% of the variance.

The second part of the questionnaire evaluates the completeness, consistency and applicability of a method components’ specification. The questions relating to “roles” and “techniques and tools” had to be excluded from the analysis as not all methods specify these components and therefore there are only a limited number of answers. By using an inter-item correlation analysis, it can be shown that the answers for completeness, consistency and applicability are highly correlated (see table 1). All
correlations are highly significant and have a value bigger than 0.6. Therefore, the answers to the three parts of the questions are combined. A factor analysis of the combined questions shows that 88% of the variance can be explained by one component. Hence the quality of the method components can be interpreted as one construct.

<table>
<thead>
<tr>
<th>Question</th>
<th>Part 1</th>
<th>Part 2</th>
<th>Correlation</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kom1</td>
<td>Completeness</td>
<td>Consistency</td>
<td>0.696</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Completeness</td>
<td>Applicability</td>
<td>0.725</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Consistency</td>
<td>Applicability</td>
<td>0.760</td>
<td>0.000</td>
</tr>
<tr>
<td>Kom4</td>
<td>Completeness</td>
<td>Consistency</td>
<td>0.728</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Completeness</td>
<td>Applicability</td>
<td>0.655</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Consistency</td>
<td>Applicability</td>
<td>0.801</td>
<td>0.000</td>
</tr>
<tr>
<td>Kom7</td>
<td>Completeness</td>
<td>Consistency</td>
<td>0.836</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Completeness</td>
<td>Applicability</td>
<td>0.762</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Consistency</td>
<td>Applicability</td>
<td>0.748</td>
<td>0.000</td>
</tr>
<tr>
<td>Kom8</td>
<td>Completeness</td>
<td>Consistency</td>
<td>0.885</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Completeness</td>
<td>Applicability</td>
<td>0.884</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Consistency</td>
<td>Applicability</td>
<td>0.927</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 1. Correlation of the method components' completeness, consistency and applicability

Finally, the reliability of the constructs is verified by calculating Cronbach’s Alpha for the four constructs “usability”, “usefulness”, “intention to use” and “quality of method components”. The results are summarised in table 2. All values are bigger than 0.8. This means that more then 80% of the variance is systematic and must be attributed to errors in measurement.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Cronbach’s Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>0.889</td>
</tr>
<tr>
<td>Usefulness</td>
<td>0.947</td>
</tr>
<tr>
<td>Intention to use</td>
<td>0.939</td>
</tr>
<tr>
<td>Quality of method components</td>
<td>0.956</td>
</tr>
</tbody>
</table>

Table 2. Reliability of constructs

Table 3 contains an overview of the results of the method comparison. In the categories usefulness, intention to use and quality of method components, SOAM received the best ranking, followed by SOMAM. For usability, MSA scored best.

<table>
<thead>
<tr>
<th>Construct</th>
<th>SOAM</th>
<th>SOMAM</th>
<th>MSA</th>
<th>SDS</th>
<th>WSIM</th>
<th>SODDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>3.84</td>
<td>3.60</td>
<td>3.88</td>
<td>3.17</td>
<td>2.67</td>
<td>1.65</td>
</tr>
<tr>
<td>Usefulness</td>
<td>4.17</td>
<td>3.85</td>
<td>3.91</td>
<td>3.47</td>
<td>3.07</td>
<td>1.64</td>
</tr>
<tr>
<td>Intention to use</td>
<td>4.27</td>
<td>4.05</td>
<td>3.30</td>
<td>3.50</td>
<td>2.25</td>
<td>1.00</td>
</tr>
<tr>
<td>Quality of method components</td>
<td>3.99</td>
<td>3.69</td>
<td>3.33</td>
<td>3.30</td>
<td>2.94</td>
<td>1.48</td>
</tr>
</tbody>
</table>

Table 3. Summary of results of method comparison (Likert scale 1-5)

By using a t-test against the value 3, the significance of a deviation from a neutral ranking was determined (table 4). Values smaller than 0.05 are counted as being significant. “+” means that the value is significantly bigger than 3. “-” means that the value is significantly smaller than 3. “o” means that no decision can be made. The results for SOAM and SOMAM are significantly bigger than 3. For
SODDM, the results are significantly smaller than 3. The significance of the value for the intention to use of SODDM could not be calculated because the standard deviation was 0.

<table>
<thead>
<tr>
<th>Construct</th>
<th>SOAM</th>
<th>SOMAM</th>
<th>MSA</th>
<th>SDS</th>
<th>WSIM</th>
<th>SODDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>+ (0.002)</td>
<td>+ (0.009)</td>
<td>+ (0.006)</td>
<td>o (0.510)</td>
<td>o (0.374)</td>
<td>- (0.009)</td>
</tr>
<tr>
<td>Usefulness</td>
<td>+ (0.000)</td>
<td>+ (0.000)</td>
<td>o (0.589)</td>
<td>o (0.060)</td>
<td>o (0.793)</td>
<td>- (0.019)</td>
</tr>
<tr>
<td>Intention to use</td>
<td>+ (0.000)</td>
<td>+ (0.012)</td>
<td>o (0.374)</td>
<td>o (0.345)</td>
<td>o (0.076)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Quality of method components</td>
<td>+ (0.000)</td>
<td>+ (0.000)</td>
<td>o (0.139)</td>
<td>+ (0.045)</td>
<td>o (0.842)</td>
<td>- (0.012)</td>
</tr>
</tbody>
</table>

Table 4. Significance of deviation of Results from median value 3

A t-test on mean equality between SOAM and SOMAM shows that no significant decision can be taken if the values are actually different (table 5). Therefore, it is not clear which of the two methods is better.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Equal variances</th>
<th>Unequal variances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>0.390</td>
<td>0.387</td>
</tr>
<tr>
<td>Usefulness</td>
<td>0.087</td>
<td>0.085</td>
</tr>
<tr>
<td>Intention to use</td>
<td>0.551</td>
<td>0.565</td>
</tr>
<tr>
<td>Quality of method components</td>
<td>0.086</td>
<td>0.083</td>
</tr>
</tbody>
</table>

Table 5. Significances of t-test on mean equality between SOAM and SOMAM

One of the architecture goals of enterprise application systems is the alignment of IT functions to business processes (Heutschi, (Heutschi 2007)). Because of the similarity of the results between SOAM and SOMAM, the answers to “The method ensures a good representation of processes in the software architecture.” were analysed separately. For SOAM, the mean value was 4.09 (agree) with a high significance (0.000) of being more than 3 (t-test on neutral value). For SOMAM, the mean value was 3.40 with a significance of 0.269. Therefore, for SOMAM it can’t be excluded that the real ranking is “undecided”. This is probably due to the fact that SOMAM doesn’t analyse business processes directly. It can be deduced that by using SOAM, business processes are better represented in the resulting SOA-system then by using SOMAM.

6 CONCLUSION

The field of methods for SOA is still poorly developed. In this article, the SOA-method (SOAM) is presented. It aims at overcoming the weaknesses of existing methods. SOAM was developed using action research. To compare SOAM to existing methods, a laboratory experiment was performed. Evaluating the results of the experiment, it can be said that the ranking of all methods is relatively close to neutral. It is important to investigate whether there is a general problem in the way methods are being described. The laboratory experiment nonetheless shows significant differences in quality between the methods. SOAM and SOMAM received the best results. MSA has high usability but comparatively low usefulness. The goal to align IT functions, especially services, to business processes, can be better reached using SOAM rather than SOMAM.

References


<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0612.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>Compatibility, Decision support systems, Service oriented architecture (SOA), Standards</td>
</tr>
</tbody>
</table>
SERVICE-ORIENTED ARCHITECTURES: MODELING THE SELECTION OF SERVICES AND PLATFORMS

Widjaja, Thomas, Technische Universität Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany, widjaja@is.tu-darmstadt.de

Buxmann, Peter, Technische Universität Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany, buxmann@is.tu-darmstadt.de

Abstract

The Service-Oriented Architecture (SOA) paradigm promises to facilitate the integration of software services provided by different vendors and thus enables users to benefit from Best-of-Breed solutions. In order to support software architects we present the Multilayer Standardization Problem (MSP) to analyze the trade-off between possibly enhanced utility versus higher assembling costs of Best-of-Breed SOA solutions. We implemented a software prototype to support decision makers during the data input and the subsequent analysis of the solution’s robustness. The MSP for the SOA-case is formulated as a linear 0–1 optimization model and extends the established Standardization Problem (SP) by modeling the user preferences and considering varying granularity as well as integration relationships in addition to communication relationships. These characteristics are common to numerous systems – thus the general MSP can serve as a basis for further research in this field.

Keywords: Compatibility, Decision Support Systems, Service-Oriented Architecture (SOA), Standards
1 INTRODUCTION

The basic idea behind the SoA paradigm is the support of business processes by IT systems consisting of services. Those services are clearly encapsulated, and loosely coupled entities, which deliver a defined business functionality. (Erl 2006, p. 290ff; Papazoglou & van den Heuvel 2007, p. 389). Current literature accredits SOA-based software systems various benefits compared to traditional monolithic systems such as enhanced agility, straightforward integration of heterogeneous IT environments, etc. (Krafzig et al. 2006, p. 251 ff). However, in order to leverage those benefits, new or adapted methods and tools are needed to support decision makers.

The technical realization of the SOA paradigm is the focus of various research efforts, see e.g. (Erl 2006; Krafzig et al. 2006). The SoA concept itself is technology-independent; however SoA is mostly seen in direct correlation with the Web-Service technology and associated standards. This technical standardization significantly decreases the costs of integrating software services implemented by different vendors, thus making the realization of Best-of-Breed software systems more feasible. (Erl 2006, p. 63) This paper focuses on the trade-off between possibly enhanced utility versus higher assembling costs of Best-of-Breed SOA solutions, i.e. solutions in which the SOA system is composed of components that are provided by different (and often specialized) vendors.

Software systems based on the SOA paradigm can be structured in two layers: Based on an integration platform (layer 0), loosely coupled services (layer 1) are combined to support business processes (see Figure 1). In this paper the term integration platform is used in a broad sense to refer to all infrastructure components of an SOA-based software system, e.g. enterprise service bus (ESB), service repository, application server, and process server. See (Krafzig et al. 2006) for a more detailed description of the integration platform. This abstract and two layered perspective is a simplification but it is possible to extend the proposed model in order to consider the SOA platform in greater detail (see chapter 5 for further research).

![Figure 1. Two layers of a SOA-based software system.](image)

In the following, we will focus our analysis on the market for Enterprise Resource Planning (ERP) software. In this market, some software vendors offer services and platforms, e.g. IBM (WebSphere), Oracle (Fusion) and SAP (NetWeaver), while others provide only a platform, e.g. Red Hat (JBoss Enterprise SOA Platform), or only services (e.g. a specialized demand forecasting service in the context of inventory management). The discussed decision problem is related to a broad literature stream analyzing the process of selecting (web) services with differing Quality of Service attributes in order to maximize user satisfaction – see e.g. (Zeng et al. 2004) for a middleware platform that addresses related issues; (Canfora et al. 2005) for a composition approach based on genetic algorithms, (Yu et al. 2007) for a related knapsack and optimal path problem; and (Blau et al. 2009) for a mechanism design approach. However, to our knowledge no other optimization model yet considers the selection of integration platforms and services and explicitly analyzes the trade-off between the possibly enhanced utility versus the higher assembling costs of Best-of-Breed SOA solutions. Such a model that investigates the decision problem from the standardization perspective could be useful for several reasons. First, it supports software architects during the selection of the services and platforms. “Intuitive” approaches can result in suboptimal solutions: An example is the selection of the set of services and the platform which provides the highest “net utility”, i.e. the highest difference between utility and necessary implementation costs (i.e. neglecting information and integration cost). Second,
the presented model can support the software architect in communicating decisions (especially in cases in which it is optimal to choose not the service preferred by the users). Third, such a model provides a first step towards a theoretical basis for the analysis of Best-of-Breed solutions in SOA-based software systems. Furthermore, the presented model can be applied to various domains beyond the SOA-case. Thus, general insights regarding the described decision problem may be even more valuable than the actual application of the model in a concrete situation. In this paper we present the Multilayer Standardization Problem (MSP) in order to support software architects determining the optimal set of services (layer 1) and platforms (layer 0). The MSP considers the following three cost categories: costs for services and platforms, costs of integrating services mutually, and costs of integrating services and platforms. We will keep the model as simple as possible in order to focus on the trade-off between possibly enhanced utility of a Best-of-Breed solution versus the higher assembly cost. Note that, the proposed model examines the described decision problem from a standardization perspective and therefore does not consider all factors that can influence the selection of services and platforms – e.g. the vendor choice is partially driven by strategic considerations and various difficult to quantify “soft factors”.

This paper is divided into five sections. Following this introduction, we will describe the MSP based on the Standardization Problem (SP). In the third section, we will present a general schema to apply the MSP to the SOA-case and a software prototype to support decision makers. Section four discusses the limitations of the presented model, and the paper closes with conclusions and possible directions for further research.

2 A MODEL TO SUPPORT THE SELECTION OF SERVICES AND PLATFORMS

In two following subsections the modeled aspects which cause the assembly cost advantage of “single-vendor” solutions are discussed. In the third subsection it is discussed how varying user preferences can result in a trade-off between the possibly enhanced utility of a Best-of-Breed solution versus the higher assembly cost. In the last subsection a possible formulation of the MSP as linear 0–1 optimization problem is presented.

2.1 Common standards facilitate the communication between services

In this section, we examine the service layer of a SOA-based software system in more detail. We focus our analysis in particular on aspects that promote a homogenous service landscape. The basic underlying assumption is that common communication standards implemented by the communication partners facilitate information exchange. There are various examples of situations in which communication standards simplify the exchange of information: a common spoken language during a discussion between two individuals, a common file format for two different word processors which exchange data, a semantic definition of business terms during the exchange of XML files between web services. We assume that all services provided by a certain vendor implement the same communication standard. Thus, the exchange of information between software services provided by the same vendor is facilitated, whereas it is not in the case of cross-vendor communication. Note that throughout this article, the term “communication standard” is used to refer to technical (e.g. data formats) as well as semantic aspects (e.g. meaning of business terms).

The economic analysis of standardization decisions is mainly shaped by the network effect theory. Network effects occur if the utility that a user derives from a good depends on the number of other users of the same kind of good. (Farrell & Saloner 1985; Katz & Shapiro 1985) Communication standards show positive network effects: e.g. the more people who speak a certain language (thus the more possible communication partners exist), the more utility this language provides to the individual “adopters”. In the SOA-case the following example is illustrative: The more information a certain functionality exchanges with (already implemented) services that are provided by a specific vendor, the more beneficial the use of a service provided by this specific vendor for the analyzed functionality is. A broad stream of literature analyzes the implications for users and vendors of network effect goods. Topics on which these works focus include start-up-phenomena (see e.g. (Oren & Smith 1981;
Rohlf's 1974)), path dependencies (see e.g. (Arthur 1989; David 1985; Liebowitz & Margolis 1995)), and "tippy networks" (see e.g. (Arthur 1989; Besen & Farrell 1994; Katz & Shapiro 1994)). Based on these insights, various strategies for users and vendors of network effect goods have been derived, see e.g. (Buxmann 2001) for selected results on the software industry. Most of the models regarding network effect theory are based on an aggregated perspective and thus abstracting from individual relations between the different actors. In contrast the Standardization Problem (SP) (Buxmann et al. 1999) explicitly considers these relationships and thus seems to be a sound foundation for the examination of the service layer. The SP considers the following trade-off: On the one hand, the adoption of communication standards (languages, file formats, or XML formats) facilitates information exchange. On the other hand, the implementation of communication standards often involves costs. The SP answers the resulting question: Which actor should implement which standard to maximize the aggregated savings through standardization with respect to the costs? Various research has been performed related to the SP; see e.g. (Buxmann et al. 1999; Miklitz & Buxmann 2007; Schade & Buxmann 2005; Weitzel et al. 2006; Westarp et al. 2000). One major focus of the research related to the SP is how the degree of autonomy of the different actors affects the structure of the optimal solutions; see e.g. (Buxmann et al. 1999; Weitzel et al. 2006). In the case of the centralized SP a superior instance determines – often based on perfect information – the set of standards each actor should implement. In contrast to that, in the case of the decentralized SP, every actor (or a coalition of actors), decides autonomously whether to implement a certain standard or not. Due to the analysis of the software architects perspective, i.e. a “central” decision maker, we use the centralized SP as a basis for our research. Furthermore several studies regarding efficient solution techniques of the SP have been performed. (Domschke & Wagner 2005; Kimms 2003)

Various aspects of the SP can be directly applied to the problem of selecting services and platforms for a SOA. Parallel to the SP the foundation of the MSP is a graph: Based on the business processes a function graph \( G^{\text{Comp}} := \left( N_{\text{Service}}, E^{\text{com}} \right) \) with \( E^{\text{com}} \subseteq \{(i, j)|i, j \in N_{\text{Service}} \text{ and } i < j\} \) is defined. The vertices of this graph \( i \in N_{\text{Service}} \) represent (IT-) functions that support the business processes. A software architect will typically use the proposed model and the respective Decision Support System (DSS) to analyze coarse-grained services and high-level business processes. Nevertheless, the level of abstraction can vary: A vertex can represent a complex function, e.g. a set of CRM functions or also a elementary function, e.g. a simple credit assessment. Based on the analyzed business process, information exchange between functions \( i \) and \( j \) may be necessary. This is modeled by the edge \((i, j)\) in the function graph. Note that not every function necessarily exchanges information with all other functions, i.e. this graph does not need to be complete. We assume that each function \( i \in N_{\text{Service}} \) can be fulfilled by services \( r \in R_{\text{Service}} \) provided by various vendors. Some vendors provide a wide variety of services (e.g. SAP and Oracle) but other (more specialized) vendors offer only services for selected functionality. Some services fulfill exactly one function (e.g. a service for the credit assessment) while other, more coarse granular services can fulfill several of the demanded functions (e.g. the core banking module of a large ERP system vendor). The parameter \( a'_i \) takes the value 1, if and only if service \( r \) can fulfill function \( i \), otherwise the value is 0. The set \( R_{\text{Service}} := \{r \in R_{\text{Service}}|a'_i = 1\} \) contains all services that can fulfill the function \( i \). The acquisition or implementation of a service \( r \) incurs costs, e.g. additional hardware, software licenses, and training. We model these costs in the parameter \( a'_i \) and refer to these costs as implementation costs. Already existing services, i.e. an installed base, can be modeled by \( a'_i = 0 \) and services that are realized as facades for legacy systems typically incur lower implementation costs than services provided by external vendors. We assume that every function has to be implemented by exactly one service to derive a valid SOA system (see Miklitz & Buxmann 2007) for a similar assumption – they propose an extension of the SP to support decision makers during the selection of IT systems in the context of Mergers & Acquisitions).

A basic assumption of the MSP is that the information exchange between services is realized via communication standards. Each service \( s \in R_{\text{Service}} \) is modeled as a bundle of a communication standard (i.e. the associated semantic and syntactic standards) \( k' \in K \) and a software product that provides certain (business) functionality. Although the SOA paradigm can be realized by several technologies, current discussion focuses on the technology of web services, see e.g. (Alonso et al. 2004; Newcomer & Lomow 2005). Thus, it is often assumed that services based on this technology can be combined
easily. We argue that this perspective is overly simplified, since the standardization to the technology of web services in particular affects the syntactical compatibility of services. However, the composition of SOA-based systems by simply plugging together existing building blocks is often inhibited by semantic differences between the services. Furthermore, various vendors extend the existing web service standards by proprietary increments. In our opinion, it is necessary to model the communication of services as point-to-point communication until a technical and semantic standard (in the respective company) exists. Therefore we assume that each vendor equips all offered services \( r \in R_{\text{Service}} \) with a certain communication standard \( k' \in K' \). Especially “big” software vendors attempt to establish de-facto standards, and niche developers are often aligned to such a standard. Therefore, “standard-ecosystems” evolve around the standards of big players. (see for related aspects (Petrie & Bussler 2008)) Figure 2 shows a simple function graph: services \( s_1 \) and \( s_2 \) are offered by the same vendor \( B \), i.e. \( k'^i = k'^3 = B \), which is indicated by the same shading. Furthermore, \( \alpha'^i = \alpha'^3 = 1 \), since \( s_2 \) supports function 1 as well as function 2.

![Figure 2 The business process as basis for the function graph.](image)

We pointed out that communication between services provided by the same vendor is “facilitated”. In order to consider this phenomenon we model information cost for each pair \((s,t)\) of service candidates \( s \in R_{\text{Service}} \) and \( t \in R_{\text{Service}} \) that can fulfill functions \( i \) and \( j \) with \((i,j) \in E_{\text{com}}\). Information costs model the cost for the information exchange, e.g. costs that are associated with pre- or post-processing of data, which is partially caused by incompatibilities between the communication standards. For a possible way to evaluate information cost and a more detailed description of the concept, i.e. the incorporation of possibly higher value of the exchanged information, see e.g. (Buxmann et al. 1999). The information costs are modeled by the parameter \( C_{ij}^p((i,j) \in E_{\text{com}} \) and \( s \in R_{\text{Service}} \) and \( t \in R_{\text{Service}} \) (see the right side of figure 2). We assume that, \( C_{ij}^p \leq C_{ij}^s \) if \( k'^i = k'^j \) and \( k'^i \neq k'^j \), i.e. information cost for the communication between services provided by the same vendor are at most equal to information cost for the information exchange between services offered by different vendors. The justification for this assumption is analogous to the argumentation of (Buxmann et al. 1999) for the SP: Technical and semantic incompatibilities between the communication standards cause additional information cost. This causes a network effect for communication standards of the services: As more functions are implemented by services provided by a distinct vendor (i.e. fulfill a distinct communication standard), the more beneficial (from the information cost perspective) the usage of services from this specific vendor (i.e. this communication standard) for the remaining functions with which information is exchanged becomes. Note that, the strength of the network effects depends on the availability, comprehensiveness, and diffusion of domain specific (open) communication standards. If a (more coarse grained) service \( r \in R_{\text{Service}} \) is able to implement two communicating functions, i.e. \((i,j) \in E_{\text{com}}\) and \( \alpha'i = \alpha'j = 1 \), cost for information exchange can be modeled but typically those costs will be close to zero.

The function graph of the MSP is not directed, though the chronological and logical structure of business process implies a directed graph. By aggregating the mutual communication cost between the service candidates of two functions \( i \) and \( j \) into one parameter \( C_{ij}^d \), the directed graph can be simplified; this is analogous to the corresponding modification in the classical SP. (Domschke & Wagner 2005)
2.2 Interplay of services and standards: The layer structure

The classical SP focuses on communication relationships among actors and examines the trade-off between standardization costs and possible savings of information cost between communication partners by standardization. The MSP extends the SP by considering a second type of relationship: the integration relationship. This extension is necessary to analyze the described decision problem in the SOA context because the costs to use a service (layer 1) depend to a significant degree on the integration platform (layer 0) the service is supposed to run on (see Figure 1). Since many platform vendors “enrich” open integration standards by various proprietary extensions, services provided by a given vendor typically operate better on that vendor’s platform than on any other platform. For example, the integration of a SAP service with the NetWeaver platform is often easier than with any other platform. These extensions allow for example better performance or ensure a higher security level. Typically, the services that are offered by the platform vendor take advantage of these extensions without any modification. Vendors that only provide services – but no own platform – often “align” (and certify) their services for use with a specific platform. Similar problems are discussed in the literature related to “Mix and Match” and the “Hardware Software Paradigm” (see e.g. Economides 1996; Einhorn 1992; Katz & Shapiro 1994; Matutes & Regibeau 1988). To model these aspects in the SOA-case, two layers are necessary: Layer 1 consists of functions \( i \in N_{\text{Service}} \) that are fulfilled by services provided by possibly different vendors in order to support a business process. On layer 0 – the platform layer – one abstract function \( n_p \) (i.e. the integration platform) is modeled (see Figure 3). Thus, the graph of the MSP for the SOA-case consists of one service layer with \( n \) vertices (layer 1) and one platform layer (layer 0) with one vertex \( n_p \). As already described, the term “integration platform” relates to all infrastructure components that are necessary to support the services. The set \( N \) of functions is defined as \( N := N_{\text{Service}} \cup N_{\text{Platform}} \) with \( N_{\text{Platform}} := \{ n_p \} \). In the following we subsume services and platforms as artifacts. The set of artifacts is defined as \( R := R_{\text{Service}} \cup R_{\text{Platform}} \). Analogous to \( R_{\text{Service}} := \{ r \in R_{\text{Service}} | \alpha' = 1 \} \) the sets \( R_{\text{Platform}} := \{ r \in R_{\text{Platform}} | \alpha'_p = 1 \} \) and \( R := \{ r \in R | \alpha'_p = 1 \} \) are defined. Analogous to services the implementation of a platform \( p \) incurs implementation cost \( d^p \). In a first step, we assume that platforms fulfill no other functions except \( n_p \), i.e. \( \alpha'_p = 0 \forall i \in N_{\text{Service}} \) and \( r \in R_{\text{Platform}} \). Furthermore, we assume that a valid solution implements exactly one platform – federated system landscapes with multiple “platforms” can be modeled in the extended model (see chapter 5).

The MSP considers vertical (in)compatibilities by integration relationships and a compatibility graph weighted with integration costs \( d^p \) (see Figure 3 for an example). In the SOA context integration costs can be interpreted as costs that are necessary to make a service executable on a certain platform (e.g. cost for the deployment and fine adjustment).

![Function graph (layer 0 and layer 1) with communication and integration relationships](image)

![Integration cost for service and platform candidates](image)

Figure 3: Integration relationships weighted with integration cost.

We assume that each service \( s \) has to be integrated with a platform \( p \) to receive a functional SOA system. The integration of a service and a platform is realized via interfaces, which fulfill “integration standards”. Note that besides the communication standards on the service layer, integration standards between the two layers also influence the assembly cost of the SOA system. For the purpose of simplification, we assume that each vendor offers exactly one integration standard for all artifacts (i.e. services and platforms). This allows us to summarize integration and communication standards for the artifact \( r \in R \) in the parameter \( k^r \in K \). Thus \( k^r \in K \) can be interpreted as vendor of artifact \( r \).
range suppliers such as SAP or Oracle offer services besides the platforms to fulfill nearly every functionality – thus in such cases a “full standardization” on both layers is possible. Pairs of services and platforms \((s, p)\) which are provided by the same vendor fulfill \(k^s = k^p\). The integration of such artifacts is often preconfigured, well documented, and supported by specialized tools. For these reasons we assume that \(d^p \leq d^q\) holds, if \(k^s = k^p\) and \(k^q \neq k^p\). If a certain service \(s\) is totally incompatible to a platform \(p\), \(d^p := M\) holds.

2.3 Preferences for services and platforms

The aspects of the decision problem already described – i.e. the possible savings of integration and communication cost by choosing only one vendor to provide both the demanded services and the platform advocate the choice of a “single vendor” solution. However, in some situations, there is no single vendor who offers artifacts for all necessary functions. Another reason to decide on more than one vendor is, that the users have different preferences regarding the services and platforms and not all preferred artifacts will likely be provided by a single vendor. In order to consider the preferences of the users, it is necessary to model the utility of the services and platforms. The preference of the user for artifact \(r \in R\) implemented in function \(i \in N\) is expressed in monetary units in the utility parameter \(b_r\). The monetary estimation of the preference can be supported by a weighted decision-matrix. The user defines a set of weighted functional and non-functional criteria upon which the potential artifacts are scored. An example of possibly relevant criteria for service candidates are maximal data transfer rate, security aspects, and the design of the graphical user interface. Note that the utility \(b_r\) of an artifact is function specific but the implementation cost \(a\) are associated with the artifact itself. This means that the cost \(a\) are incurred if artifact \(r\) fulfills at least one function. Note that the model allows to consider “function-specific” costs by subtracting them from the utility \(b_r\) of the artifact - receiving a “implementation specific net-utility”. For simplification, we assume throughout this article that all costs are “artifact-specific”. This aspect is related to a major difference between communication and integration relationships. Communication relationships exist between implementations of functions by artifacts, whereas integration relationships exist directly between artifacts, i.e. the service-platform pairs \((s, p)\). Furthermore, the integration relationship is directed: a distinct service needs a platform to operate, not vice versa. If the first service is used, at least one platform has to be implemented to host this service and then for subsequent implemented services the usage of this platform is “free of charge”. This points out the infrastructure character of the platform. Since every function \(i \in N = N_{Service} \cup N_{Platform}\) has to be fulfilled by one artifact in the presented version of the model, this infrastructure character is hidden.

(Chou 2007) advertises that in hybrid systems, i.e. systems that are composed of components provided by different vendors, utility losses due to incompatibilities are possible. We observed in the SOA-case that not every service-platform combination \((s, p)\) harmonizes in the same manner. For example, the data transfer rate of a certain service \(s_i\) could be vastly higher than the maximum data transfer rate of a platform \(p_i\). This means that the whole combination of \(s_i\) and \(p_i\) can only operate at the data transfer rate of the platform. Thus the relevant characteristics of the complete system are determined to a certain extent by the composition of the system and not only by the addition of the individual utilities. We propose the following approach: During the assessment of the artifacts, an “ideal” environment should be assumed – thus no utility losses due to incompatible complementary components should be considered. Afterwards the assumed utility losses can be integrated by additional costs added to the parameter \(d^p\).

2.4 Model for the MSP in the SOA-case – A mathematical formulation

We introduce four binary variables \(w^s, x^s, y^p, z^p \in \{0, 1\}\) for the formulation of the MSP as a linear optimization problem of the MSP. Table 1 summarizes the introduced notation.

The objective function maximizes the utility derived by the artifacts implementing the functions with respect to the implementation, information, and integration costs. Note that all terms are expressed in monetary units.
Table 1 Notation of the model for the MSP

The model for the SOA-case MSP is formulated as follows:

Maximize $F(w, x, y, z) = \sum_{i \in N} \sum_{r \in R^i} b^r_i x^r_i - \sum_{i \in N} \sum_{r \in R} \alpha^r_i w^r_i - \sum_{(i, j) \in E^{com}} \sum_{s \in R^i} \sum_{t \in R^j} c_{ij}^{st} y_{ij}^{st} - \sum_{s \in R^i} \sum_{p \in R^i_{platform}} d_{sp}^{op} z_{sp}^{op}$ [1]

s.t.

$x^r_i - w^r_i \leq 0 \quad \forall i \in N$ and $r \in R^i$ [2]

$\sum_{r \in R} x^r_i = 1 \quad \forall i \in N$ [3]

$y_{ij}^{st} - x^r_i \leq 0 \quad (i, j) \in E^{com}$ and $s \in R^i$ and $t \in R^j$ [4]

$y_{ij}^{st} - x^r_i \leq 0 \quad (i, j) \in E^{com}$ and $s \in R^i$ and $t \in R^j$ [5]

$\sum_{s \in R^i} \sum_{t \in R^j} y_{ij}^{st} = 1 \quad \forall (i, j) \in E^{com}$ [6]

$w^r_i - \sum_{p \in R^i_{platform}} z_{sp}^{op} \leq 0 \quad \forall s \in R^i$ [7]

$z_{sp}^{op} - x^r_i \leq 0 \quad \forall s \in R^i$ and $p \in R^i_{platform}$ [8]

$w^r_i, x^r_i, y_{ij}^{st}, z_{sp}^{op} \in \{0, 1\} \quad \forall (i, j) \in E^{com}$ and $i' \in N$ and $p \in R^i_{platform}$ and $r \in R^i$ [9]
The set of constraints [2] guarantees that variable \( w' \) has to take the value 1 if at least one implementation of the artifact \( r \) is part of the solution. Note that \( w' \) can take the value 1 if \( a' = 0 \), even if the artifact is not implemented. Constraint [3] guarantees that each function is implemented by exactly one artifact. This constraint is necessary since all functions have to be fulfilled by a service in order to receive a valid SOA. Note that the constraints [4] and [5] are similar to the formulation of the SP by (Domschke & Wagner 2005). Constraint [6] ensures that each pair of communicating functions incurs information costs. Constraint [7] ensures that each service is hosted on one platform. Constraint [8] guarantees that, if a platform hosts any services, this platform must be implemented. The set of constraints in [9] specifies the binary domains of the variables.

3 THE APPLICATION OF THE MSP IN THE SOA-CASE

Figure 4 shows the general schema of the application of the MSP in the SOA-case. We implemented a Eclipse-based software prototype to support decision makers during data input (e.g. import of BPEL-based process descriptions), to solve the MSP and to analyze the solution’s robustness (figure 5).

![Figure 4: The application of the MSP in the SOA-case. The whole process is supported by the DSS.](image)

The input parameters of the MSP are classified in four groups: Input 1) the communication graph \( G^{com} = (N_{service}, E^{com}) \) on the service layer (layer 1), Input 2) information regarding the artifacts (i.e. implementation cost \( a' \), functions that the artifact can implement \( a'_i \), function-specific utility \( b'_i \), and vendor of the artifact \( k' \)), Input 3) information cost \( c^{ij}_w \) for the relevant pairs of services, and Input 4) integration cost \( d^{ij}_p \) for the service-platform combinations. In the first step, the functions that should be covered by the SOA system must be identified, i.e. the set \( N \). Note that, a finer granular analysis of the functions induces a higher complexity during the data collection and solution process of the resulting MSP instance. Based on the relevant business processes the communication relationships \( (i, j) \in E^{com} \) between the functions have to be assessed. The software prototype allows to import existing BPEL process descriptions (e.g. created in other process modeling tools). In the next step, different service and platform candidates \( R = R_{service} \cup R_{platform} \) must be assigned to the functions in order to determine \( a'_i \). It seems to be useful to consider only those service and platform candidates that are able to fulfill mandatory requirements of the decision makers (especially with regard to security and data transmission aspects). The user has to assess for each artifact the implementation cost \( a' \) and the (expected) utility \( b'_i \). Based on the expected intensity of communication and the degree of the compatibility of the communication standards, the information cost \( c^{ij}_w \) for all pairs of functions \( (i, j) \in E^{com} \) are assessed and finally, the integration costs \( d^{ij}_p \) for each service-platform combination are estimated. Based on these parameters the (under the given assumptions) optimal configured SOA is determined by the open source Mixed Integer Programming (MIP) solver “lp_solve” (version 5.5). (see http://lpsolve.sourceforge.net for details)
The configuration is optimized with respect to the information cost \( c_{ij}^w \), integration cost \( d_{ij}^w \), the user preference \( b_{ij}^r \), and the standardization costs \( a_{ij}^r \) of each artifact and function. The implemented DSS allows to conduct a Monte Carlo analysis with various graphical outputs to visualize the robustness of selected solutions.

4 LIMITATIONS OF THE PROPOSED MODEL

The limitations of the MSP model are similar to the limitations of the classical SP models, which are well studied in the literature. According to e.g. (Schade & Buxmann 2005) three types of limitations of the SP model exist; the data problem, the complexity problem and the implementation problem. The data problem refers to the difficulty collecting accurate data for the necessary parameters. Note that the MSP requires additional parameters compared to the SP. In general, the user must consider to which extent additional costs incurred by data collection (in particular regarding the mentioned “soft factors”) are justifiable by the improvement of data-quality. Especially the determination of parameter \( b_{ij}^r \) is related to the general problem of evaluating an IT-system’s economic value. However, IT-investment-decisions are economic decision-problems and therefore economic methods should be applied. Preliminary experiences from two case studies currently being conducted indicate that it is better to estimate the parameters in workshops with enterprise architects, rather than solely based on the (existing) business case for the investment decision. If detection and assignment of utility in a specific instance of the problem is too expensive, the cost-minimal solution for this SOA system can be achieved by defining \( b_{ij}^r := 0 \forall i \in N \text{ and } r \in R^i \). In this case, the parameter \( d_{ij}^w \) does not incorporate costs for utility losses due to incompatibilities in hybrid systems. The complexity problem arises in situations where bigger instances of the SP are solved, i.e. complexity increases when the number of vertices in the graph or the number of available artifacts increases. (Schade & Buxmann, 2005) By transformation into a warehouse location problem (WLP) it can be shown that the SP with more than two standards is NP-hard. (Domschke & Wagner, 2005) The implementation problem refers to the difficulties that arise when the implementation of a certain configuration is to be enforced.

The aim of the MSP is to analyze the described decision problem from a standardization perspective, knowing that various other relevant perspectives exist which are also of possible interest; e.g. it is observable that some software architects prefer to ensure a certain heterogeneity of the IT architecture to reduce dependency on single software vendors. Although the suggested model is a static one, some parameters can exhibit dynamic aspects, e.g. operating costs. These dynamic rates of the parameters can be derived by estimation and discounting. Nevertheless the presented model doesn’t allow an iterative transition between different IT-architecture landscapes. A further general limitation of the proposed model is the simplified (i.e. deterministic and linear) objective function.
5 CONCLUSIONS AND AVENUES FOR FURTHER RESEARCH

The aim of this article was to analyze the optimization problem regarding the choice of services and platforms in a software system based on the SOA paradigm in order to support decisions of software architects. We address two interdependent questions of the decision makers, which arise during the composition of SOA-based software systems: Which set of services should be chosen, and which platform should host those services? We presented an application of the Multilayer Standardization Problem (MSP) to the two layered SOA-case, its formulation as a linear 0–1 optimization problem and a according Decision Support System (DSS). The MSP extends the Standardization Problem (SP) by incorporating differing user preferences (utility of the artifacts), varying granularity of services and considering integration relationships in addition to communication relationships.

The general MSP is able to analyze more complex systems (n layers), which is useful to model the integration platform in a more detailed way, i.e. to model the integration platform as a system which itself contains communication and integration relationships. Another application of the general MSP is the examination of other domains e.g. a PC system with three layers: layer 0 (hardware), layer 1 (operating system), layer 2 (application software). A further possible direction for future research is the formulation of a dynamic decision model and the investigation of decentralized multilayer standardization problems. It would also be interesting to develop enhanced solution techniques for large problem instances – possible approaches include reformulations to provide tighter upper bounds for pruning in a Branch and Bound algorithm, or the development and use of (meta)heuristics. This article focused on the software architect’s perspective during the composition of SOA-based ERP systems. Future research regarding the analysis of strategies for software vendors’ can use the MSP to anticipate the reaction of users. In this context, strategies regarding compatibility and bundling (service/service as well as service/platform) are interesting topics.

Acknowledgments

We thank the anonymous reviewers for their constructive comments. We also thank Stephan Faßbender, Joachim Kanbach and Florian Weichand for their valuable contributions to the implementation of the presented DSS. Thomas Widjaja’s research is supported by a research grant from the FAZIT-Stiftung.

References


Proceedings ECIS 2009
Design Factors for Service-oriented Architecture Applied to Analytical Information Systems: an Explorative Analysis

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0666.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Intelligence, Service oriented architecture (SOA), Empirical study, Data Warehousing</td>
</tr>
</tbody>
</table>
DESIGN FACTORS FOR SERVICE-ORIENTED ARCHITECTURE
APPLIED TO ANALYTICAL INFORMATION SYSTEMS:
AN EXPLORATIVE ANALYSIS

Dinter, Barbara, University of St.Gallen, Institute of Information Management, Müller-Friedberg-Strasse 8, 9000 St.Gallen, Switzerland, barbara.dinter@unisg.ch
Stroh, Florian, University of St.Gallen, Institute of Information Management, Müller-Friedberg-Strasse 8, 9000 St.Gallen, Switzerland, florian.stroh@unisg.ch

Abstract

Today’s analytical information systems demand innovative architecture concepts in order to address requirements like flexibility and faster time-to-market. Service-oriented architectures (SOA) as a current trend might meet these challenges. So-called BISOA describes the approach of deploying SOA in Business Intelligence systems. We identified two research questions aiming at finding insights about the interdependencies of BISOA and the organization’s system landscape: What are the dominant design factors of BISOA and what are the distinct realization approaches of BISOA? The paper answers the questions by means of empirical research. Besides the factor describing the degree of BISOA realization, three further design factors resulted from the analysis: excellence in embedded Business Intelligence (BI), process orientation, and excellence in data management. Depending on the factors distinct approaches of real-world organizations to deploy SOA in BI systems were derived. These profiles allow us to gain insight which aspects have to be considered besides mainly technical oriented implementation issues for BISOA. Surprisingly, there does not exist such a strong correlation between BISOA and operational BI as often assumed.

Keywords: Business Intelligence (BI), service-oriented architecture (SOA), operational BI, empirical analysis.

1 INTRODUCTION

Service-oriented architecture (SOA) dominates as a recent trend the discussion about today’s IT architectures. It is currently establishing as an accepted and useful concept for information system architectures in organizations (Baskerville et al., 2005; Kraftig et al., 2005; Schelp & Winter, 2007). Organizations are faced with major challenges in order to adapt fast changing organizational structures in a flexible way and to support new or modified business processes. The challenges encompass not only business aspects but also technical issues such as reducing complexity of the application landscape. Service-oriented architectures are aimed at providing support in this context.

On the other side, analytical information systems represent meanwhile an essential component of the enterprise application landscape. Business intelligence (BI) as the underlying concept is used as an umbrella term to describe the processes and systems dedicated to the systematic and purposeful analysis of an organization and its competitive environment. Also analytical information systems demand increasingly innovative architecture concepts in order to address new challenges like realizing operational BI approaches (cf. Section 2.2). Furthermore, organizational units responsible for development and operations of analytical information systems have to assess if the SOA paradigm might apply to analytical information systems as well and have to design a possible integration of
SOA in the current architecture landscape. In the following, we designate the term “BISOA” to the situation that SOA is applied to analytical information systems and BI systems, respectively.

According to a survey of Ventana Research about the status of SOA in BI, only 33% of organizations reported they believe their internal IT personnel have the knowledge and skills to implement BI services (Everett, 2006). However, the decision, if and how to apply BISOA and to develop an appropriate procedure model requires know how about the interdependencies between BISOA and the current system landscape.

We have reviewed both scientific and practitioner oriented publications in order to assess the state-of-the-art of BISOA. Whilst there are numerous articles published by consultancies, research and advisory firms, and software vendors, only little work can be attributed to the scientific community. The paper at hand aims at providing insights to the aforementioned interdependencies, based on the results of an empirical analysis. Moreover, it tends to assist the selection and adaptation of frameworks from information systems research, which may support the introduction of BISOA concepts in organizations. In the past, design research artefacts including models and methods were rather generic and did not differentiate enterprise specific needs. Thus, concepts in design research have evolved which aim at enhancing generic problem solving approaches such as adaptable conceptual models (vom Brocke, 2007; vom Brocke & Buddendick, 2006) or situational method engineering (Becker et al., 2007; Ralyté & Rolland, 2001). By means of these approaches, existing models and methods can be refined to ensure that a model or method fits better to problem-specific requirements than a generic one does. The study at hand investigates relevant design factors, which influence the application of relevant design research artefacts (methods or methods fragments of e.g. a BISOA introduction). In addition, it tends to reveal BISOA profiles these artefacts are applied for. Consequently, the following two research questions will be discussed:

1. What design factors dominate the realization of BISOA?
2. Based on these design factors: Which distinct approaches to BISOA are pursued by real-world organizations?

The remainder of the paper is organized as follows: Section 2 provides an overview of SOA and its deployment in analytical information systems (“BISOA”), respectively. In Section 3 the empirical investigation for identifying design factors as well as BISOA approaches using data collected from a written survey is described. Following, Section 4 discusses an interpretation and discussion of the survey results. Section 5 concludes and provides an outlook on further research.

2 THE CONCEPT OF SOA AND BISOA

2.1 The SOA paradigm

The main idea of SOA is to encapsulate business functionality into small loosely coupled services, which – once implemented on a software level – may be flexibly orchestrated to match the particular and ever changing business requirements (Krafzig et al., 2005). Consequently, rather than providing a technical solution, the SOA concept presents a business driven architecture paradigm. SOA aims at designing an ideal IT architecture, which is not built by monolithic applications but by standardized components (services) that are designed according to underlying business functionalities.

Many expectations concerning various and extensive benefits come along with the service-oriented paradigm, in particular addressing the need for improvement of existing IT architectures. The main benefits of SOA are amongst others:

**Increased reuse rate:** Many business functionalities that are realized in software are reused in various business processes. By means of SOA these software functionalities can be used by several processes
via standardized interfaces. Consequently, redundant development can be avoided (Endrei et al., 2004).

**Reduction of complexity**: SOA facilitates the decoupling of business process control from the technical functionality. The coordination mechanisms are shifted to an integration layer and are therefore separated from the functional components. In addition, the interface standardization allows a reduction of the number of interfaces.

**Increased manageability**: Several applications might simultaneously benefit from further development of services – assumed an appropriate version management has been established. As the service functionalities are encapsulated via interfaces the underlying software might be (ex)changed without requiring adaptations in the corresponding components.

**Faster time to market**: Due to more efficient deployment mechanisms SOA enables the IT departments to react faster to changing business requirements. Consequently, agility and flexibility are enhanced (Bieberstein et al., 2005; Schelp & Winter, 2007).

Evolutionary extensibility, cost savings, investment protection, and outsourcing potentials are considered as further advantages. A detailed discussion about SOA can be found e.g. in (Endrei et al., 2004; Krafzig et al., 2005; Keen et al., 2004).

### 2.2 SOA for BI systems (BISOA)

The aforementioned Ventana Research survey confirmed the relevance of BISOA. More than 81% of respondents said SOA was important to BI due to its powerful combination of business and IT benefits. From the business side, nearly 50% felt that BI services would help make information more broadly available, while also improving the business's ability to respond faster to change. From the IT side, nearly 66% felt that BI services would help IT departments to respond to business needs in a better way, with another 33% viewing ease of integration and lower lifecycle management costs as key benefits (Besemer, 2007). These benefits are similar to the generic SOA advantages as presented in Section 2.1.

BI systems can act either as service provider or as service consumer. In addition, functionalities within BI systems can be realized via services. Nowadays, there is a (more or less) common understanding which services can be differentiated in BI systems / architecture, although classification criteria, terminology, and the service types in details still vary (Besemer, 2007; Gordon et al., 2006; Martin & Nussdorfer, 2007; Wu et al., 2007). The most popular service types are given below:

**Sourcing (Backend) services**: Operational data sources are connected to the data warehouse via dedicated extraction services or via already existing services provided by operational applications (Gordon et al., 2006; Keith et al., 2007).

**Transformation services**: Transformation functionality within ETL (extraction, transformation, loading) processes is realized by means of services. Examples are aggregations, encoding, etc. (Gordon et al., 2006; Wu et al., 2007).

**Analytical (Frontend) services**: BI data and products are made available to end users and other applications via services. An overview of analytical services like reporting, dashboard, alerting services can be found amongst others in (Gordon et al. 2006; Martin 2006; Martin & Nussdorfer 2007).

**Infrastructure services**: Tasks with infrastructure character such as meta data, master data, and data quality functionalities are realized by means of services (Gordon et al., 2006).

In addition to these use cases, which correspond to the typical data warehouse architecture layers, operational BI is often mentioned as an application area for BISOA, e.g. in Besemer (Besemer 2007; Blasum 2006; Eckerson 2007; Keny & Chemburkar 2006). There is not yet a common understanding about the term operational BI. Some authors focus on specific aspects, others like Eckerson (2007),
consider operational BI as an umbrella term: “Operational BI delivers information and insights to a 
broad range of users within hours or minutes for the purpose of managing or optimizing operational or 
time-sensitive business processes”. Eckerson (2007) presents a survey conducted about operational BI. 
Respondents being asked amongst others for main challenges in deploying operational BI, rated 
“Architecting the system” to the top. Moreover, other authors point out the significance of SOA to 
near/real time data delivery (data warehousing) (Abrahem, 2007; Wu et al., 2007).

Considering the various fields of applications for SOA in analytical information systems mentioned 
above and driven by the demand for innovative BI architecture concepts, it seems worth the effort to 
consider BISOA as an alternative or enhancement of established architectures. In the next section we 
will elaborate the predominant design factors and analyze their impact on the realization degree of 
BISOA.

3 EXPLORATIVE ANALYSIS

3.1 Research method

Explorative analysis was used as the underlying research method. By means of a factor analysis we 
identified the predominant design factors (indicators) for BISOA realization. This method is used in 
order to extract a small number of relevant mutually independent factors from a multiplicity of 
variables (of a data set) (Härdle & Simar, 2003). In the field of factor analysis two different 
approaches with different goals can be distinguished: confirmatory factor analysis (CFA) and 
exploratory factor analysis (EFA) (Thompson, 2004). A CFA approach requires “a firm a priori sense, 
based on past evidence and theory, of the number of factors that exist in the data, of which indicators 
are related to which factors, and so forth” (Brown, 2006). In contrast to that, researchers using the 
EFA approach may not have any specific expectations regarding the number of underlying factors. 
The explorative nature of the article at hand implicates therefore the application of EFA.

A cluster analysis of the extracted factors was then performed in order to identify different BISOA 
approaches. Clustering as a combinatorial data analysis technique investigates “a set of objects in 
order to establish whether or not they fall […] into groups […] of objects with the property that 
objects in the same group are similar to one another and different from objects in other groups” 
(Gordon, 1996). Initially, these groups are unknown and need to be determined. Various clustering 
methods for different purposes are available. They can be categorized by the types of algorithm used 
to obtain the clusters. Most important clustering techniques include agglomerative, divisive, 
incremental, direct optimization, and parallel algorithms (Gordon, 1996). According to Härdle & 
Simar (2003) agglomerative algorithms have the largest significance in practice. Thus the clustering 
method of the paper at hand is based on such an algorithm. Starting with n clusters, each containing a 
single object, an agglomerative algorithm reduces the number of clusters by merging the two most 
similar ones at each step. In our context, the cluster analysis aimed at determining the correlations 
between the indicators (reflecting the design factors) and BISOA approaches in organizations and 
consequently, at evaluating the relevance of these indicators.

3.2 Data collection and selection

Data for the empirical analysis was collected by means of a written survey that was conducted at a 
practitioner conference on data warehousing and business intelligence held in Switzerland. The 
conference was attended by 137 specialists and executives with primarily large and medium-sized 
companies in the German-speaking area. The questionnaire used for the survey was designed to 
answer the research questions (cf. Section 1), i.e. to assess design factors for BISOA and to identify 
BISOA approaches. Prior to the survey, the questionnaire was revised by experts from both the 
scientific community and the entrepreneurial world in terms of completeness and comprehensibility.
The items of the questionnaire cover a broad range of layers relevant to the context of BISOA (cf. Section 3.3). The respondents were asked to indicate their agreement with several statements on a five-tiered Likert scale (0 to 4) from the perspective of the organization they are working for. There was a dedicated time slot during the event to fill in the questionnaire. The objectives, structure, and terminology used in the written survey were explained to the attendees. A total of 68 questionnaires were returned. This corresponds to a return rate of approximately 49.6%. If a data set was incomplete (11 questionnaires) or apparently inconsistent (checked by control questions; 6 questionnaires), the questionnaire was discarded. Fifty-one duly completed questionnaires were used as foundation for the analysis, resulting in an overall return rate of about 37.2%. The data set is considered to constitute an adequate basis for an explorative analysis.

Respondents of the survey were employees from organizations in the German-speaking area. Large and medium-sized organizations accounted for the largest share: 23.5% of all organizations have 1000-5000 employees, 36.5% more than 5000 employees. In particular the industry sectors banking (22.0%), insurance (11.8%), and retail (8.8%) were mainly represented.

3.3 Description of the data set

For the purpose of information system analysis in organizations frameworks, differentiating several layers, are often used in order to facilitate a transparent overview of complex correlations between information systems and objects, respectively (e.g. processes) (Alter, 2004; Winter & Fischer, 2007; Zachman, 1987). The deployment of BISOA affects several levels in such frameworks. Therefore, the questionnaire of the empirical analysis incorporated with regards to its content and structure all layers being relevant in this context, namely process level, integration level, and application level.

The process level, which is addressed by all approaches of aforementioned authors, includes all processes along the organization’s value chain. That level was covered in the questionnaire by asking e.g. about the definition of process output and process owners. The application layer, also part of the aforementioned models, subsumes all software artefacts like software services and data structures. The questionnaire addressed this level by primarily technical questions, e.g. about functionality or usage of analytical information systems (Wixom et al., 2008). Finally, the integration layer according to (Winter & Fischer, 2007), links the process layer and the application layer and enables the integration of software components and processes. Considering this layer, questions also addressed integration options of business processes and BI applications, such as support of process execution through analytical information or process monitoring (Kueng, 1999) or interface configuration. In addition to questions focusing on these three layers, statements were asked describing the degree of BISOA realization – in order to derive the corresponding correlations.

3.4 Factor analysis

As stated above, the factor analysis method of this empirical research is based on EFA (cf. section 3.1) in order to identify the predominant design factors for BISOA (cf. research question 1 in Section 1). The overall objective is to find a common term embracing multiple items. The input for the factor analysis comprised of 15 items. Prior to the EFA, the adequacy of the data set is verified using two criteria. Variables (items) are suitable for factor analysis, if and only if the anti-image of the variables is as low as possible. In such a case the off-diagonal elements of the anti-image covariance (AIC) matrix are as close as possible to zero. As suggested by (Dziuban & Shirkey, 1974) a correlation matrix can be seen as unsuitable for factor analysis, if the percentage of the off-diagonal elements unequal to zero (> 0.09) in the AIC matrix is 25 % or more. The criterion of the data set at hand is about 18.5%.

As second verification criterion the measure of sampling adequacy (MSA), proposed by (Kaiser & Rice, 1974), can be applied. MSA represents an indicator for the extent, to which the input variables belong together, and therefore provides information on whether a factor analysis can reasonably be
performed or not. Kaiser and Rice (1974) appraise a value of 0.7 or more as “middling”, i.e. the data set is considered to be appropriate for applying factor analysis techniques (Kaiser & Rice, 1974). According to (Kaiser & Rice, 1974) a value of 0.7 is to be seen as average and 0.8 as good. In the present case the value is 0.802 which justifies the data set as suitable for the factor analysis.

Within the fields of factor analysis two main techniques can be distinguished – the principal components and principal axes factors analyses, which can be counted by far to the most commonly used factor extraction methods (Thompson, 2004). The higher the number of variables of the data set the more similar the results of principal components and principal axes factor analyses will be (Ogasawara, 2000). According to (Thompson, 2004) principal components analysis (PCA) has some additional preferable properties and is probably the most frequently used EFA extraction method for empirical studies. Thus, PCA is used as factor extraction method in the paper at hand. In order to determine the desirable number of factors, two statistical verification methods are applied, the Kaiser-Guttmann criterion and the scree test. According to the Kaiser-Guttmann criterion (Guttman, 1954; Kaiser & Dickman, 1959) the number of factors to be extracted should equal the number of factors with eigenvalues larger than 1 (cf. Table 1). This results in the extraction of four factors explaining 72.3% of the total variance. The scree plot points to the same solution.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Total</th>
<th>% of Variance</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.521</td>
<td>43.474</td>
<td>43.474</td>
</tr>
<tr>
<td>2</td>
<td>1.899</td>
<td>12.662</td>
<td>56.136</td>
</tr>
<tr>
<td>3</td>
<td>1.349</td>
<td>8.993</td>
<td>65.130</td>
</tr>
<tr>
<td>4</td>
<td>1.071</td>
<td>7.142</td>
<td>72.272</td>
</tr>
<tr>
<td>5</td>
<td>0.750</td>
<td>5.001</td>
<td>77.273</td>
</tr>
<tr>
<td>6</td>
<td>0.634</td>
<td>4.224</td>
<td>81.497</td>
</tr>
<tr>
<td>7</td>
<td>0.556</td>
<td>3.707</td>
<td>85.203</td>
</tr>
<tr>
<td>8</td>
<td>0.491</td>
<td>3.274</td>
<td>88.478</td>
</tr>
</tbody>
</table>

Table 1. Eigenvalues

As final step the nature of the underlying constructs has been clarified by applying the Varimax method as the most common factor rotation method with Kaiser normalization (Thompson, 2004). The rotated component matrix is depicted in Table 2. Items are assigned to a factor if the factor loading adds up to at least 0.5 (Härdle & Simar, 2003). The four factors vary in the number of items with a range from three to four items.

<table>
<thead>
<tr>
<th>Item description</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedicated projects aiming at deploying SOA in BI projects.</td>
<td>0.843</td>
<td>0.083</td>
<td>0.193</td>
<td>0.252</td>
</tr>
<tr>
<td>SOA is deployed (selectively) in the context of upcoming BI projects.</td>
<td>0.833</td>
<td>0.164</td>
<td>0.125</td>
<td>0.028</td>
</tr>
<tr>
<td>Know how regarding the deployment of SOA in BI systems is acquired.</td>
<td>0.794</td>
<td>0.199</td>
<td>0.222</td>
<td>0.066</td>
</tr>
<tr>
<td>Service-oriented architectures are applied to operational systems.</td>
<td>0.745</td>
<td>0.071</td>
<td>-0.036</td>
<td>0.292</td>
</tr>
<tr>
<td>Analytical information is used for automatic process execution.</td>
<td>0.310</td>
<td>0.845</td>
<td>0.064</td>
<td>0.144</td>
</tr>
<tr>
<td>Analytical information supports the execution of business processes.</td>
<td>0.043</td>
<td>0.796</td>
<td>0.272</td>
<td>0.063</td>
</tr>
<tr>
<td>Analytical information and process information are combined and jointly interpreted.</td>
<td>0.203</td>
<td>0.749</td>
<td>0.110</td>
<td>0.366</td>
</tr>
<tr>
<td>For all relevant business processes all relevant performance indicators are measured.</td>
<td>0.046</td>
<td>0.622</td>
<td>0.384</td>
<td>0.334</td>
</tr>
<tr>
<td>All activities and their dependencies are defined for all relevant business processes.</td>
<td>0.159</td>
<td>0.195</td>
<td>0.852</td>
<td>0.215</td>
</tr>
<tr>
<td>Process output is defined for all relevant business processes.</td>
<td>0.170</td>
<td>0.134</td>
<td>0.834</td>
<td>0.237</td>
</tr>
<tr>
<td>Process owners are defined for all relevant business processes.</td>
<td>0.111</td>
<td>0.205</td>
<td>0.725</td>
<td>0.201</td>
</tr>
</tbody>
</table>
There exists a mature master data management. 0.174 0.044 0.189 0.790
There exists a mature data quality management for BI systems. 0.233 0.349 0.123 0.758
There exists a mature meta data management for BI systems. 0.088 0.226 0.422 0.650
There exists a mature data ownership concept. 0.183 0.266 0.305 0.635

Table 2. Results of factor analysis (rotated component matrix)

The following paragraphs discuss the identified factors and give a short interpretation for every factor.

**Factor 1: Excellence in BISOA**

Four items were found to have significant impact on the first factor. Their common denominators are maturity aspects of BISOA realization in organizations. According to our analysis, there are four indicators for the achievement of excellence in BISOA: First, dedicated projects aiming at deploying SOA in BI projects, second SOA is (selectively) deployed in the context of upcoming projects, third, know how of introducing SOA concepts into BI system architectures, and forth, SOA usage degree within operational systems. The last indicator seems to be obvious since in most organizations BISOA will probably not be addressed independently of other SOA activities. Ideally all (BI)SOA activities are synchronized to achieve synergy effects.

**Factor 2: Excellence in embedded BI**

The second factor is made up by four items that essentially account for embedded BI, also called “composite applications” in (Eckerson, 2007). Our findings suggest that the level of excellence in embedded BI depends on the usage of analytical information for automatic process execution and the support of analytical information for the execution of business processes. In addition, excellence in embedded BI is positively influenced by the degree of the combination of analytical information and process information. The degree of the measurement of relevant performance indicators for relevant business processes has impact on this factor as well. Especially the first three items are similar to the understanding of operational BI (i.e. the integration of BI and business process management), as described e.g. in (Marjanovic, 2007). Eckerson (2007) (cf. Section 2.2) differentiates several levels of operational BI. The variables loading on the second factor correspond to one of these levels: “facilitate processes” by embedding BI into operational applications.

**Factor 3: Process orientation**

Three items exhibit high loadings on the third factor. In summary, they can be characterized by the term “process orientation”. All variables describe relevant preconditions for alignment of business activities to business processes: All activities and their dependencies as well as process output and process owners are defined for all relevant business processes.

**Factor 4: Excellence in data management**

Finally, four items covering relevant data management aspects were found to have substantial positive impact on the fourth factor, namely master data management, data quality management, meta data management, and data ownership concept.

### 3.5 Cluster analysis

In order to identify distinct realization approaches of BISOA, cluster analysis was applied on the data set using the calculated factor values of the four previously identified factors as input data. The Ward algorithm and the squared Euclidean distance have been used as fusion algorithm and distance measure, respectively. In literature the Ward algorithm is recognized as an efficient partitioning mechanism (it reveals the appropriate number of clusters with a similar number of observations in each cluster at the same time) generating good clustering results (Hair Jr et al., 2006; Ward Jr, 1963).
Thus, it is applied in the study at hand. Starting with individual cases each representing a single cluster, Ward’s method continues by combining them into clusters until each and every case belongs to the same cluster. For determining which clusters have to be merged next, the sum of the squared Euclidean distance between each case and the mean of its cluster is minimized. While also other clustering algorithms and distance measures have been tested best results in terms of interpretability, context, and purpose of the study at hand could be realized by means of the Ward method in combination with the squared Euclidean distance. The so-called dendrogram (cf. Figure 1) provides a visualization of the hierarchical clustering process. By means of the dendrogram the number of clusters to be built for a particular clustering problem can be graphically derived. In the context of the study at hand this heuristic suggests that the construction of four clusters (representing four distinct realization approaches of BISOA) is the most reasonable solution, marked in the dendrogram between the two dotted lines.

![Figure 1. Results of cluster analysis (dendrogram)](image)

Table 3 exhibits the arithmetic means and the standard deviations of the four calculated factor values of the four clusters. A graphical representation of the arithmetic means is added in form of pie charts (please note: the pie charts are representing relative, not absolute values).

<table>
<thead>
<tr>
<th></th>
<th>Excellence BISOA</th>
<th>Excellence embedded BI</th>
<th>Process orientation</th>
<th>Excellence data mgt.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Arithmetic mean</td>
<td>Standard deviation</td>
<td>Arithmetic mean</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>0.76</td>
<td>0.66</td>
<td>0.30</td>
<td>0.70</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>0.67</td>
<td>0.99</td>
<td>-1.05</td>
<td>0.89</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>-0.61</td>
<td>0.79</td>
<td>0.66</td>
<td>0.94</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>-0.74</td>
<td>0.68</td>
<td>-1.42</td>
<td>1.27</td>
</tr>
</tbody>
</table>

Table 3. Results of cluster analysis (arithmetic means and standard deviations of factor scores for each cluster)

Based on this information, the clusters can be characterized as follows:

**Cluster 1: Data management experts**

Those 21 organizations (i.e. the majority of all respondents) that are grouped into the first cluster are characterized by the most advanced degree of BISOA realization and by high levels of excellence in data management but by average level of excellence in embedded BI and of process orientation. These results correspond e.g. to (Gordon et al., 2006) who claim data management aspects like data ownership, enterprise data, etc. as critical success factors for BISOA.

**Cluster 2: Process orientation experts**

The second cluster is made up of 9 organizations. Although the cluster is characterized by nearly the same degree of BISOA realization as the first cluster, the remaining design factors differ significantly. In contrast to the first cluster these organizations exhibit a high level of process orientation but low levels of embedded BI and data management.
**Cluster 3: Embedded BI experts**

The third cluster merges 14 organizations that are characterized by a low degree of BISOA realization, but the highest excellence level of embedded BI. On the other hand the maturity level regarding data management and process orientation is rather low.

**Cluster 4: Freshmen**

Nine organizations where embedded BI and process orientation get the lowest attention in comparison to the other clusters are grouped in this cluster, which is characterized by the lowest degree of BISOA realization. The maturity level of data management is on average.

According to the values determined in the cluster analysis, the four realization approaches to BISOA can be arranged in a two-dimensional matrix (cf. Figure 2). “Excellence in data management” (factor 4) is being depicted on the horizontal axis and “Excellence in BISOA realization” (factor 2) is being displayed on the vertical axis. For both dimensions, high and low parameters (i.e. high and low levels of implementation) are distinguished for clarity. Thus, the classification scheme resembles a 2x2 matrix. Within each of the four segments, we furthermore differentiate between the “Process orientation (factor 3, horizontal axis) and “Excellence in embedded BI” (factor 2, vertical axis). Cluster 4 (“Freshmen”) might also be assigned to the lower left side of the matrix as its maturity level of data management represents an average value, which cannot be clearly allocated to high or low values.

**Figure 2. Realization approaches of BISOA**

4 DISCUSSION AND INTERPRETATION OF RESULTS

We have identified four relevant BISOA profiles in Section 3. In the following, two aspects regarding the results of the empirical analysis and the cluster analysis will be discussed in detail. Further interpretation of the results might be useful and interesting.

**Improvement of the BISOA maturity**

Figure 2 illustrates the fragmentation of surveyed organizations into two main groups: One group is characterized by a high degree of BISOA realization and the second group by a low degree, respectively. However, within these groups there is no homogeneity concerning the maturity levels of the remaining (design) factors. Obviously, there is no single and definite approach to deploy BISOA or to increase BISOA maturity levels, respectively. According to the analysis results, a high BISOA maturity level corresponds either to excellence in data management or to a high degree of process...
orientation. There are no organizations (at least among the respondents) that are characterized by high maturity levels of both, data management and process orientation, or even of all three factors (including embedded BI).

In the survey we also asked to indicate the same statements for the future. These values are not in the focus of the paper at hand and of the empirical analysis presented so far. However it is evident that nearly all organizations aim at improving the degree of BISOA realization. This is also true for organizations with a high degree of BISOA realization in the cluster analysis (cluster 1 and cluster 2). This might be explained by the reason that these factor values have to be regarded as relative, not absolute. I.e., even organizations belonging to cluster 1 and 2 do not realize a true mature level of BISOA so far. According to the analysis results it might be concluded that activities to increase the degree of BISOA realization might be positively influenced by simultaneous improvement of the remaining (design) factors (data management, process orientation, and embedded BI). Regarding the fact that very few organizations are characterized by mature levels of all three “dimensions” (factors), it seems that organizations with comparatively mature BISOA (cluster 1 and cluster 2) still have significant potential for further improvement regarding BISOA.

Relevance of BISOA for operational BI scenarios

As mentioned in Section 2.2 many authors link SOA with operational BI and praise operational BI as the main use case for BISOA. The analysis results have revealed that this coherence cannot be seen as evident as it is often claimed. At least regarding the facets of embedded BI / composite applications – according to (Eckerson, 2007) one stage in operational BI - do not correlate unambiguously to BISOA. In cluster 1 and cluster 2, which include organizations with a comparatively high mature BISOA approach, both enterprises types, with an advanced degree of embedded BI and with a low degree, have been found. Accordingly, the same situation can be observed in cluster 3 and cluster 4 (organization with a comparatively low mature BISOA approach). Thus, BISOA does not necessarily constitute one or even the only driver for operational BI and vice versa. At least it seems that coherence and interdependencies between BISOA and operational BI are not as straight and definite as often assumed, especially in practice. In fact, other factors also have impact, namely the degree of process orientation and the maturity of data management, as the study at hand has shown. One-track argumentation regarding BISOA and operational BI and significant correlations in-between should consequently be seen more carefully, at least if any coherencies are only assumed and not somehow proved.

5 SUMMARY AND OUTLOOK ON FUTURE WORK

Deploying the SOA paradigm to analytical information systems, promises progress in realizing the vision of closed action loops between operational systems and analytical information systems. However, BISOA so far is not addressed extensively by scientific community. Consequently, we identified two research questions aiming at finding insights about the interdependencies of BISOA and an organization’s system landscape. A factor analysis and a cluster analysis were applied in order to identify predominant design factors for BISOA in an explorative way. We also determined typical scenarios for BISOA and explored the impact of relevant EA layers on the shaping of those scenarios.

The resulting design factors might support the application of models and methods in information systems research for organizations that are planning or currently realizing BISOA. The factors represent dimensions in which BISOA profiles differ from each other. As mentioned in Section 4, approaches for BISOA deployment vary in the surveyed organizations. Thus, models and methods that tend to support the deployment phase should be adaptable in order to address the design factors and different maturity levels (“BISOA profiles”) elaborated in the study at hand. For instance, data management as an essential aspect should be incorporated in a method while considering the specific maturity level of data management in the organization.
Moreover, the design factors can serve as a guideline and provide recommendations for actions. Due to their identified impact on the maturity level of BISOA, they clarify which issues should be addressed in addition to the usually technical dominated implementation aspects of BISOA. In other words, it is not recommended to regard BISOA as an isolated IT project – indeed, the coherencies to other aspects (represented by the design factors) should be considered carefully.

Surprisingly, the cluster analysis depicted that correlations between BISOA and operational BI (at least when focusing on the facet of embedded BI) are not as evident as often supposed. It seems that additional factors influence BISOA simultaneously.

The analysis results might serve as the basis for further research. On the one hand, it seems worth to further detail the scenarios and to elaborate their characteristics. Possible development paths for BISOA and more detailed recommendations may result in a comprehensive procedure model and a maturity model, respectively. On the other hand, as already mentioned, in the survey the respondents were also asked about their assessments for the future. Since this information is not included in the empirical analysis so far, further research might generate new interesting insights. Finally, the survey might be expanded to further organizations in order to put the empirical analysis on a firmer footing.

References


Bid price control and dynamic pricing in Clouds

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0297.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Grid Computing, Pricing, IT Services, Heuristics</td>
</tr>
</tbody>
</table>
BID PRICE CONTROL AND DYNAMIC PRICING IN CLOUDS

Arun Anandasivam, University of Karlsruhe, Englerstr. 14, 76131 Karlsruhe, arun.anandasivam@kit.edu
Marc Premm, University of Karlsruhe, Englerstr. 14, 76131 Karlsruhe, marc.premm@student.kit.edu

Abstract

The term Cloud Computing represents a paradigm for offering different kind of Web services, which can be dynamically developed, composed and deployed on virtualized infrastructure. This work will extend the concepts known from the revenue management to the specific case of Cloud Computing and propose two models, bid price control and a variant of dynamic pricing, that will compete with the commonly used static pricing. Both models will try to maximize revenues by controlling the availability or price of every offered fare class. The aim is to understand from a Cloud Computing company’s perspective, how decisions about the pricing and the optimal allocation of the given resources for the various Cloud Services can be supported. As expected, simulation results show that an optimally adjusted dynamic pricing model will outperform any pricing model with static prices and will simultaneously contribute to slightly smoother resource utilization in some cases. However, we will see that the adjustment itself is difficult to realize, and if conducted suboptimal, it may also have certain disadvantages compared to static prices. In combination with a reasonable product differentiation, the bid-price method performed very solid and in nearly any case better than the pure static pricing model.

Keywords: Cloud Computing, revenue management, pricing, IT services
1 INTRODUCTION

Cloud Computing has recently become very popular as a new paradigm to shift IT resources and software from locally independent computers to a more collaborative level (Hayes, 2008). The hub-and-spoke system known from the client-server model is replaced by a geographically distributed cluster system. The concept of cloud computing comprises many existing technologies and architectures like on-demand or utility computing, software as a service and AJAX. Providers like Amazon or Rackspace offer products on the infrastructure level to run individually configured operating systems. Server farms remain idle most of the time. 99 percent of the entire computing capacity of a company is only used 10 percent of the time (Weiss, 2007). Additional revenue can be earned by utilizing resources. The resource demand is mostly both dynamic and unpredictable. Defining demand based pricing policies can influence the behavior of price sensitive customers (Desiraju and Shugan 1999). Customers, whose utility is highly depending on prices, would shift their computational jobs to a time, where resources are cheap. Business customers, though, are less flexible and are thus less price-sensitive. Their utility is increased by getting the job done on time or getting a better service. When prices are relatively fixed and demand outweighs supply, providers have to manage their capacity by deciding whether to accept and incoming request or waiting for a future request demanding a higher valued service. This will probably induce higher revenue. This problem is also known as the capacity management problem from the Revenue Management domain (Talluri and van Ryzin, 2004). Solutions have been successfully applied to the airline sector in the majority of cases (McGill and van Ryzin, 1999).

In this paper, we analyze how Revenue Management concepts can be useful to the Cloud Computing domain. The aim is to understand from a Cloud Computing company’s perspective, how decisions about the pricing and the optimal allocation of the given resources for the various Cloud Services can be supported. Our contribution is threefold. At first, the requirements for applying Revenue Management to Cloud Computing have to be analyzed. In section three we derive a mathematical model to determine appropriate policies when to accept or reject requests for Cloud Services. Two models are compared, namely bid price control and dynamic pricing. As a supplementary benchmark, we use a static pricing model. The proposed bid price model and dynamic pricing models are evaluated via simulation in section four. Section five concludes the paper.

2 RELATED WORK

2.1 Cloud Computing

The term Cloud Computing is currently used in various ways and is often confounded with the term Grid Computing. Grid Computing accrued in the mid 1990s and originally denoted a scientific network to share computation power for computationally intensive jobs. The main characterization of a Grid is the distribution of a computing job in a somehow connected network. Accordingly, jobs for Grids are divided in several small jobs which are distributed to independent servers or desktop computers (Foster et al., 2001). This opens the possibility to share resources between institutions that are dispersed among different geographical locations. However, sharing computing resources is also a central aspect of Cloud Computing, but with a focus on virtualized instances that usually run on a server cluster (Buyya et al., 2008). A server cluster typically distributes computation power among several locally connected servers and hosts the virtual instance, which has been allocated to a certain customer. In scientific Grids the institutes are provider and consumer similar to P2P networks. Cloud Computing is commercially driven and most of the providers do not consume the resources they offer. The role between provider and consumer is currently disjointed. Some providers offer services in Clouds, which are based on other services like Rackspace, Rightscale or MorphExchange. In particular, services in this case are distinguished. For example, the product MorphExchange is a
platform providing a Tomcat server for uploading and hosting Java web applications (Platform as a Service). It uses Amazon’s Simple Storage Service (Infrastructure as a Service) for storing the applications. On top these services, software or data can be provided as a service as well like StrikeIron offering a data service in combination with Salesforce’s Customer Relationship Management software service. Lawton describes the cloud applications (Software as a Service) as Web-based applications accessed via browsers but with the look and feel of desktop programs (Lawton 2008). Skillicorn already emphasized in 2002 the advantages of composing simple services together to a new Web service with an added value for the customer (Skillicorn 2002).

Another aspect is the relationship between Clouds and Grids. Grid and Cloud Computing can be viewed as integrating several instances of one or multiple Clouds in a Grid. For example, we can use Grid technology for connecting multiple Clouds that do not have to be allocated at the same geographical location to combine computation power and furthermore distribute utilization (Boss et al., 2008). This definition implies a standard protocol to interact between several cloud instances. However, this is strictly depending on the providers, who supply Web Services to their products. Moreover, currently the Grid is more academia driven, while Cloud Computing is focusing on commercial application (Weinhardt et al., 2009). Earlier approaches from IBM or SUN Microsystems to commercialize Grid Computing have not been successful yet.

2.2 Revenue Management

The term Revenue Management is most commonly used for the theory and practice of maximizing expected revenues by opening and closing different fare classes or dynamically adjusting prices for products (Talluri and van Ryzin, 2004). The development of scientific research methods in this discipline started after the deregulation of the American Airline industry in 1978. They relaxed restrictions over standardized prices and profitability targets enabling dynamic pricing and resource allocation, although the first approach was back in 1972 by Littlewood (Littlewood, 1972, Belobaba, 1987). (Netessine and Shumsky, 2002) define Yield Management as a part of Revenue Management although boundaries between both are often ambiguous.

The perishability of the offered products is one of the main characteristics of Revenue Management. For example, a hotel cannot save up one room for the next day. Instead, an unused room becomes worthless without creating any revenue (Netessine and Shumsky, 2002). To get an efficient return from utilizing Revenue Management techniques the arriving customers have to be segmented into different classes. Every customer segment can be distinguished by different needs and thus by higher or lower reservation prices. Consequently, the main goal in Revenue Management theory is to find the combination of customers which seems to provide the highest possible revenue (Kimes, 1989). Therefore, we have to choose between opening a certain product class for sale or protecting it for a more profitable customer. However, future demand is not certain what causes this problem to get far more complex. If the vendor decides to protect a product for future demand, he takes the risk of ending up without selling the product (Goldman et al., 2002, Netessine and Shumsky, 2002). Although the vendor would face lower risk by immediately using the possibility to sell the product, we assume for all models presented in this work that these transactions are repeated often enough to justify a risk-neutral approach (McGill and van Ryzin, 1999, Bitran and Caldentey, 2003).

2.3 Application of Revenue Management in Clouds

The application of Revenue Management to other domains is only feasible, if several requirements are fulfilled. The requirements enable a comparison to traditional Revenue Management approaches which will be useful to formulate and optimize the unique optimization problem analogously to established strategies.

a. Time horizon: Like the Amazon EC2, the duration of every service is defined as a finite time period. With this assumption, we can handle every full hour as a separated product, what makes it comparable with a seat on one flight leg in airline industry or a room for one night in a hotel (Kimes, 1989).
b. Perishability: The perishability of the offered products is one of the main characteristics that have to be fulfilled for an appropriate use of Revenue Management (Weatherford and Bodily, 1992). The usage of resources on servers is limited to a point of time. Afterwards they are worthless, since they are not storable. As long as enough IT resources are left to provide at least one product, the revenue is still not maximized.

c. Production Inflexibility: IT resources are largely fixed and can only be extended at disproportional costs in relation to their running expenses. Product supply is thus limited and not extensible in a specific time horizon (Weatherford and Bodily, 1992). As most Cloud Computing providers offer their services on relatively huge server cluster. The integration of an additional resource like buying a new server will incur high additional fixed costs due to integration and increasing maintenance cost for the provider of infrastructure services.

d. Possibility of booking future products: Advance reservation has some relevant benefits (Boss et al., 2008, Dube et al., 2005). On the one hand, it gives the user the possibility to ensure a prospective computation demand by reserving the required products for the desired time. On the other hand, it provides the seller with the ability to easily discriminate customers by their valuation.

e. Customer segmentation: The offered products have to be differentiated by adding or restricting certain features to reach the desired customers for every price class. In most cases, the demand for these price classes differs particularly in their valuation for special product features and price sensitivity (Talluri and van Ryzin, 2004). For example a service level might indicate the minimum percentage of availability. Additionally, we might also identify different needs for resource combinations between price segments and with future booking enabled. Moreover, restrictions for reservation changes can be adopted commonly known from the airline industry.

f. Multiple products based on same resources: For Cloud Computing we have the desired advantage of flexible products. The resources of a Cloud Computing center are able to provide different service offerings using the same IT resources. Several products or product bundles can be defined, e.g. containing different amounts of CPU power or memory and defining individual service levels as aforesaid. The flexibility of the initial capacity provides plenty of resource combinations (Bitran and Caldentey, 2003).

h. Overbooking: The Cloud Computing overbooking procedure would be to sell more of the computing and storage capacity than the computing center has. In this case, not every costumer will exploit its reserved resources completely. Overbooking of Cloud Computing resources allows more flexibility. (Urgaonkar et al., 2002) show that the usage of overbooking techniques can increase utilization drastically: Already an overbooking rate of just 1% may increase the utilization of the whole computing cluster by a factor of two without losing meaningful availability guarantees.

It is important to consider customer needs. On the one hand, we believe, that in contrast to most traditional Revenue Management industries there is a market for open-end requests, which explicitly do not have certain end dates. That is because these virtual computing resources are likely to replace local computers which are in most cases also bought without a fixed time horizon for its usage. This issue is considered in our model. On the other hand, these requests would lead to the loss of an accurate resource schedule for future booking requests, so we will assume for the following, that there are no open-end requests, instead these have to be formulated as a long-term request with a finite time horizon (Boss et al., 2008).

3 PRICE DETERMINATION

The goal of a provider is to identify the appropriate price to maximize his revenue and satisfy the consumer demand. To give the customer the highest possible flexibility, we will allow every possible hardware configuration based on the given resources $k = 1...K$. This approach can already be observed
in the industry. For example the Enterprise Cloud of terremark\(^1\) has the capability to offer individually configured servers. The user can assemble the desired product by choosing in each case a certain amount of processing power, storage size and memory size.

### 3.1 Scenario

Products are requested over time. Within a product planning period \(T_p\) resources can be planned and allocated at \(t_p = \{0, \ldots, T_p\}\). Every product and thus every resource \(k\) for every point in time \(t_p\), will be offered in different fare classes \(c\). The price differences between classes will be justified with different service levels. For example, the higher valued class A will ensure 99.9\% availability while the lower valued class B will only provide 99\%. Other services may be faster reaction times or higher security specifications like redundancy checks. Following (Bitran and Caldentey, 2003), the connection between price and demand dependent on the time \(t\) will be expressed by the stochastic demand function \(D(t, p, \xi)\) with noise parameter \(\xi\). This assumption implicates myopic customer behavior.

Furthermore, we will not enclose any approach that considers demand between fare classes as dependent. The introduction of probability dependencies like passenger diversion or degradation costs (Botimer and Belobaba, 1999) would make the problem far more complex and too hard to compute. Variable costs of a Cloud Computing center are relatively low. Because servers and computers in general only have a minimal abrasion, the costs of one additional customer can be reduced to the energy consumption costs that increase because of the higher utilization. However, empirical studies have shown that an idle server needs nearly 90\% of the average power consumption (Fan et al., 2007). Thus variable costs will be neglected in this paper. Consequently, with marginal costs being close or equal to zero, the marginal profit equals the turnover and so every sold product will contribute to the gained profit, regardless of the price.

### 3.2 Basic model

A common assumption is that for every time \(t\) there is at most one request from any customer. However, with Cloud Computing services being a time dependent service, the time interval of the offered product has to be predetermined. In the majority of the current business cases these are full hours (e.g. Amazon’s EC2 product). To overcome this problem and to relax this assumption, a second time scale \(t_r = 0, \ldots, T_r\) is introduced. Simultaneously a restriction is required for the number of arriving requests per time interval which will be set to \(R\). Then, the connection between the product planning time \(t_p\) and the request time \(t_r\) can be defined as \(t_r = \left\lfloor \frac{t_p}{R} \right\rfloor = t_p\). Accordingly, the time horizon for the request time \(T_r\) can be formulated as \(T_r = R T_p\). This term allows us, on the one hand, modeling the complex situation and, on the other hand, using the assumption that there is at most one request in every time interval \(\Delta t\) between to sequent points of \(t_r\).

The capacity left for sale in \(t_r\) will be represented by the matrix \(X_{t_r} = (x_{k,t_p}^{t_r})\) for every request time \(t_r\). As we reckon that initial resource capacity \(\Gamma_k\) of product \(k\) will not change over the time horizon \(T_p\), we can define the initial capacity matrix for every resource \(k\):

\[
X_0 = (x_{k,t_p}^0) \text{ with } x_{k,t_p}^0 = \Gamma_k \forall t_p \in \{1, \ldots, T_p\}.
\]

The next important step is to unify a booking request. As we have stated above, the request has to include the amount of desired resources for every possible resource \(k = 1, \ldots, K\). So, we define a vector \(r\) with elements \(r_k\) for every resource \(k\). Additionally, we also need the information about the requested start time \(s \in \{t(t_r)+1, \ldots, T_p\}\) and the duration of the service \(d \in \{1, \ldots, T_p - s + 1\}\). Consequently, every submitted request has to contain the vector \(r\) as well as the variables \(s\) and \(d\). The

\(^1\)http://www.theenterprisecloud.com/
dependency of these variables is expressed in the request matrix $R^{tr}$ for every request time $t_r$ as follows:

$$R_{t_r} = (r_{k,t_p}) = \begin{pmatrix} 0 & \ldots & 0 & r_{1,s} & \ldots & r_{1,s+d-1} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & r_{K,s} & \ldots & r_{K,s+d-1} & 0 & \ldots & 0 \end{pmatrix}$$

Let $R^{tp}_{t_r}$ denote the column $t_p$ of $R_{t_r}$. This definition will be important for the model formulation and will even allow requests that have varying resource needs over time. The goal of Revenue Management is to maximize the profit. Therefore, let $V(t|X)$ be the expected revenue at time $t$ and with the state of available resources $X$. Next, we present different models to determine the expected profit.

### 3.3 Static pricing

The first model that we will look at fixes all prices for the whole time horizon. Therefore, it has to set the price vector $p^c = (p^c_0, \ldots, p^c_K)^T$ which contains prices for every resource $k$ for every fare class $c$. Marbach introduced a similar problem which tries to optimize the general static pricing problem for a network service provider. However, they do not consider multiple resources (Marbach, 2004). Because prices are fixed for the whole time, the optimization for the static pricing model has to be done only once, more precisely before the first time interval of the time horizon. To accomplish this task, we have to solve the general optimization problem, depending on the price-response-function $D_c(t, p, \xi)$ for class $c$ like it has been introduced above:

$$\max_{p_k} V_{t_r}(X_{t_r}) = \sum_{t_{p=1}}^{T_p} \sum_{k=1}^{K} \sum_{c=0}^{C} p_k E[D_c(t_p, p_k^c, \xi)]$$  \hspace{1cm} (1)

As long as we do not know anything about $D_c(t, p, \xi)$ yet, we assume that all prices $p_k^c$ are given by the market and we have no possibility to change them at any point.

### 3.4 Bid price model

The bid-price model can be seen as an extension to the static pricing model introduced above (Goldman et al., 2002). They both have in common, that prices for every fare class and for every resource are fixed over the whole time horizon. Consequently, to find optimal prices, a similar optimization problem has to be solved (Williamson, 1992, Talluri and Ryzin, 1998). For every request matrix $R^{tr}$ that arrives at request time $t$, with start point $s$ and duration $d$, we can define the revenue that is gained by accepting this request as

$$\sum_{t_{p=1}}^{s+d-1} (p^c \top R^{tp}_{t_r})$$  \hspace{1cm} (2)

In case the provider would accept a delivered request $R^{tr}$, he will gain the profit like we have just seen, but simultaneously the available resource for the next request time $t, +1$ will lower exactly by the amount of the request to $X_{t_r+1} = X_{t_r} - R_{t_r}$. So, the expected revenue for the remaining capacity $X_{t_r+1}$ will be

$$V_{t_r}(X_{t_r+1}) = V_{t_r+1}(X_{t_r} - R_{t_r}).$$  \hspace{1cm} (3)

However, the provider, who receives a request, does not have very much room for different strategies. He simply has to choose whether to accept or reject an incoming request. The decision is defined by the indicator variable $u \in U = \{0,1\}$ with one for acceptance and zero for rejection. With the preceding formulas, we can formulate the Bellman equation similar to (Talluri and Ryzin, 1998) to find the optimal decision $u^*_{t_r}$

$$V_{t_r}(X_{t_r}) = \max_{u_{t_r} \in U} \sum_{t_{p=1}}^{s+d-1} (p^c \top R^{tp}_{t_r} u_{t_r}) + V_{t_r+1}(X_{t_r} - R_{t_r} u_{t_r})$$  \hspace{1cm} (4)

s. t. $x_{ij} - r_{ij} \geq 0 \forall i, j$
The bid-price approach is very intuitive by comparing immediately realizable revenue with their opportunity costs. Let \( S_k^p = (s_{k,t_p}) \) a \( K \times T_p \) matrix where the element \( s_{k,t_p} \) equals 1 and every other element equals 0. With the assumption that the function for the expected revenue \( V_t(X) \) has a derivative \( \frac{\partial}{\partial x_k} V_t(X) \) for the resource \( k \), the approximated condition for accepting a request for a single resource \( k \) at time \( t_p \) with price \( p_k^c \) is (Talluri and van Ryzin, 2004):

\[
p_k^c \geq V_t(X_{t_p}) - V_t \left( X_{t_p} - S_k^p \right) = \frac{\partial}{\partial x_k} V_t(X_{t_p}) = \pi_{t_p}^k. \tag{5}
\]

Next, (5) has to be extended for a whole request, including start time \( s \), duration \( d \) as well as multiple resources \( k \). This simply can be done by summing up all relevant prices for the requested fare class \( c \) and comparing it with the total sum of all relevant bid-prices for the given request. This approach allows us to compensate resources that may be sold under value with ones that will be sold for a higher price than expected. For pre-calculated bid-prices vectors \( \pi_{t_p} = (\pi_{t_p}^1, \ldots, \pi_{t_p}^K) \) we can reduce the necessary inequality that a request for fare class \( c \) has to satisfy to

\[
\sum_{t_p=s}^{s+d-1} p^c T_{t_p} \geq \sum_{t_p=s}^{s+d-1} \pi_{t_p}^c R_{t_p}^f . \tag{6}
\]

This condition ensures that the provider will only sell resources in fare classes that at least gain the revenue that he is expecting to gain. Consequently, an incoming request should be accepted if and only if the inequality (6) is fulfilled. We can formulate the optimal decision function as

\[
u^*_r = \begin{cases} 1 & \text{if (6) is satisfied and } x_{k,t_p}^r - t_{k,t_p}^r \geq 0 \forall k, t_p \\ 0 & \text{otherwise.} \end{cases} \tag{7}
\]

Because we only know the actual request at time \( t_r \) we cannot directly use the recursive definition of (4). Instead, we have to find a heuristic that uses another method to get the expected revenue of any remaining capacity \( X \). In turn, the knowledge of the expected revenue together with (5) will enable the calculation of every necessary bid-price \( \pi_{t_p}^k \).

For the following, we want to present one of several possibilities to value a certain state \( X \) of the available resources \( k \). In any case, we need to model the expected demand to say anything about the valuation of the resources that are left. Thus, a demand function \( D(k, t_p) \) is required that will return the estimated demand for resource \( k \) at time \( t_p \). However, the valuation cannot exceed the available capacity stored in the vector \( X_{t_p}^r \), and the demand function has to be adapted by

\[
D(k, t_p) = \min \left[ \bar{D}(k, t_p), x_{k,t_p}^r \right]. \tag{9}
\]

Let \( p_k^f(C | C = c) \) the probability of any demand for resource \( k \) at time \( t_p \) to equal fare class \( c \) considering that higher valued classes have to be handled primarily to the disfavor of the lower valued ones. The expected revenue can be calculated by

\[
V_t(X_{t_p}) = \sum_{t_p=t_1(t_r)+1}^{T_p} \sum_{k=1}^{K} \sum_{c=0}^{C} p_k^c D(k, t_p) P_k^f(C | C = c) \tag{10}
\]

s.t. \( \sum_{c=0}^{C} P_k^f(C | C = c) = 1 \forall k \).
### Algorithm 1 Transform $\tilde{D}(c)$ to observe capacity restrictions

**Require:** expected demand $\tilde{D}(c)$ for fare class $c$

**Ensure:** expected demand $D(c)$ under capacity restrictions

1. $X \leftarrow$ capacity left
2. for $c = 0$ to $C$
   1. if $X > \tilde{D}(c)$ then
      1. $D(c) \leftarrow \tilde{D}(c)$
      2. $X \leftarrow X - D(c)$
   2. else
      1. $D(c) \leftarrow X$
      2. $X \leftarrow 0$
3. endfor

### 3.5 Dynamic pricing

The following Pricing Model does not try to derive a benefit from denying requests for lower valued fare classes, instead it attempts to adjust prices for every fare class. Consequently, we can increase the price for a product to maximize profits rather than just limiting the capacity of a product and leaving the rest of the demand unused. The intuitive issue behind this approach is that fluctuations in demand can be compensated while simultaneously increasing revenues.

The frame of the optimization problem is very similar to the bid-price model of the preceding section. However, the big difference will be the variable that we are going to optimize. In both preceding models we had to decide whether to accept an incoming request or not, what was represented by the decision variable $u$. In this case, we will primarily accept any incoming request that will not exceed our remaining capacity, but we will influence demand by adapting the price. Consequently, the variable that we have to optimize is the price for every fare class $c$. As we will not allow every real number as a possible price for a fare class $c$, we assume that we have to select one element of the set $\mathcal{P}_{c,k}$, which contains all possible prices for fare class $c$ and product $k$. The use of such a discrete set is very common in business practices. Especially conventional retailing industries often use price steps of at least one full monetary unit or in higher price regions even steps of integer multiples of those (Talluri and van Ryzin, 2004).

The difference to static pricing models is that we have to know something about the relation of price and demand to solve most optimization problems. Since we do not know the exact price-demand function, we assume that we have such a function $\tilde{D}(k, t_p, c)$ which returns the demand dependent on the price $p$. Analogous to the bid-price model, we have to transform this function with the algorithm 1 or with (9) to fulfill capacity restrictions. A probability distribution has to distribute the demand to the different fare classes (similar to the bid-price model). The revenue calculation also has to be slightly changed to include the price dependent demand function and different prices for every time interval. So the optimization problem that has to be solved is $\max_{p \in \mathcal{P}_{c,k}} p \cdot D(k, t_p, c)$. As this optimization problem has to be done for every resource $k$, fare class $c$ and planning time $t_p$, we can formulate the Bellman equation that will recursively return the expected revenue:

$$V_{t_r}(X_r) = \sum_{t_p=s}^{s+d-1} \sum_{k=1}^{K} \sum_{c=0}^{\mathcal{P}_{c,k}} \left( \max_{p \in \mathcal{P}_{c,k}} p \cdot D(k, t_p, c, p) \right) + V_{t_{r+1}}(X_r - R_{t_r}). \quad (11)$$

There are still two practical problems to directly use this term: On the one hand, we have to know or approximately assume a specific relation between price and demand. On the other hand, it is obvious that this calculation will be very complex, as we have to solve at least $d \cdot (C + 1) \cdot K$ different optimization problems and we still do not know the valuation of the future time period $t_{r+1}$. Therefore, we want to present a heuristic that will adjust prices depending on the variables $k$, $c$ and $t_p$. This heuristic will calculate a parameter $\gamma$ for every resource $k$, time $t_p$ and fare class $c$ by setting a
specific relation between the expected demand $D(k, t_p, c)$, the capacity left $x_{k,t_p}$ and the time that is left until the requested start time $s - t_r(t_r)$:

$$
\gamma(k, t_p, c) := \frac{\text{Expected Demand}}{\text{Capacity Left}} \ln(\text{Time To Go}) = \frac{D(k, t_p, c)}{x_{k,t_p}} \ln(s - t_r(t_r)).
$$

We assume that every parameter $\gamma(k, t_p, c)$ is positively correlated to the valuation of the customers and can serve as an approximation for the price-demand relation that we still do not know. Let $P_{c,k} = \{\hat{p}_{c,1}^k, \ldots, \hat{p}_{c,P}^k\}$ the set of possible prices for fare class $c$ and resource $k$ with the ascending order $\hat{p}_{c,1}^k > \hat{p}_{c,2}^k > \cdots > \hat{p}_{c,P}^k$. We can define the function

$$
p(k, t_p, c) := \begin{cases} 
\hat{p}_{c,1}^k & \gamma(k, t_p, c) > \tau_1(k, c) \\
\hat{p}_{c,2}^k & \tau_1(k, c) \geq \gamma(k, t_p, c) \geq \tau_2(k, c) \\
\vdots & \\
\hat{p}_{c,P-1}^k & \tau_{P-1}(k, c) \geq \gamma(k, t_p, c) \geq \tau_P(k, c) \\
\hat{p}_{c,P}^k & \text{otherwise}. 
\end{cases}
$$

(13)

to get the proper price for resource $k$, time $t_p$ and fare class $c$ with the threshold values $\tau_i(k, c) \forall i \in \{1, \ldots, P\}$. We assume that all $\tau_i(k, c)$ will be defined manually at the beginning of our time horizon, primarily based on empirical knowledge. A possible orientation might be the expected value of the parameter $\gamma(k, t_p, c)$, so every $\tau_i(k, c)$ might be expressed as

$$
\tau_i(k, c) := \alpha_i E[\gamma(k, t_p, c)]
$$

(14)

with the freely chosen real multiplier $\alpha_i > 0$ and the expected value $E[\gamma(k, t_p, c)]$ of $\gamma$ for every $t_p$.

4 Evaluation

4.1 Scenario

As requirements mostly cannot be foreseen even by the customers, we will only allow requests for a fixed amount of every resource $k$ for a certain time period $\Delta t = d$. Forecasting of future demand and utilization are essential for overbooking decisions in the specific Cloud Computing case. Since this task is not in the focus of this paper, so some relatively simple, but intuitive approaches will be applied. Demand is independently distributed over time, so a very intuitive step is to use historic demand data to estimate future demand. To eliminate smoothing effects, not all the historical data will be considered, but a limitation to the last $r$ values. In this case, the historical data consists of the observed demand $D(k, t, c)$ for product $k$, time $t$ and fare class $c$. The expected demand at the request time $t_r$ for every future time $t_p \geq t_r(t_r)$ can be expressed as

$$
D_{t_r}(k, c) = \frac{1}{N_D} \sum_{t=t_r-N_D}^{t-r-1} D(k, t, c)
$$

(15)

where the last $N_D$ values of the observed demand are taken into account. The variable $N_D$ might be increased for a smoother demand forecasting or decreased for a quicker adaptation to market changes.

For utilization, we will assume that the utilization behavior will change over the time of day, while the progress will be similar from day to day. For overbooking decisions in a Cloud Computing environment, we have to set some kind of a service level which indicates that how many resources may be additionally sold. Let $\bar{U}_n(t, k)$ the recognized usage of customer $n$ for the product $k$ at time $t$. With $N_D$ last values and resource $k$, we can define the average value and the variance of the past utilization as

Proceedings ECIS 2009
With a predefined service level $SL$ and the assumption that the utilization of every product follows a normal probability distribution, we can solve the following equation and adapt the available capacity accordingly to ensure that we can provide our services with the probability of $SL$:

$$P(\hat{u} \leq SL) = \int_{-\infty}^{\infty} \frac{1}{\sigma_{k,t,r} \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{t - \mu_{k,t,r}}{\sigma_{k,t,r}} \right)^2} dt \leq SL$$  \hspace{1cm} (17)

We have already defined the behavior of the sellers by formulating the different pricing models above, but furthermore, we have to explicitly determine the buyers’ behavior. We will implement two different types of buyers, namely short term buyers who tend to request the product last-minute and long term buyers who are more likely to reserve earlier (Shen and Su, 2007). The amount of delivered workloads over the course of a day is modeled according to (Calzarossa and Serazzi, 1985).

### 4.2 Simulation results

The first simulation set tested the performance of each seller in a monopolistic environment under different demand situations (e.g. 10 short-term buyers). The planning horizon $T_p$ of every seller has been set to 1500 what approximately equals two month when one tick is considered as a full hour. While providers have no information about demand and utilization at the beginning of each simulation run, the first ticks are not very representative, so only the time span from 300 to 1020 has been considered for the evaluation. This time span contains 720 ticks or hours what equals one month. For utilization and demand forecasting, parameters have been set to $ND = 50$, $NU = 5$ and $SL = 99.5\%$. Three possible fare classes were defined with the price vector $p_1^c$ used by the static and bid-price seller for resource 1 (Table 2). Other resources are adapted according to the relation used in the Amazon EC2. In contrast to the other models the dynamic seller needs to have a price matrix which includes prices for every resource and fare class. For the simulation, the prices have been chosen in accordance with the prices for the static models with additional ones below and above. Additionally, the threshold values $T$ have to be set to define price change behavior for every price class $c$ and resources $k$ (Table 3, setting I).

![Figure 1](https://example.com/figure1.png)  \hspace{1cm} (a)

**Figure 1.** Utilization and revenue of a dynamic provider facing 20 short and 5 long term customer for 100 timeslots (excerpt for $t = \{300, \ldots, 400\}$) (Figure a) and difference between simulation setting I and II (Figure b).
It can be observed that the dynamic seller performs particularly well in cases where demand is high. Even when demand is low, the dynamic seller achieved some acceptable outcome. The worst case, when there were only ten short term buyers, the dynamic seller had only 3.0 % lower revenue than the static seller. In the best case, he even gained 78.9 % more revenue. Also the bid-price seller performed surprisingly good and achieved as expected always higher revenue than the static seller, even when demand was low. In these cases, he even outran the dynamic seller (Figure 4a). It can be summarized that both revenue management adaptations performed very solid in the monopolistic environment and managed to significantly increase revenues in nearly any cases. However, the threshold values \( \tau \) for the dynamic seller have been adjusted on simulation results. When we chose to slightly change these values for every price class \( c \) and resources \( k \) (Table 3, setting II), the dynamic seller behaves different and results in considerably decreased revenues in most cases (see Figure 1b). Consequently, an optimally or nearly optimal adjustment of the dynamic seller has to be done what may be difficult and expensive to achieve in some cases. Figure 1a illustrates the utilization and revenue trend of a dynamic seller. The time period is exemplarily set to 300...400. We can also observe how the utilization over the time of day changes.
CONCLUSION

Cloud providers like Amazon, Google or Rackspace have to consider to whom which products are offered at a certain time to increase revenue, when resources are scarce. Future incoming requests are oftentimes unpredictable. Revenue Management concepts enable to implement a decision policy to accept or reject requests and to influence the behavior of the consumers. Price sensitive consumers will shift their demand to low-demand time, when prices are low as well. Depending on the utilization of his resources the provider can set prices or decide about accepting a request.

At first we analyze if revenue management requirements are fulfilled by a Cloud Computing scenario. The scenario comprises several products using the same pool of multiple resources. Then, we compare dynamic pricing and bid price control according their revenue outcome. In contrast to traditional Revenue Management approaches, we used a matrix to interpret a request and introduced a second time scale which allows multiple requests in one interval of the planning horizon. The proposed model allows the service providers to decide about accepting or rejecting incoming service requests based on expected revenue, expected demand and a given set of resources.

In future, the modeling of buyer according to the customer choice behavior is an interesting challenge. Then, dependencies between fare classes have an impact on the revenue. Customers, who are not allocated, could probably decide to change to a higher fare class which increases the revenue. Furthermore, the bid price as well as the dynamic algorithm can be optimized by more complex heuristics and learning algorithms.

References


A MODEL OF PREFERENCE ELICITATION: THE CASE OF DISTRIBUTED RESOURCE ALLOCATION

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0381.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Agent-based model, Economics of IS, Simulation, Grid Computing</td>
</tr>
</tbody>
</table>
A MODEL OF PREFERENCE ELICITATION: THE CASE OF DISTRIBUTED RESOURCE ALLOCATION

Stößer, Jochen, Institute of Information Systems and Management, Universität Karlsruhe (TH), Englerstr. 14, 76131 Karlsruhe, Germany, jochen.stoesser@kit.edu

Neumann, Dirk, Chair for Information Systems Research, Albert-Ludwigs-Universität Freiburg, Kollegiengebäude II, Platz der Alten Synagoge, 79085 Freiburg, Germany, dirk.neumann@is.uni-freiburg.de

Abstract

Market mechanisms are deemed promising for distributed resource allocation settings by explicitly involving users into the allocation process. The market considers the users’ and providers’ valuations to generate efficient resource allocations and prices. In theory, valuations are assumed to be known to the user. In practice, however, this is not the case. It is a complex burden for both users and providers to assess their true valuation for a certain combination of resources and services and to efficiently communicate this valuation to the market.

This paper contributes to the theory of designing distributed allocation models in that (i) we propose a model for preference elicitation, which allows users and providers to assess their valuations as a function of their resource requirements and strategic considerations, (ii) we show how this model can be encoded within so-called bidding agents which interact with the market on behalf of the user, and (iii) we evaluate our approach in a numerical experiment to illustrate how the bidding agent adapts to the dynamic market situation. As this evaluation shows, the model outperforms technical schedulers and can thus be used for decision support in electronic markets.

Keywords: Distributed Resource Allocation, Preference Elicitation, Automated Trading.
1 INTRODUCTION

In recent times, enterprises have been facing increased pressure to host resource-demanding applications that exhibit fluctuating utilization patterns. At the same time they are forced to provide flexible infrastructures at low cost. A meta-study by IT critic Nicholas Carr (2005) revealed that enterprises only utilize on average 40-50% of their data centres’ storage capacity and only 10-35% of their available processing power.

Grid and cluster technologies help enterprises deal with this dilemma. Enterprises no longer need to build up vast server farms in order to be able to serve few peak loads on their system. Instead, they only need to accommodate the basic load and tap into external computing, storage and application services on demand. The possible cost savings have been estimated to amount to 30% of total IT spending (Minoli 2004). Such grids can be located within enterprises to serve the various business units and departments, or spread across organizational boundaries to allow for dynamic resource (re-) allocations across enterprises. Prominent examples for the industry take-up of grid technology are the initiatives of Sun Microsystems, Google, IBM and Amazon. With its network.com platform, Sun offers computing resources for a fixed price of $1 per CPU hour. With Amazon’s Elastic Compute Cloud (http://aws.amazon.com/ec2) and its Simple Storage Service (http://aws.amazon.com/s3), users can configure virtual servers and remotely run their applications and store their data on Amazon’s infrastructure, only paying for what they use.

However, so far only few customers have adopted these offers. While Sun and Amazon both employ fixed pricing schemes, there is doubt whether this static model can adequately deal with the dynamic nature, which is inherent to grids as demand and supply are highly fluctuating. With fixed prices, users cannot express their priorities (i.e. valuations) for resources, thus leading to inefficient allocations and unrealized profit (Lai 2005).

Market mechanisms are deemed promising to provide a better fit to this dynamic setting by explicitly involving users into the allocation process. The market considers the users’ and providers’ valuations in order to generate efficient resource allocations and prices. Consequently, the market takes the responsibility of producing prices, which reflect the true scarcity of resources by adequately matching demand and supply. While market mechanisms exhibit compelling features, two important building blocks are missing which hampers their practical use: preference elicitation and automated bidding. It is a complex burden for both users and providers to (i) assess their true valuation for a certain combination of resources and services and to (ii) efficiently communicate this valuation to the market.

While we focus on grid settings, our results are fundamental for all market-based business relationships, where users have to formulate bids. Literature frequently assumes that the preferences are fully known to the user. We will relax this strong assumption by proposing a model for preference elicitation. As such, this paper contributes to the theory of designing distributed allocation models in that

- We propose a model for preference elicitation, which allows users and providers to assess their valuations as a function of their resource requirements and strategic considerations.
- We show how this model can be encoded within so-called bidding agents, which interact with the market on behalf of the user.
- We evaluate our approach in a numerical experiment to illustrate how the bidding agent adapts to the dynamic market situation and how even fuzzy approaches can lead to more efficient allocations than technical schedulers.

The paper is structured as follows. In Section 2, we briefly discuss previous work on preference elicitation and automated bidding. Subsequently, we introduce the tasks of a bidding agent and existing approaches to preference elicitation from the marketing research domain in Section 3. In Section 4, we present our model which illustrates one approach for how such bidding agents can combine preference elicitation and trading protocols to dynamically adapt to the situation on the market. Clearly, this model represents just one approach; others are also possible. Thus, we need to thoroughly evaluate our approach via numerical experiments. This evaluation is covered in Section 5, where we present a sample market mechanism and
show the interdependency between the mechanism, the competition in this market, and the bidding agent’s behaviour. Section 6 summarizes the paper and points to future research directions.

2 RELATED WORK

As pointed out in the introduction, we combine two aspects of the interaction between human users and the market for resources and services: preference elicitation and automated trading. Probably the largest body on preference elicitation stems from the domain of combinatorial auctions (cf. Conen and Sandholm 2001, Zinkevich et al. 2003, Parkes 2005, Nisan and Segal 2005). However, this previous work focuses on a separate issue: If users know their valuation, but communication between the users and the market is costly, how to efficiently query users for their valuations given the specific structure of the underlying allocation problem. This problem is inherent to combinatorial auctions, where the user has to submit valuations for all \(2^n-1\) possible bundles which can be composed of \(n\) goods. Different from this literature, we use the term preference elicitation to denote the users’ problem of determining their true valuation, i.e. questions such as “What am I willing to pay for using a server with application X, a dual-core processor and 2 GB of memory for one hour?” There is currently not much research available in this area, which is surprising as it is a prerequisite for any market-based approach.

Preference elicitation has been widely studied in the area of decision / negotiation support systems, where participants have to decide about their preference for an alternative or a good that is specified by multiple attributes. Typically, Multi-Attribute Utility Theory (MAUT) is used to aggregate the single attributes’ values to an overall score or ranking (e.g. Bichler and Kaukal 1999). In MAUT, the negotiator’s utility function is modelled as the weighted sum over the attributes’ utility levels. In the INSPIRE (Kersten and Noronha 1999) and NEGOISST (Jertila and Schoop 2005) negotiation systems, a so-called hybrid conjoint analysis is applied (Green and Krieger 1996) to determine the attributes’ relative importance. We defer a detailed discussion of conjoint analysis to Section 3.1. Vetschera (2007) investigates whether the preference structure of negotiators has an impact on their behaviour and the ultimate outcome.

Similar to our approach, Byde et al. (2003) also try to combine the problems of determining the value for specific resources and of allocating these scarce resources via markets in an automated manner so as to maximize the overall system value. In their setting, a utility centre operator has signed (long-term) service level agreements with various users and now has to decide how to serve these requests based on its actual situation. Byde et al. (2003) do not introduce bidding agents on the users’ side, but solely on the provider’s side. Each service request is represented by an application agent and a business agent. The application agent is responsible for monitoring and predicting the applications resource consumption. The business agent then translates this prediction into an estimated business value (from the provider’s perspective) based on the underlying service level agreement, agreed metrics, penalties, rewards etc.

MacKie-Mason and Wellman (2006) study the automation of the user-market interaction by means of trading agents. By equipping users with such (at least partially) automated tools, the communication with the market can be drastically simplified since human users do not constantly need to monitor the market outcome and update their requests. One prominent outcome of this research is the TAC trading agent competition (www.sics.se/tac/) where research teams compete in designing trading agents for a specific market mechanism.

In summary and to our best knowledge, hitherto there has been no research that tries to combine the issues of preference elicitation and automated trading from the user’s perspective in an auction context.

3 PREFERENCE ELICITATION

We propose the use of bidding agents that essentially have three levels of “intelligence”:

1. **Prediction of resource demand**: The bidding agent learns the technical requirements of the grid application. This information could either be static, e.g. the user could manually configure the agent with the required amount of processing power, memory, and storage. But the agent could also dynamically adapt
its information, e.g. by estimating the resource requirements based on historical information using statistical methods (Degermark et al. 1997, Smith et al. 1998).

2. **Preference elicitation**: The bidding agent estimates the user’s valuation for a specific application and combination of resources. This will be the core of this paper.

3. **Automated bidding**: The bidding agent interacts with the market in an automated fashion in order to obtain the right resource level – possibly according to a workflow of tasks – and considering the current market situation to achieve the best possible price.

The agent can then use information about the market outcome and the application’s resource consumption to refine its estimates. For example, after the market has cleared, the agents can refine their bidding strategies. Furthermore, the user could review the bidding agent’s preference elicitation by rating its accuracy, and the agent can compare the applications resource consumption with its prediction.

While we will focus on bidding agents for users in the remainder, note that the same principles can be applied to design bidding agents for the provisioning side of the market. We model our setting as a multi-attribute combinatorial allocation problem. A grid application is typically characterized by multiple attributes, such as the required amount of CPU, memory, storage, bandwidth, software libraries etc. The application can only be run if all such required resources are obtained in the right quality (Schnizler et al. 2008, Subramoniam et al. 2002).

More formally, we assume that an application $A$ can be characterized by a finite set of attributes $X^A = (X^A_1, X^A_2, ..., X^A_n) \in \mathbb{N}^n = \Omega$, where $\Omega$ is the space of possible application profiles.

**Multi-attribute valuation function**: Let application $A$ be characterized by the finite set $X^A$ of attributes. Then the user’s valuation $v^A$ of user $i$ for this application $A$ is a function $v^A: \Omega \rightarrow \mathbb{R}$ with

\[
v^A(x) = \begin{cases} 
> 0, & \text{if } x_j \geq x_j^A, j = 1, ..., n \\
= 0, & \text{else}
\end{cases}
\]

In the following, we will assume that the bidding agent has some estimate of application $A$’s resource profile $x^A$, and we will focus on the preference elicitation and the automated bidding as well as the iterative refinement of the estimated user valuations.

The bidding agent essentially faces two decision or learning problems:

- What is user $i$’s valuation function?
- How should the agent interact with the market to obtain the resource set $x^A$ so as to maximize its user’s benefit?

In this section, we will discuss several preference elicitation approaches which can be used to estimate a user’s valuation for a specific application and a combination of attributes.

3.1 **Conjoint Analysis**

Conjoint analysis has its roots in the 1960’s and 70’s and has become the most widely used technique in marketing research to measure consumer preferences (Luce and Tukey 1964, Green and Rao 1971, Marder 1999). Conjoint analysis allows to ask the user “What if” questions. The technique begins by defining the number of attributes and possible values, so-called “attribute levels”. The aim of the conjoint analysis is to estimate the user’s value for each attribute. The analysis essentially creates a number of attribute combinations and attribute levels, so-called “profiles” (e.g. specific products), and asks the user to evaluate these profiles, e.g. by ranking or rating the profiles. For instance, a computing server may have the attributes CPU, memory, and bandwidth. Conjoint analysis would then generate specific server profiles, e.g. (price/hour, CPU, memory, bandwidth) = ($1, 2, 2 GB, 10 MB/s), ($2.5, 4, 3 GB, 100 MB/s) etc.

The key to conjoint analysis is the “partitioning assumption” (Marder 1999), which refers to the aggregation of the values for the attribute levels makes up the user’s value for a specific profile. To simplify the analysis, one possibility is to follow the approach by Keeney and Raiffa (1976) and to use
additive independent valuation functions, where each resource (attribute) represents an additive term multiplied by a scale factor which encodes the relative importance of this resource attribute in comparison to the other resource attributes:

Additive independent valuation function: The attributes $X_1^A, \ldots, X_n^A$ are mutually independent and $v^A$ can be written as

$$v^A(x) = \begin{cases} \alpha_1^A x_1 + \alpha_2^A x_2 + \cdots + \alpha_n^A x_n, & \text{if } x_j \geq x_j^A, j = 1, \ldots, n \\ 0, & \text{else} \end{cases}$$

with scale factors $\alpha_j^A \in \mathbb{R}, j = 1, \ldots, n$.

Conjoint analysis estimates the user’s value for a certain attribute level by performing regression analyses on the user’s feedback to the presented attribute profiles.

The main drawback of this method is that the full-profile method presents all possible profiles to the user, which is clearly infeasible already for small numbers of attributes and attribute levels. Thus, as pointed out above, especially negotiation support systems employ hybrid and adaptive models, which restrict the number of profiles and mathematically determine the best profiles to present to the user (Green and Krieger 1996). Moreover, conjoint analysis relies on the partitioning assumption to be able to link the user’s evaluation of the profile to the valuations for the individual attributes. But it is problematic to assume a common relationship between the overall profile and the attributes across all users.

3.2 Analytical Hierarchy Process

In contrast to the conjoint analysis method which aims at determining the value of a certain attribute, the analytical hierarchy process tries to determine the relative importance of a certain attribute among a set of attributes (Saaty 1980). The user has to compare the relative importance of attributes in a pairwise manner based on a predefined scale from 1, 3, ..., to 9 (and $1/3, \ldots, 1/9$ respectively), e.g. “CPU is 3 times as important as memory”. Ultimately, a matrix is generated containing all pairwise comparisons, from which a vector with the various attributes’ relative importance is computed, e.g. (CPU, memory, bandwidth) = $(\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$. Saaty has also proposed a method to check the consistency of the user’s comparisons based on the matrix’s eigenvalue.

The analytical hierarchy process suffers from the large number of pairwise comparisons that the user has to perform to generate the matrix from which the relative weights are computed. With $n$ attributes, the user essentially has to do $n - 1$ comparisons. Furthermore, while the method ultimately gives the relative weights of the attributes, the question remains about the value of a specific combination of attributes.

In summary, both conjoint analyses and the analytical hierarchy process are not appropriate for practical auction-based settings with a large number of attributes and attribute values and frequent trades. Furthermore, they depend on the assumption about a specific structure of the user’s valuation function, e.g. an additive independent valuation function, which seems somewhat arbitrary.

4 THE MODEL

Thus, we propose to take an evolutionary approach. This approach is based on the assumption that a consumer who wants to buy some good does generally not exactly know his valuation for this good, i.e. if asked, the user would not be able to state a valuation. Instead, the user only has a rough estimate of his own true valuation. That is, when the consumer is confronted with an offer and a specific price, he decides whether to accept the given price, or to continue his search and look for alternative offers.

We adopt a similar approach in that the user initially indicates the request’s priority, e.g. “high”, “medium” or “low”. After each market clearing, the agent presents the results to the user, and the user may indicate whether he is satisfied with the outcome or not. Based on this user feedback, the agent iteratively updates...
its estimate of the user’s valuation. So step by step, the agent approximates the results of a full-profile conjoint analysis and successively refines its estimates.

The evolutionary preference elicitation process goes as follows:

1. **Initialization:** In contrast to presenting all possible profiles to the user a priori, the agent initially assumes a valuation $v^A_0$ for application $A$ based on the application’s resource specification $x^A$, the current price for each resource which is published by the market mechanism in a price vector $p_0 \in \mathbb{R}_{+}^n$, and an initial priority $\theta^A_0 \in \mathbb{N}$ set by the user: $v^A_0(\theta^A_0, p_0, x^A) = \theta^A_0 p^T_0 x^A$.

If the price is only published for a bundle of resources, we assume that the market imputes and publishes prices for the individual resources (cf. Xia et al. 2004).

For example, a given application $A$ requires 2 CPUs and 1 GB of memory for one minute, i.e. $x^A = (CPU, Memory, Runtime in mins)$ and $x^A = (2, 1, 1)$. CPU has last been traded for $1 per hour, consuming 1 GB of memory for one hour cost $2. The user indicates a medium priority $\theta^A_0 = 2$. The bidding agent then initiates its estimate of the user’s valuation for this task as $v^A_0(\theta^A_0, p_0, x^A) = 2 \cdot \left( \frac{51}{60} \cdot 2 + \frac{32}{60} \cdot 1 \right) \approx 0.133$.

This naïve approach might be extended by applying case-based reasoning (Hu and Haddawy 1998) or neural networks (Haddawy et al. 2003).

Subsequently, for each run $j$ of the application:

2. **Bidding:** The agent bids on the market according to its estimate $v^A_j(\theta^A_j, p_j, x^A)$ and some bidding strategy.

3. **Refinement:** There are two possible outcomes of the bidding process:
   a. **Successful:** The agent obtains the necessary resources at an overall price $p_{j+1}^T x^A = \sum_{i=1}^{n} p_{j+1,i} x^A_i \leq v^A_j$, and reports this information to the user. The user then indicates whether he is satisfied with the outcome or not, and the agent refines its estimate.

   If the user indicates that he was satisfied with this price ($\theta^A_{j+1} = 1$), the agent does not update its estimate.

   If the user indicates that the price was too high ($\theta^A_{j+1} = 0$), the agent updates its estimate of $v^A$ according to some update factor $\rho \in [0, 1]$:

   $$v^A_{j+1}(\theta^A_{j+1}, p_{j+1}, x^A) = \begin{cases} v^A_j(\theta^A_j, p_j, x^A), & \text{if } \theta^A_{j+1} = 1 \\ \rho p_{j+1}^T x^A, & \text{if } \theta^A_{j+1} = 0 \end{cases}$$

   b. **Unsuccessful:** The agent was not able to obtain the necessary resources at a price $p_{j+1}^T x^A = \sum_{i=1}^{n} p_{j+1,i} x^A_i \leq v^A_j$.

   If the user indicates that the price was indeed too high ($\theta^A_{j+1} = 1$), the agent does not update its estimate. If the user indicates that he would have preferred to pay that price rather than not getting the resources ($\theta^A_{j+1} = 0$), the agent updates its estimate of $v^A$ with the current price:

   $$v^A_{j+1}(\theta^A_{j+1}, p_{j+1}, x^A) = \begin{cases} v^A_j(\theta^A_j, p_j, x^A), & \text{if } \theta^A_{j+1} = 1 \\ p_{j+1}^T x^A, & \text{if } \theta^A_{j+1} = 0 \end{cases}$$

In our example, assume the agent was able to obtain the resources at a (total) price of $p_1 = $0.08 < $0.133 = v^A_0$ and the user indicated that this price was too high (Case a and $\theta^A_1 = 0$). Then the agent updates its estimate to $v^A_1(\theta^A_1, p_1, x^A) = 0.9 \cdot $0.08 = $0.072$ for an update factor of $\rho = 0.9$. 

Proceedings ECIS 2009
The agent will iteratively try to converge its estimate to the user’s true valuation and at the same time assist the user in identifying that valuation by presenting the user a series of simpler decision problems rather than forcing the user to directly reveal his valuation.

It is important to note that this approach does not depend on any assumption about the structure of the user’s valuation function. But similar to conjoint analysis and the analytical hierarchy process, one critique may be that the user still needs to give feedback to the agent. But obviously such feedback loops cannot be avoided if the agent is to at least approximate the user’s true valuation. However, we hypothesize that the agent can already obtain a good estimate with just a few such feedback loops. It will be an interesting question for future research if this assumption is valid, and if so, how many feedback loops are required to obtain a “good” estimate.

5 EVALUATION

As pointed out above, we hypothesize that grid users will generally not be able to state an exact number if asked for their valuation for a certain application. Instead, users will only approximate this valuation over time with the bidding agent’s assistance. This inherently precludes a “hard” evaluation of the proposed evolutionary approach to preference elicitation, as we cannot simulate this process. However, we will evaluate two complementary aspects of grid markets:

- We will show that the use of grid markets leads to efficiency gains over simple technical schedulers, even if the bidding agent uses a simple rule for initializing its estimate about the user’s valuation as presented above.
- We illustrate for a sample setting how the bidding agent strategically misreports the user’s (estimated) valuation so as to maximize its user’s benefit.

Of course the market outcome and the agent’s behavior depend on the actual market mechanism. Thus, before we turn to our analysis, we will briefly introduce a sample mechanism.

5.1 The Market Mechanism

The following mechanism has been proposed by Stößer et al. (2007) and is tailored towards usage in grids (Amar et al. 2007). Market mechanisms essentially consist of three elements: the bidding language which determines how users and providers specify their resource requests and offers, the allocation scheme which assigns resource requests to offers, and the pricing scheme which determines corresponding payments.

5.1.1 The Bidding Language

A user \( j \) who would like to submit a computational job to the grid system reports the job’s characteristics \( (v_j, c_j, m_j, s_j, e_j) \) to the market mechanism where \( v_j \in \mathbb{R}_+ \) denotes \( j \)'s valuation (i.e. \( j \)'s maximum willingness to pay) per unit of computing power and time slot, \( c_j \in \mathbb{N} \) and \( m_j \in \mathbb{N} \) the minimum required amount of computing power and memory respectively, and \( s_j \in \mathbb{N} \) and \( e_j \in \mathbb{N} \) specify the job’s estimated runtime (start and end). We require the market mechanism to make atomic allocations in the sense that each job can only be executed if there are sufficient resources available in all requested time slots. Furthermore, jobs can potentially be migrated between several compute nodes over time but each job can only be executed on one node at a time.

A provider \( n \) who would like to contribute a compute node to the grid system reports the node’s characteristics \( (r_n, \bar{c}_n, \bar{m}_n, \epsilon_n, \lambda_n) \) to the market mechanism where \( r_n \in \mathbb{R}_+ \) specifies this node’s (pretended) reservation price per unit of computing power and time slot, \( \bar{c}_n \in \mathbb{N} \) and \( \bar{m}_n \in \mathbb{N} \) the maximum amount of computing power and memory available on this node, and \( \epsilon_n \in \mathbb{N} \) and \( \lambda_n \in \mathbb{N} \) the time frame during which the node can be accessed. Given sufficient resources, we assume that each node is able to execute multiple jobs in parallel.
5.1.2 **The Allocation Scheme**

The objective of the market is to generate efficient resource allocations. This can be modelled as a combinatorial allocation problem as proposed in Stößer et al. (2007). However, due to the resulting computational complexity, Stößer et al. (2007) propose a greedy heuristic as follows:

1. The resource requests are sorted in non-ascending order of their valuations \( v_j \). Resource offers are sorted in non-descending order of their reservation prices \( r_n \).

2. The heuristic loops over the ranked list of requests and successively constructs a feasible allocation schedule by assigning the requests with the highest valuations to the cheapest offers.

5.1.3 **The Pricing Scheme**

Subsequently, the allocation schedule generated by this heuristic needs to be complemented by corresponding prices. K-Pricing was introduced in Schnizler et al. (2008). The basic idea is to distribute the welfare generated by the allocation algorithm between users and resource providers according to a factor \( k \in [0,1] \). For instance, assume an allocation of resources from a specific provider to a specific user. The user values these resources at \$10\) while the provider has a reservation price of \$5\). Then the (local) welfare is \$10 - \$5 = \$5\) and \( k \cdot \$5\) of the surplus is allotted to the user – thus having to pay \$10 - k \cdot \$5 – and \((1 - k) \cdot \$5\) is allotted to the provider – thus receiving \$5 + (1 - k) \cdot \$5\). Besides allowing for fairness considerations, the main advantage of K-Pricing is that it can be determined in polynomial runtime. On its downside, however, it only yields approximately truthful prices and payments on both sides of the market, i.e. users and providers may benefit from misreporting their valuations.

5.2 **Data Generation**

The problem with numerical evaluations of grid markets is that there is no log of real-world workloads which contains all parameters according to our bidding language specified above, in particular as regards the users’ valuations. We thus created synthetic workloads using the distributions specified in Table 1, which is in line with Feitelson (2002).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Request start time ( s_i )</td>
<td>Binomial B(5, 0.5)</td>
</tr>
<tr>
<td>Request CPU requirement ( c_j )</td>
<td>Binomial B(5, 0.5)</td>
</tr>
<tr>
<td>Request memory requirement ( m_j )</td>
<td>Lognormal L(4, 0.15)</td>
</tr>
<tr>
<td>Request valuation ( v_j )</td>
<td>Uniform U(1, 30)</td>
</tr>
<tr>
<td>Offer start time ( e_n )</td>
<td>Binomial B(4, 0.5)</td>
</tr>
<tr>
<td>Offer availability ( \lambda_n )</td>
<td>Binomial B(8, 0.5)</td>
</tr>
<tr>
<td>Offer CPU ( \tau_n )</td>
<td>Binomial B(10, 0.5)</td>
</tr>
<tr>
<td>Offer memory ( m_n )</td>
<td>Lognormal L(5, 0.2)</td>
</tr>
<tr>
<td>Offer reservation price ( r_n )</td>
<td>zero reservation prices</td>
</tr>
</tbody>
</table>

*Table 1.* Distributions used for generating the synthetic workloads.

We analyze four settings with varying degrees of competition: 20 requests and 20 offers, 40 requests and 20 offers, 60 requests and 20 offers, and 80 requests and 20 offers. For each of these four settings, we generated 200 problem instances. In the remainder, we will report the averages across these 200 runs to account for stochastic outliers.

5.3 **Data Analysis**

5.3.1 **Efficiency Gains from Preference Elicitation**

In this subsection, we analyze the performance of our preference elicitation scheme with respect to allocative efficiency, which is defined as the aggregated utility across all users and providers: How well does our preference elicitation scheme perform (i) compared to the market mechanism presented in
Subsection 5.1 if users know their valuations and truthfully submit these to the mechanism and (ii) compared to a technical scheduler which does not consider user preferences? As mentioned above, it is not possible to evaluate our iterative preference elicitation approach as such. Instead, we will evaluate the initialization step in which the agent estimates the user’s valuation based on the application’s resource requirements, current market prices, and a fuzzy priority assigned by the user.

We configure the bidding agent to initialize its estimate of the user’s valuation as follows. Since the heuristic from Subsection 5.1 ranks resource requests according to their valuation per CPU and timeslot, the bidding agent initializes its estimate of the user j’s true valuation as

$$v_{j0}(\theta_{j0}, p_0, x_j) = \theta_{j0} p_0 \cdot \text{CPU}.$$  

We assume that users can distinguish between three job priorities \{1,2,3\}, e.g. “high”, “medium” and “low” priority. The initial user priorities are (recall that we draw the true valuations from U(1, 30))

$$\theta_{j0} = \begin{cases} 1, & \text{if } 1 \leq v_j \leq 10 \\ 2, & \text{if } 11 \leq v_j \leq 20 \\ 3, & \text{if } 21 \leq v_j \leq 30 \end{cases}$$

The effect is that, assuming all agents apply this initialization rule, the requests get clustered into three blocks, requests with \(v_{j0}(\theta_{j0}, p_0, x_j) = p_0 \cdot \text{CPU}\). \(v_{j0}(\theta_{j0}, p_0, x_j) = 2p_0 \cdot \text{CPU}\), and \(v_{j0}(\theta_{j0}, p_0, x_j) = 3p_0 \cdot \text{CPU}\). For example, let job j require one CPU and 100 MB memory for three timeslots. The user’s true valuation is $3 per timeslot. One CPU has last been traded for $1 per timeslot. Then the user indicates an initial priority of 1, and the agent estimates a valuation of \(v_{j0}(\theta_{j0}, p_0, x_j) = 1 \cdot 3 = 1\). Ultimately, assuming truthful behaviour, the agent reports \((v_{j0}(\theta_{j0}, p_0, x_j), c_j, m_j, s_j, e_j)\) to the market.

Technical schedulers only aim at maximizing resource utilization or balancing the system load. As a proxy for such technical schedulers we implemented a scheduler which randomly assigns jobs to feasible nodes that offer sufficient resources in the right timeslots.

The numerical results with respect to allocative efficiency are reported in Table 2.

<table>
<thead>
<tr>
<th>Number of requests and offers (ratio of allocated requests)</th>
<th>Efficiency with true valuations</th>
<th>Efficiency with initialized preference elicitation</th>
<th>Efficiency of randomized (technical) scheduler</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 requests, 20 offers (97.3%)</td>
<td>3.687.08</td>
<td>3.679.11 (99.8%)</td>
<td>3.664.66 (99.4%)</td>
</tr>
<tr>
<td>40 requests, 20 offers (82.4%)</td>
<td>6.852.75</td>
<td>6.693.97 (97.7%)</td>
<td>6.013.25 (87.7%)</td>
</tr>
<tr>
<td>60 requests, 20 offers (61.6%)</td>
<td>8.734.67</td>
<td>7.885.82 (90.3%)</td>
<td>6.561.02 (75.1%)</td>
</tr>
<tr>
<td>80 requests, 20 offers (48.8%)</td>
<td>9.776.49</td>
<td>8.366.91 (85.6%)</td>
<td>6.716.23 (68.7%)</td>
</tr>
</tbody>
</table>

We successively increased the competition in the market, indicated by the first column. With 20 requests and 20 offers, almost all requests get allocated. But with 80 requests and 20 offers, only 48.8% of the requests can be accommodated. The economic performance of technical schedulers will decrease as competition in the market increases. If competition is low, technical schedulers will be able to accommodate most of the requests, so the negative effect of not considering the users’ valuations is comparably small. But obviously this negative effect increases with the scarcity of the resources. This general reasoning is confirmed by our results. For the most competitive setting with 80 requests and 20 offers, the technical scheduler only achieves 68.7% of the efficiency which the market mechanism achieves with true valuations. Interestingly, despite the simple initialization rule and without considering the iterative refinement of the agents’ estimates, the market mechanism combined with bidding agents still achieves 85.6% of this benchmark. This result underlines the case for considering economic principles in the allocation of grid resources, even if the hard informational assumptions regarding user valuations are relaxed.

5.3.2 Strategic Considerations with Automated Bidding

In this subsection, we will illustrate the general idea for how bidding agents can implement strategic behaviour when interacting with the market in order to maximize the users’ benefit. As outlined in Section 3, this is a learning task in itself which is complementary to preference elicitation. We assume the bidding...
agent to have an estimate \( v_{j0}(\theta_{j0}, p_0, x_j) \) about user \( j \)’s true valuation. We model the one-dimensional action space of the agent to consist of a factor \( \beta \in \mathbb{R}_+ \). Consequently, the agent reports \( (\beta \cdot v_{j0}(\theta_{j0}, p_0, x_j), e_j) \). The question is how the agent should choose \( \beta \).

As is common in agent-based simulations, this choice can be modelled using reinforcement learning, e.g. Q-learning (Watkins 1989). This is basically a trial-and-error approach, where the agent tests different strategies (i.e. choices of \( \beta \)) during the exploration phase. After this exploration phase, the agent chooses the best strategy according to some policy, e.g. the \( \beta \) which resulted in the highest utility (the so-called epsilon-greedy selection policy). Then, during the exploitation phase, the agent plays this best strategy. In our setting, the agent may initially submit the true (estimated) valuation \( v_{j0}(\theta_{j0}, p_0, x_j) \) to the market. Then it successively explores its strategy space by deviating from this true valuation, e.g. by only bidding 50% of its true estimate, 55%, 60%, up to 150%.

The results of this strategy (\( \beta \in [0.5, 1.5] \)) for our numerical evaluation are illustrated in Figure 1, where we again employed the heuristic above complemented by K-Pricing with \( k = 0.5 \). The graphs show the utility relative to truthful bidding for the four settings.

![Figure 1. Efficiency gain from misreporting.](image)

The results show that, for each setting, it is the best strategy on average to report the true estimate. Intuitively, the higher the competition, the higher the risk of not being allocated anymore if the agent reports a valuation below its true estimate, thus yielding a smaller utility for the user. The discrete priority values \( \theta_{j0} \in \{1, 2, 3\} \) are reflected in the step-like shapes of the relative utility when underbidding \( (0.5 \leq \beta < 1) \). Assuming all agents apply the same initialization step with \( n \) discrete job priorities, we get \( n \) clusters of requests with identical \( v_{j0}(\theta_{j0}, p_0, x_j) \) in the ranking of the allocation heuristic. Consequently, as soon as \( \beta \cdot v_{j0}(\theta_{j0}, p_0, x_j) \) falls within the next lower cluster in the ranking, we get a drop in the relative utility.

If the agent overbids, there are two possible consequences. If the request was allocated with truthful bidding, then the request will still be allocated if the agent reports a higher valuation. However, with K-Pricing the user will receive a smaller fraction of the true welfare, and thus have a smaller utility. If the request was not allocated with truthful bidding, the request either remains outside of the allocation if the agent reports a higher valuation, thus leaving the user’s utility unchanged, or the request gets allocated, yielding a positive utility as long as \( \beta < 2 \) with zero reservation prices (cf. K-Pricing rule). However, as the results show, the overall effect of overbidding is slightly negative on average. This property is important, as agents have no incentive to misrepresent their priority.
Overall, the results indicate that the agent should truthfully report $v_{j0}(\theta_{j0}, p_{a}, x_{j})$ during the exploitation phase in our experimental setting.

6 CONCLUSION & OUTLOOK

The idea of applying market mechanisms to the allocation of scarce resources in dynamic and distributed settings is certainly not new. In such settings, markets exhibit many desirable features, in particular they promise to lead to more efficient allocations than traditional technical approaches. However, we argue that markets are facing two main obstacles:

- Users will generally not exactly know their true valuations for the resources, but only have a fuzzy estimate.
- Users cannot be expected to continuously monitor and interact with the market, but this interaction is costly. For instance, human users incur opportunity costs and have bounded perceptivity.

Surprisingly, these two barriers have largely been neglected in the vast body of literature on market mechanisms. Instead, existing work simply assumes perfect information on the users’ side and costless interaction with the market. The aim of this paper was to take a first step towards overcoming this obvious gap between theory and real-world settings. We briefly discussed previous work on preference elicitation and automated bidding in Section 2. In Section 3, we introduced the learning tasks of bidding agents and existing approaches to preference elicitation from the field of marketing research. At the core of this paper, we proposed a new model for preference elicitation and automated trading in distributed settings. To evaluate the initialization step of this model and the strategic behaviour of bidding agents in more detail, we performed a numerical experiment with a sample market mechanism, whose results we reported and interpreted in the previous section. In summary, these results underline the case for considering market mechanisms in the allocation of distributed resources, even if the hard informational assumptions regarding user valuations are relaxed. Instead, we propose the use of bidding agents, which (i) assist the user in finding his true valuation using an evolutionary approach, and (ii) shield parts of the underlying infrastructure’s and the market’s complexity from the user by acting on his behalf.

This work suggests several natural extensions for future research. The preference elicitation task of the agent requires further analyses, in particular as regards the iterative refinement of the agent’s estimates. How should the agent be parameterized, e.g. as regards the update parameter $\rho$ and the discrete user priorities $\theta_{j0}$? As pointed out above, the use of case-based methods (Hu and Haddawy 1998) and neural networks (Haddawy et al. 2003) might be interesting approaches. And how many iterations does this process take to get within x% of the user’s true valuation? Moreover, the automated trading of bidding agents offers interesting research questions. How does the agent behaviour depend on the specific market mechanism and vice versa? And what exploration and exploitation strategies should such agents apply? These questions can certainly not be answered with our rather static analysis, but call for agent-based simulations.

References


Smith, W., I. Foster. and V. Taylor (1998). Predicting Application Run Times Using Historical Information. IPPS/SPDP 1998 Workshop on Job Scheduling Strategies for Parallel Processing, Orlando, Florida, USA.


MAKING MONEY WITH CLOUDS: REVENUE OPTIMIZATION THROUGH AUTOMATED POLICY DECISIONS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0588.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Decision making, Economics of IS, Grid Computing, Pricing</td>
</tr>
</tbody>
</table>
MAKING MONEY WITH CLOUDS:

REVENUE OPTIMIZATION THROUGH AUTOMATED POLICY DECISIONS

Püschel, Tim, Albert-Ludwigs-Universität Freiburg, Chair for Information Systems Research, Platz der Alten Synagoge, 79085 Freiburg, Germany, tim.pueschel@is.uni-freiburg.de

Arun Anandasivam, Universität Karlsruhe (TH), Institute of Information Systems and Management (IISM), Englerstraße 14, 76131 Karlsruhe, Germany, arun.anandasivam@iism.uni-karlsruhe.de

Stefan Buschek, Universität Karlsruhe (TH), Institute of Information Systems and Management (IISM), Englerstraße 14, 76131 Karlsruhe, Germany, stefan.buschek@student.kit.edu

Neumann, Dirk, Albert-Ludwigs-Universität Freiburg, Chair for Information Systems Research, Platz der Alten Synagoge, 79085 Freiburg, Germany, dirk.neumann@is.uni-freiburg.de

Abstract

Many existing IT applications exhibit strongly varying demand patterns for resources. Accommodating an ever increasing and highly fluctuating demand requires continuous availability of sufficient resources. To achieve this state at reasonably costs, a high degree of flexibility with respect to the given IT infrastructure is necessary. Facing this challenge the idea of Cloud computing has been gaining interest. In so-called Clouds resources such as CPU, storage and bandwidth can be bundled into a single service, which is offered to Cloud users. These services can be accessed in oblivion of the underlying IT infrastructure. This way Cloud Computing facilitates the introduction of new products and services without large investments in the IT infrastructure.

Cloud Computing is a promising approach with a high impact on business models. One aspect of business models is clearly the revenue model, which defines how prices should be set to achieve predefined revenue level. The decision about accepting or denying requests has a high impact on the revenue of the provider. In this paper we analyze two approaches that support the cloud provider in its decision. We show that predefined policies allow increasing revenue compared to widely used technical models such as first-come-first-serve.

Keywords: Cloud Computing, Revenue Management, Decision Mechanisms
1 INTRODUCTION

Many existing IT applications (e.g. data mining, portfolio analysis or numerical simulations) exhibit strongly varying demand patterns for computing resources. Demand for computing resources is further increased by the introduction of new products and processes. Accommodating an ever increasing and highly fluctuating demand requires that sufficient resources are available. Building an adequate infrastructure requires significant investments in the infrastructure and instructing and hiring of personnel. However, many enterprises face increasing competition on global markets. This leads to the challenge to make use of new applications, reduce process times, and introduce innovative services for customers on one side and cut the costs of their IT-infrastructures on the other side (Carr, 2005).

To achieve this goal it is necessary to maintain a very high degree of flexibility with respect to the IT infrastructure. Facing this challenge the idea of Cloud computing has been gaining interest lately. The term Cloud Computing was initially coined by Amazon.com, which were among the first companies to offer Cloud services on a large scale. In the Cloud resources (e.g. processing power, storage and bandwidth) can be bundled as services, which are offered to Cloud users. These Cloud services can be accessed without knowledge of the underlying infrastructure. Cloud Computing also allows the introduction of new products and services without large investments in installing or upgrading of the IT infrastructure. The New York Times managed to convert 4TB of scanned images containing 11 million articles into PDF files in very little time using Hadoop, a framework for distributed applications and Amazon’s Cloud services. The actual conversion process took only 24 hours cost merely 240 US-$. New Projects and websites like Justin.tv or SmugMug.com would not have been able to adapt to such a rapidly growing user base without the help of Cloud services. Don MacAskill, CEO of SmugMug, estimated his company saved about 339,430 US-$ in 7 month, spending only 84,255 US-$ on Amazons Simple Storage Service. According to analysts at Gartner “The projected shift to Cloud Computing will result in dramatic growth in IT products in some areas and in significant reductions in other areas.

In order to remain attractive for customers Cloud services must meet user demands concerning price, availability and quality of service. The more providers offer their services, the more likely it is that they can be accessed in quantity and quality. Therefore it is important to attract more providers. However, Cloud providers will only offer their services if they can realize a sufficient level of profit. Providers have to plan their resource usage carefully and be aware of dynamic changes of the incoming requests for their services. With state-of-the-art technology, this assimilation is hampered, as the components facilitating the deployment of the services are not designed to incorporate economic issues (e.g. price, expected demand). To improve performance in the commercialization of distributed computational resources decisions about the supplied resources and their management should be based on both technical and economic aspects (Kenyon and Cheliotis 2004).

In recent times, several research projects have started to develop price-based resource management components supporting the idea of grid or utility computing, which could be applicable for Cloud providers in some cases. Those approaches are devoted to scheduling by using price mechanisms such as auctions. Clearly, this means that technical issues such as resource utilization are ignored for scheduling. Furthermore, these approaches require a tight integration with the infrastructure of the provider, as well as the interface to the customer. In cases where a provider offers a wide range of services and customers might use different providers at the same time, it is often not feasible to undertake a complex and time consuming bidding process. Technical resource management systems typically offer the possibility to include priorities for users. Purely price-based schedulers do not distinguish important from less important partners, as only current price matters for the allocation. In this work we propose two mechanisms which increase revenue of Cloud providers by deciding which job or service requests are accepted and or rejected. The first mechanism provides an elegant way to introduce utilization based pricing in real world systems. Integration of concepts from revenue
management in the second mechanism allow increasing overall revenue even in fixed-price scenarios which dominate Cloud Computing so far.

This paper is novel because it proposes mechanisms for fixed price, as well as pay-as-you-bid scenarios which can be used in real-time and advance reservation scenarios. At the same time our mechanisms can be implemented in real world scenarios with relatively few changes to existing architectures. With increasing competition among Cloud providers work covering this topics is highly relevant in the research area of information systems. The paper is structured as follows. In the next section we introduce motivational scenarios and the requirements they impose. Subsequently we present work related to our model. Section 4 contains a description of our approaches and the models they are based on. An evaluation based simulations follows in section 5. The paper ends with a conclusion and brief outlook of future work.

2 MOTIVATIONAL SCENARIOS AND REQUIREMENTS

In this section we present three major real world scenarios, which motivate the use of automated policy decisions. The scenarios impose different requirements on the decision making unit concerning the timeframe, which is available for determining the decision.

2.1 Motivational Scenarios

Scenario I: A service provider offers a number of different cloud services on the market, similar to like Salesforce.com or mor.ph. Most of these services have to be accessible within very short time spans. Each service has a different profile of resource requirements. Some service are computationally intensive, others require large amounts of memory or storage. Currently, the provider has no mechanisms to adequately deal with utilization approaching maximum capacity. All incoming jobs are accepted as long as there is capacity left. Because demand for the different services vary greatly situations where one type of resources (e.g. CPU) operate at nearly full capacity while others still have a significant buffer left arise regularly.

Scenario II: A company Y has a lot of computationally intensive tasks in their business processes. They have built and maintain a number of large datacenters providing different services for their own business processes. Due to the nature of their business utilization of the resources is very uneven. The company has accumulated important knowledge in the management of their infrastructure and can provide the services efficiently. Providing these services has become a core competence for the company. Therefore, they consider offering these services to external customers as a new product. This allows them to enter a new business segment, achieve a more even utilization of their resources and at the same time take further advantage of economies of scale. However, the sale of services to external users must not have any negative impact on their own business processes. Especially, it should not result in decreased resource availability for internal users. Because of this, jobs users from within the organization should always be preferred when deciding which jobs are accepted.

Scenario III: The service provider offers a wide range of cloud services for different areas of application, such as data mining, numerical simulations or rendering. Like in the first scenario the profile of resource requirements varies greatly between these services. However, in contrast to scenario I the service provider has more time to take a decision whether to accept a job or not, since resources are reserved in advance. Furthermore jobs are significantly larger and cost more, justifying the introduction of more complex optimization methods. We assume that the provider has not the option to dynamically change his prices. He rather has fixed prices like Joyent.com offering a subscription model or Flexiscale.com with a pay-as-you-go model. An important aspect is to consider the time of arrival, when jobs come into the system. The request time for a service has a great impact on the forecasting models and the revenue expectations. We assume that large jobs are often well planned and scheduled more in advance than small jobs.
2.2 Requirements

From these scenarios the research question arises “What pricing mechanisms can be used to increase provider revenue?”. These mechanisms have to consider different circumstances such as real-time decisions or further constraints like a strict minimum capacity available for internal users. Naturally there are myriads of options whose study covers a very wide research area. In this paper we investigate mechanisms, which try to increase revenue by accepting or denying service requests. We consider fixed price situations as well as simple pay-as-you-bid mechanism. We restrict our analysis to these types of mechanisms as they are relatively simple to implement in real world systems.

Scenario I imposes some requirements on a mechanism. To effectively deal with this scenario a mechanism must have very low computational cost; it simply needs to calculate the decision within a very short time span. Furthermore, the mechanism should take utilization of each resource type into account, when deciding whether or not a job is accepted. For this scenario we assume the application of a pay-as-you-bid mechanism. Scenario II adds the need to guarantee one class of users a certain capacity or implement some kind of prioritization of some users. The requirements for the third scenario are slightly different. In this case there is more time to take the decision. Therefore a mechanism can do a more complex optimization as long as it leads to higher revenue. Scenario III uses a fixed price mechanism. Using the requirements of all three scenarios it is obvious that it would be very difficult - if not impossible - to find a single mechanism which excels in all situations. Therefore we present two different mechanisms which are deemed promising to work in our scenarios. There exists no one-size-fits-all-mechanism. Efficiency is widely used as a requirement and metric. However, Cloud providers’ main focus is on revenue not efficiency. Since they decide which mechanisms to implement we focus on revenue optimization in this paper.

3 RELATED WORK

There are two main streams of work existent that deal with the problem of admitting job requests. The first one considers technical and rule based optimization. The second one investigates the use of concepts from Revenue Management.

Technical and rule based optimization stems from the computer science perspective, as we will see in the following: Ferguson, Nikolaou, Sairamesh, and Yemini (1996) discuss the application of economic theories to resource management. Aiber, Gilat, Landau, Razinkov, Sela, and Wasserkrug (2004) present an architecture for autonomic self-optimisation based on business objectives. Elements of client classification such as price discrimination based on customer characteristics have been mentioned in other papers (Newhouse et al. 2004 and Buyya 2002). They did however not consider other discrimination factors. Chicco, Napoli, and Piglione (2006) describe data-mining algorithms and tools for client classification in the electricity grids but concentrate on methods for finding groups of customers with similar behaviour. An architecture for admission control on e-commerce websites that prioritizes user sessions based on predictions about the user’s intentions to buy a product is proposed by Poggi, Moreno, Berral, Gavaldà, and Torres (2007). Boughton, Martin, Powley, and Hormann (2006) present research on how workload class importance should be considered for low-level resource allocation.

One approach to realize advance reservation and end-to-end quality of service is the Globus Architecture for Reservation and Allocation (Foster et al. 1999). This approach uses advance reservations to achieve QoS. Another way to achieve autonomic QoS aware resource management is based on online performance models (Kounev et al. 2007). They introduce a framework for designing resource managers that are able to predict the impact of a job in the performance and adapt the resource allocation in such a way that SLAs can be fulfilled. Both approaches do not consider economic factors such as pricing.
As resources, e.g. processing power, storage and bandwidth, are usually bundled as services in the Cloud, providers have to plan their resource usage carefully and be aware of dynamic changes of the incoming request. Although, advance reservation diminishes the unpredictable requests, users can still submit important jobs on-demand and cancel queued jobs or their reservation (no-shows). The provider can apply revenue management strategies to enhance revenue and optimally allocate his resources to the consumers. He can set booking limits for his services and accept a certain amount of customers. An acceptance strategy is required due to the competition for the resources by the services. The resources in a cluster are accessed by different services. It is important to know which service request should be accepted now to gain enough revenue in the future. The described problem is an instance of the dynamic inventory network capacity control (NCC) for finite time horizon $T$. Therefore work about Revenue Management concepts for IT is a key part of our related work.

The first paper analyzing Revenue Management concepts for cluster systems was published by Dube et al. (2005). In the suggested model one resource is offered for different prices. By assuming the customer behavior follows a logit model, the authors analyzed an optimization model for a small number of price classes and provided numerical results.

Cancelations and no-shows reduce the efficiency of resource usage. Sulisto et al. (2008) analyzed how overbooking strategies can be applied to maximize revenue. Different prices were charged for one resource. Three overbooking policies were implemented and compared via simulation. The benefits of overbooking for shared hosting platforms were emphasized by Urgaonkar (2002) as well. He did not optimize the revenue by classifying different services, but only the throughput rate.

Anandasivam and Neumann (2009) presented a framework for applying Revenue Management in Grid Computing. They gave an introduction and outlined some requirements, which have to be fulfilled. Their theoretic model comprises bundles of resources and shows how they can be priced. This framework can be adapted for Cloud Computing as well. However, none of these papers analyzes capacity planning strategies of resource bundles via simulations.

Nair and Bapna (2001) introduced Revenue Management concepts for a similar application domain: Internet Service Provider. The provider has to decide whether to accept an incoming customer or to reject. The application domain is different from Cloud Computing as it does not take advance reservation and bundles into account. Customers can only instantly get an internet access.

Policy based decision modeling and capacity planning for fixed prices and advance reservation have been partially studied. Nevertheless, in the context of cloud computing, requirements are different than the domains studied in the literature before. The settings were already outlined in the scenarios. In the next section we present our own model comprising QoS and client classification for online allocation and decision support for advance reservation.

4 \hspace{1cm} \textbf{THE MODEL}

One of the key questions for resource providers is whether a job should be accepted at the time and for the price for which it was submitted. The standard approach used nowadays is to always accept a job if there is enough capacity available to fulfill it. Capacity is allocated on a first come first serve basis. However, since our main objective is to maximize the revenue of the provider those jobs that achieve this objective should be chosen. This can be done by solving the following maximization problem:
Where $T$ is the set of all timeslots; $J$ is the set of available job requests; $p_i$ is the price paid for job $i$; $x_i$ is a binary allocation variable indicating whether job $i$ was accepted or rejected; $c_{ir}(t)$ is the capacity required by job $i$ in timeslot $t$; and $gc_r(t)$ is the total capacity available of resource $r$ during timeslot $t$.

The problem can be formulated and solved as a linear program. However there are two problems when choosing this approach. The first one is that this is an instance of the knapsack problem, to be more specific the temporal knapsack problem. As such it is NP-hard and therefore computationally intractable. Furthermore to solve the problem, it is necessary to have the information on the future jobs, the capacity and runtime they require, the prices, etc. In general this information is not available before a job is submitted.

We propose two approaches to solve this problem. The first approach uses policies as a heuristic. This approach has very low computational costs, decisions can be calculated very fast and useful policies can be already generated with very limited information about future demand. Therefore this approach can be used in a very wide range of scenarios, even real-time scenarios where decisions have to be taken immediately. The second approach introduces more complex mechanisms used in classic revenue management. It requires more time and resources for calculations than the first approach as well as a forecast for the demand. Therefore it cannot be used in such a wide range of scenarios as the first approach. In return, however, in those scenarios where it can be used it promises higher revenue than the first approach.

4.1 Policy based heuristic

The first approach is based on previous work presenting an earlier version of an architecture for an economically enhanced resource manager (Püschel et al. 2007). The idea is to use policies as a heuristic for maximizing the revenue without having the exact information about future jobs. These policies can be based on requirements from Service Level Agreements (SLA), company policies and information gained from historic workload traces, utilization curves, prices, etc. If available, such policies can also be based on forecasts or information about the statistical distribution of job size, runtime and prices. This approach is especially well suited for motivational scenario I and II.

In this work we will extend two aspects of our policy based approach. The first aspect is dynamic pricing based on various factors, such as projected demand or utilization. The second aspect is job priority based on client classification. The idea behind utilization based pricing is that when utilization is high demand for the resources is high as well and therefore customers are willing to pay higher prices for resources. If utilization is low lower prices are charged to attract more customers.

In scenarios where various services with different profiles of resource requirements are offered the (reservation) price can be calculated based on the overall utilization or on the utilization of each resource type (e.g, CPU, memory, storage). If prices are calculated based on utilization of the individual resource types those services that need a lot of capacity of scarce resources have a higher price increase than services which only need little. Such a pricing mechanism reduces the risk that some resources have a lot of free capacity but no jobs can be accepted because one type of resources is used to capacity.

For certain users it may be very important to always have access to the resources. This is the case in our third motivational scenario; sufficient resources have to be available for jobs from internal users. To achieve this a type of client priority on job acceptance can be used. When the utilization of the system is low jobs from all classes of clients are accepted but when the utilization of the resources rises and there is competition between the clients for the resources, jobs from certain clients are preferred. There can be two types of priorities: strict priorities and soft priorities. Strict priority means that when a certain utilization threshold is reached only jobs from clients with this priority are accepted. Soft priority means jobs from clients with priority are generally preferred but standard clients have the chance to outbid clients with priority. Therefore clients can still submit jobs when the threshold is reached as long as they pay a higher price.
To keep the system adaptable it includes a Policy Manager, a component which stores and manages the different policies, such as policies concerning pricing and client classification. Policies are formulated in the SWRL, the Semantic Web Rule Language (Horrocks et al. 2004). SWRL combines the Web Ontology Language (OWL) and the Rule Markup Language (RuleML) and is de facto standard concerning rules languages that offer semantic integration. A simple example rule from a pricing policy in SWRL is the following rule which expresses that if the utilization is between 71% and 100% there is a surcharge of 5 on the general reservation price:

\[
\text{Utilization(?utilization)} \land \text{InsideUtilizationRange(?utilization, "71\% - 100\%")} \\
\Rightarrow \text{SetSurcharge(?utilizationsurcharge, "5")}
\]

The overall architecture for an economically enhanced resource manager is implemented in a prototype in an EU project; the extensions presented in this work are implemented in a simulator.

4.2 Bid price optimization

4.2.1 General description

Each offered service represents a booking class, which has a fixed price. The provider has to decide, if a service request should be accepted or denied. Thus, a limit defining how many requests are operable for each booking class has to be identified, which is known as capacity control. Nested booking limits allow preventing that bookings for services with high revenue are being rejected in favor of bookings with low revenue. They define how much capacity is reserved for a certain booking class. Every service has limited access to the resources like CPU, storage or bandwidth. Due to multiple resources a nested booking limit control must be defined for each resource. This is called virtual nesting control (Talluri and van Ryzin 2004). It is difficult to forecast demand appropriately for virtual classes. The requirement of mapping services to virtual classes increases the complexity. Furthermore, the assumption that demand for low-class services occurs earlier than high-class products is quite static, and when demand is stochastic the strict low-to-high order is less appropriate. It is assumed that customers book the services for usage at a certain time in the future. The low-before-high arrival is expressed by a high booking probability of low-fare services at the beginning of the booking period. This probability decreases during the time period. Contrarily, the probability of booking high-fare services is low at the beginning of the booking period, and increases until the end of the time period.

For the application of bid-price controls, at any point in time a simple threshold value for each resource has to be stored. Bid prices are interpreted as an approximation for the opportunity cost (Bertsimas and Popescu 2003) for reducing the resource capacities, which are needed to satisfy incoming service requests. Moeller et al. (2007) describes bid prices as monetary values of single capacity units, and the resource demands of a request weighted with the corresponding bid prices must be summed. If this sum exceeds the revenue yielded by the sale of one unit of the respective product, the request is rejected, otherwise it is accepted. Regular updating of bid-price values is necessary to guarantee a continuous precision of the bid prices. Less accurate bid prices can lead to accept/reject decisions of minor value. Continuously updated bid prices are based on the current booking situation at a certain point in time \( t \in \{ T, T-1, \ldots, 0 \} \). That is, if a high amount of capacity has already been sold, the bid prices turn out to be higher.

The decision of accepting or denying a request depends on the policy how the decision is made. Capacity control and dynamic pricing known from Revenue Management are an instance of linear programming models. Since speed of computation matters, bid price control is an approximation method to define policies very fast. It provides a good estimate, but not always an optimal solution. Especially in the NCC setting, the calculation of the optimum increases exponentially with the number of resources \( m \) and products \( n \) (Talluri and van Ryzin 2004). In the following we assume that the provider has three resources \( h \in \{1\ldots m\} \) with \( m=3 \) available and offers three services \( i \in \{1\ldots n\} \) with \( n=3 \). For example resources are CPU, storage and bandwidth, whereas services can be low medium and high instances like the Amazon Web Services. Resources have to be quantifiable. The
The bid-price function is based on two parts. The first part ($\alpha_h \cdot c_{ht})$ is responsible for the increase of the respective bid price over time. $c_{ht}$ is the amount of capacity of a resource $h$ reserved at time $t$. If a request for a service $i$ is accepted, the bid price of a resource increases by the delta of the value $c_{ht}$, which equals $a_{hi}$ capacity units, and on the value of $\alpha_h$. Thus, the bid price of a resource $h$ only is only increased by the acceptance of incoming requests. This corresponds to the fact that the amount of available resources decreases due to sales and hence become more expensive.

The second part of the formula ($\beta_h \cdot u_{hi}$) decreases the bid price for every occurring request. A decrease is required because if some requests are rejected, some future requests can be accepted again. If there was no decrease in the function, the bid price only would rise, and from a certain point every future request would be rejected, and no more sales could take place. $u_{hi}$ is the capacity required to satisfy the demand for the products $i \in A_h$ until $t$. The demand until $t$ for a product $i \in A_h$ can be calculated by $D_{iT} - D_{hi}$. It requires forecasts of the total demand per product $i$ ($D_{iT}$), as well as forecasts of the demand-to-come ($D_{hi}$) for every point in time $t$ until the end of the booking period. $u_{hi}$ is calculated by

$$\sum_{i \in A_h} a_{hi} \cdot (D_{iT} - D_{hi})$$

For every product its demand for resource $h$ $a_{hi}$ is multiplied with the expected demand until $t$ $D_{iT} - D_{hi}$, and the sum of these products is taken. The values of $u_{hi}$ increase over time as more demand
is realized. The bid-price decreases with time proceeding. The two parts of the resource-oriented bid-price function make sure that the total value of a bid price $\pi_{ht}$ increases only, if a request is accepted. The increase amplifies if the amount of reserved capacity $c_{ht}$ is high.

The parameters $\alpha_h$ and $\beta_h$ attach importance to variables. They have a strong impact on the accept/reject decisions. For instance, very high values for coefficient $\alpha_h$ and very low values for coefficient $\beta_h$ lead to more frequent reject decisions because the increase of the bid-price function turns out too high. This would imply losses in revenues due to rare sales. In the opposite case, very low $\alpha_h$ values and very high $\beta_h$ values result in an "accept all" policy, and thus, it can happen that capacity is mostly sold to low-fare classes (requests for low-level services) also leading to potentially lost revenues. Because of these reasons promising values for the control variables should be obtained by simulation-based optimization.

An enhanced approach to adopt the bid-price function is to artificially raise the bid prices depending on the current degree of reservation. Thereby, low-class requests occurring at the beginning of the booking period can be rejected explicitly in order to reserve capacity for requests yielding higher revenues. This kind of policy can be realized using a set of variables $\gamma_i$ which depend on the average of the degree of reservation denoted by $z_t$ calculated each time a request occurs by

$$z_t = \frac{\sum_{h=t}^{\infty} z_h}{t}.$$ Given $z_t$, the provider can set utilization thresholds which specify when to use a specified $\gamma_i$, e.g. as follows:

- If $z_t < 0.4$: use values $(0.12, 0.02, 0.01)$ for $\gamma_i, i \in \{1,2,3\}$.
- If $0.4 \leq z_t < 0.75$: use values $(0.02, 0.025, 0.05)$ for $\gamma_i, i \in \{1,2,3\}$.
- If $0.75 \leq z_t \leq 1$: use values $(0.005, 0.05, 0.1)$ for $\gamma_i, i \in \{1,2,3\}$.

$\gamma_i$ is added to control variable $\alpha_h$, thereby increasing the bid prices.

$$\pi_{ht} = (\pi_{ht})^\alpha_h \cdot (\alpha_h + \gamma_i) - \beta_h \cdot \frac{c_{ht}}{U_{ht}}$$

In the case when demand strongly fits to the assumptions stated above, this contributes to more frequent rejections of service-1 requests at the beginning of the booking period, and to more frequent rejections of service-3 requests close to the end of the booking period.

5 EVALUATION

For the evaluation we considered the following three types of services with requirements for processing power, memory, and storage according to table 1. For each of the proposed mechanisms a scenario matching the requirements from the motivational scenarios was created and a simulation was run. The same service types and their resource profiles were used in both cases. The first simulation was based on real workloads, for the second mechanism a very challenging scenario using fixed prices fixed runtimes and relatively constant rates of arrival were used.

<table>
<thead>
<tr>
<th>Service</th>
<th>CPU</th>
<th>Memory</th>
<th>Storage</th>
<th>Mean Price/Fixed Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>4</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>16</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>8</td>
<td>8</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 1. Types of services and their resource profiles.
5.1 Policy based heuristic

The approach was evaluated using generated workloads as well as workloads based on real world scenarios, from the Parallel Workload Archive (Feitelson 2007). The setting of the evaluation matches the requirements for the first motivational scenario. The runtimes and offsets between submit and start time were generated using a lognormal distribution. For the workloads based on real world scenarios, the SHARCNET log, which was graciously provided to the Parallel Workload Archive by John Morton (john@sharcnet.ca) and Clayton Chrusch (chrusch@sharcnet.ca), was used. The SHARCNET log was chosen as a basis because it contains a large variety of jobs with different runtimes, numbers of used CPUs, and varying submit and start times. Job runtimes where rounded down to full hours to allow a timeslot based allocation. After filtering invalid jobs and those jobs which were considered too long or too short 566,701 jobs were left and used for the scenario. For each of these jobs a service type was randomly drawn. Based on these jobs ten joblists with different prices were generated. A capacity of 8200 for each of the resource was used to be able to accommodate some of the larger jobs from the workload. A simple pay as you bid mechanism was assumed and pricing information was generated using a normal distribution with the mean price from table 1 and variance of 0.5.

<table>
<thead>
<tr>
<th>Policy</th>
<th>FCFS</th>
<th>Utilization</th>
<th>Strict Priority</th>
<th>Soft Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. Revenue</td>
<td>54190907.26</td>
<td>88697433.21</td>
<td>80480832.63</td>
<td>82579338.5</td>
</tr>
<tr>
<td>Percentual Increase</td>
<td>63.38%</td>
<td>48.51%</td>
<td>52.39%</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Revenue increase using policy based approach.

As can be seen in table 2 the average revenue the benchmark, a simple first-come-first-serve (FCFS) policy, was 54190907.26. A policy using utilization based pricing managed to increase the revenue on average by 63.38% resulting in an average revenue of 88697433.21. Introducing strict priority without charging preferred users for the privilege obviously reduces the achieved revenue. However since the preferred users in our scenario are internal users it is reasonable that they get the privilege without a charge. The average revenue for a policy combining utilization based pricing with strict priority we achieved in our simulation is 80480832.63 which is an increase of 48.51% compared to the FCFS policy. Since soft priority allows outbidding preferred clients revenue is higher than in a policy using strict priority, however this results in slightly decreased resource availability for preferred users. A policy with soft priority and utilization based pricing resulted in a revenue increase of 52.39% compared to FCFS. Runtime for one simulation run with 566,701 job requests varied between 55 and 60 seconds, each job is accepted or rejected within fragments of a second. This makes the mechanism suitable for scenarios where job acceptance has to be decided in less than a second.

5.2 Bid price optimization

For the evaluation of bid price optimization we assumed the setting in scenario II. Instead of prices drawn from a normal distribution the prices were set according to table 1. Naturally, this leaves significantly less room for revenue increase compared to scenario I. Thus the bid price approach focuses more on forecasting and capacity planning. The bid prices are in every timeslot. Parameters like the current amount of capacity reserved, the total expected demand and the already realized demand are taken into account. For each incoming request at time t, the bid prices are updated. Bid prices increase, if the amount of reserved capacity rises due to the acceptance of requests. They decrease, if incoming requests are rejected and no additional capacity is being reserved. Given the bid prices, an approximation of the opportunity cost is calculated by weighting the resource demand of an incoming request with the bid prices. If the revenue yielded by the sale of one unit of the respective service exceeds this approximation, the request is accepted. This simulation setting differs from the policy-based approach described above, since the bid price optimization considers advance reservation of resources, while in scenario I resources are instantly consumed. The capacity of a resource provider was fixed to a maximum of 750 units for every resource. The simulation was performed with a
booking period of length T=2000 for five job files. The SA-BP model yielded an average revenue of 1860.5. The average revenue of a FCFS policy was 1810.08 (increase of 2.79%). In the best case the revenue increase was 11.26%. The performance of SA-BP is based on the accuracy of the forecasting data. With inaccurate forecasting models, the revenue of SA-BP can even be lower than the FCFS. The runtime of SA-BP is similar to the FCFS computation. This is due to the heuristic approach of SA-BP. More complex revenue management models require long runtimes to calculate the optimal allocation. In both policies the runtime for 2000 jobs were about 6 seconds. In practice, a longer time horizon is necessary for calculation, which will increase the runtime exponentially.

The reason why SA-BP performed better than FCFS was due to rejections of most service-1 requests in the first third of the booking period. Thus, more capacity was kept free for later arriving service-2 and service-3 requests. The SABP3 policy accepted more service-1 requests in the second third of the booking period. Furthermore, the acceptance of service-3 requests decreased over time, thus enabling more accepts of later occurring service-1 requests. The FCFS policy accepted all incoming requests in the first third of the booking period until the capacity was completely reserved. SA-BP rejected early arriving service-1 requests, and thereby, later arriving requests for services 2 and 3 could be accepted.

5.3 Discussion

The evaluation shows that the proposed mechanism can increase revenue. The policy based mechanism fulfils all requirements imposed by the first two motivational scenarios. Revenue is significantly increased and computational costs are very low. The concept of strict priorities enables Cloud providers to guarantee availability of a predetermined amount of capacity for certain users. Unless these users are charged for this privilege it results in a revenue decrease compared to the pure utilization based pricing. Soft priorities can be used to mitigate this effect; however, they can result in a reduction of resource availability for preferred users. The bid prize optimization mechanism can be used to increase revenue in motivational scenario three. Even in a challenging setting with fixed prices and the same profit margin for all services it still managed to deliver revenue increases. Computational costs are higher than for the policy based approach but still acceptable for many scenarios.

6 CONCLUSION

Cloud Computing is a promising approach with a high impact on business models. One aspect of a business model is revenue and how prices should be set to achieve predefined revenue goals. The decision about accepting or denying requests has a high impact on the revenue of the provider. Revenue Management concepts known from the airline comprise interesting policies. In this paper we analyze two approaches. Predefined policies allow increasing revenue compared to static models like first-come-first-serve. Such policies include utilization based pricing as well as client classification.

We furthermore present how to apply more complex revenue management concepts in Clouds. When price are relatively fixed, an optimization of the capacity over bid price concepts seems to be a promising approach. A bid price optimization enables to determine the minimum price a consumer has to pay for requesting a service. The complexity arises when services utilize the same resources. The bid price is applicable for a scenario, where advance reservation is possible. Our proposed policies and models were validated via simulation. For the first of our proposed mechanisms we were able to use real world workload traces. By using these real workload traces the models are applicable in practice to some extent. However we still had to generate pricing information, therefore complete external validity of our results cannot be assumed. We showed that well-defined pricing strategies can notably increase revenue. Future work will comprise an implementation in existing middleware considering architectural design and technical feasibility. Moreover, the heuristics and policies have to take learning algorithms into account to analyze, whether it will lead to better results.

References


A TEACHING CASE: TOWARDS BRIDGING DISCIPLINARY DIVIDES IN IT EDUCATION

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0488.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Teaching Case</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Learning, Education, Applied learning, Teaching Case</td>
</tr>
</tbody>
</table>
A TEACHING CASE: TOWARDS BRIDGING DISCIPLINARY DIVIDES IN IT EDUCATION

Abstract

Information Technology (IT) now plays a central role in almost all aspects of business. However to best realise the potential of IT, it is vital that educators adopt strategies that challenge rather than reinforce existing disciplinary divides in IT education and business. This paper supports the application of active learning, via the use of a serial teaching case, towards bridging disciplinary divides in IT education. Specifically, the case calls for the design, development and implementation of an Integrated Management Information System. Students are therefore required to demonstrate analytical as well as technical skill-sets in the areas of IT, accounting and business. The paper describes the business context facing AHN Ltd. and illustrates the pedagogic use of the case as a means of integrating IT with aspects of accounting and business knowledge, towards helping educators contribute to answering the calls from academia (Panko, 2008) and industry (Chan and Reich, 2007) for graduates with a hybrid of business and IT skill-sets.

Keywords: Business Case, Active learning, Linkage, Knowledge.
1  INTRODUCTION

Active learning approaches, such as case-based learning and problem-solving, have been advocated as ways of fostering deeper learning (Boyce et al., 2001; Biggs, 1994). Organizations such as Microsoft and Financial Services have been using problem-solving scenarios such as simulations to both test and train employees. Use of such scenarios and role play will through real world problem-solving enhance the learner’s logical reasoning, numeric abilities and spatial thinking. Realising the potential of such methods however requires active engagement from educators and learners alike (Baskerville, 2008; Healy and McCutcheon, 2008). For many educators willing to engage in such methods, the lack of appropriate materials and guidance is often perceived as a barrier. This paper presents one approach to active learning, via the use of the AHN Ltd. case study.

AHN Ltd. is a serial teaching case based on the ongoing analysis of AHN Ltd., a manufacturer of electronic toys. The case, described in the next section and discussed in section 3, is used as a fictional business scenario whereby students assume the role of an entity tendering for and then offering a prototype IT solution to meet the business needs of AHN Ltd. The materials are structured in a manner whereby a given instructor can focus on one particular aspect or disciplinary perspective (e.g. that of a programmer or of a management accountant); alternatively the case can be used more generally, to link IT to business, thereby mitigating the possibility of silos or islands of knowledge being created around the individual modules of a given programme of study. Use of the case also facilitates a key pedagogical component of any learning experience: that of real world problem-solving, where the problem is not phrased like a question from the textbook.

Problem-solving skills require the use of a number of different learning strategies and types of knowledge. The learner’s own experience, internal mental models, and other ‘cognitive structures’ are necessary to ‘construct’ their own knowledge when faced with new information or different situations (Baskerville, 2008; Reeves, 1997). This case is organised to provide the student with a detailed description of the company. Attention is drawn to the inadequacies of the current IT infrastructure to enable students to identify and develop solutions to address these. The case has been successfully used on a module in a postgraduate programme in IT and Management Accounting. Used in total or in part the case illustrates and develops the practical as well as the theoretical alignment of business and IT. Individual lecturers can decide based on resource constraints and student capabilities to make elements we have included in the teaching note available to their students, for example Figure 1 or multiple components of Figure 2. Section 4 describes, in detail, the approach used to teaching the AHN Ltd. case.

2  THE TEACHING CASE: AHN LTD.

AHN Ltd. makes electronic toys. In recent years production and sales levels have grown rapidly and without corresponding attention being paid to the information systems (both financial and non-financial) within the organization. What was once a small operation now has a turnover of €35 million and close to 100 employees. AHN Ltd wishes to embrace the philosophy of ‘continuous improvement’ throughout all its operations and administrative activities, stream-lining manufacturing processes, reducing costs and improving quality and customer satisfaction. A web presence is also under serious consideration.

There are three main product lines: ‘Space Contender’, ‘Corporate’ and ‘Executive’. Details of the three products are as follows:

- **Space Contenders**
  This a hand-held game device aimed primarily at the Christmas toy market; thus demand is seasonal, peaking in the months of September, October and November. The ‘Space Contenders’ line is a cost
leader in the marketplace and a replacement/substitute for a highly popular imported product. AHN Ltd. produces the product to stock, primarily to use up excess resources as they arise. The current sales price is €20 per unit.

- **Corporate**
  This is an executive hand-held game device, consisting of a basic plastic console unit for which individual game modules can be purchased. Basic units are mass-produced, and then individually tailored to customer requirements. A key aspect is the customization of the accompanying gaming software. AHN Ltd. offers its customers the opportunity to incorporate their logos and promotional features into the games, providing them with an unusual and innovative means of advertising. Selling price is dependent on the quantity of an order and the amount of customization required. The ‘Corporate’ line is sold mainly, though not exclusively, through corporate advertising agencies to large multinational clients. Orders are large and demand is fairly constant throughout the year.

- **Executive**
  The ‘Executive’ line is similar to the ‘Corporate’ model, but cheaper in price. Distribution is mainly through office supplies wholesalers.

The current manufacturing facility was purpose built to accommodate the initial products, ‘Space Contender’ and ‘Corporate’. A state-of-the-art plastic moulding machine was installed to produce unit casings. Polymer resin compound is moved from the raw material storage area and delivered to the moulding machine, where it is processed into moulded casing units. A second moulding machine was purchased to facilitate production of ‘Executive’ casings. Finished casings are then stored until required for assembly. Production of finished units takes place along three modern assembly lines. The process of loading software onto each of the products is largely automated and currently operates at 60% capacity. Completed units are then sent to Quality Control for Inspection/Testing. Defective products are returned to the assembly area for re-processing. Checked and passed Finished Goods are placed in pre-printed boxes and stored for shipment. The entire warehouse area is fully racked and spacious, and complies with all current Health and Safety regulations.

2.1 **Sales Activity**

AHN Ltd. has three markets regions: Republic of Ireland, Northern Ireland and the U.K. Five full-time sales persons service this market. Each sales representative is provided with a company car and paid a basic salary, plus commission (calculated monthly), equal to 5% of gross sales revenue. AHN Ltd.’s total customer base of approximately 1,760 clients includes an estimated 310 dormant accounts. Sales Invoices number (on average) 1,100 per month.

Sales Orders are posted, phoned or faxed by customers or Sales Representatives to the Sales Support/Customer Services personnel. Orders received are written up manually on Sales Order/Dispatch Forms. These forms come in pre-printed books with four carbon copies. Original Sales Order/Dispatch Forms are passed (physically) four times daily, to the Production Director. Copy orders are passed to the Sales/Marketing Director, the Accounting Office and the final copy is retained in the Sales Order/Dispatch Form Book. The Production Director passes orders for in-stock finished goods items to the Credit Controller for Credit Checking. Sales Orders passed by the Credit Controller are then forwarded to Dispatch and given to warehouse personnel to use as a picking list. Filled orders are deposited in the Dispatch Area and warehouse personnel sign off the Sales Order/Dispatch Form. The Dispatch Clerk also signs off on the Form and makes two photocopies of it. The original is enclosed with the goods for delivery; one copy is placed on file in Dispatch and the other photocopy is passed to Accounting for entry into Invoicing on the accounting system. Invoices are batch processed and update takes place every evening. During update the system can be accessed.
only for information purposes – no data can be entered or edited. Invoices are printed daily and posted to customers. Statements are printed at the end of every month and posted to customers. Sterling invoices however are typed separately and are currently not entered into the accounting system at all. At the end of each accounting period, sterling sales activity is accounted for via a series of journal entries to the monthly accounts.

Goods returned by customers are checked by the Storeman and written up on a pre-printed (triplicate) Goods Returns/Credit Claim Form describing the reason for claim: typically breakages, rejected goods, wrong location. The original document is sent to the customer, one copy is kept in the Stores and the other copy is passed to Accounting for processing. Goods Returns/Credit Claim Forms are entered into the accounting at the end of each month before customer statements are printed and posted. During peak periods the paper trail can break down. Before further expansion can take place this shortcoming must be addressed.

2.2 Production Activity

Production scheduling is done on a weekly basis, using an excel spreadsheet, and is normally based on budgeted sales for the period. The Production Manager gives a (verbal) daily production report to the General Manager. Production output is also recorded on Output Sheets, which are filled in twice daily and used by the Storeman to maintain finished goods stocks in the accounting system. Invoicing updates automatically update finished goods stock levels. However the present system ignores items salvaged from damaged reworked product returns returned to finished goods stocks for re-sale.

AHN Ltd. deals with 149 suppliers. Where possible, raw materials are sourced locally. All transactions are conducted in Euros. Purchase Orders for raw materials are created manually using pre-printed Purchase Orders. AHN Ltd. generates an average of 180 purchase orders per month and issues approximately 162 cheques each month, although the organization does hope to implement paperless transactions within two years in compliance with Irish Payment Services Organization (IPSO) regulations. Authorized Purchase Orders are posted to suppliers, with a copy held by the Storeman for comparison purposes when goods are received. Typically however the Storeman signs for all goods received as being ‘received unchecked’ given the distance between his office and the goods inward section of the production facility. At the end of each shift, the Storeman uses the delivery dockets to enter the receipts of that day on the accounting system. He then forwards the delivery dockets to the Accounting office.

2.3 Existing Information Systems

AHN Ltd. is equipped with a dated Nixdorf computer system that currently runs linked Invoicing, Payroll, Accounts Receivable, Accounts Payable, and General Ledger on a group of four networked terminals. When the system is in ‘end-of-period’ mode, data cannot be entered or edited and system querying is limited and slow. It takes approximately four hours to run all the reports (assuming all goes well and the run is not accidentally interrupted); therefore the company does not generate reports on a weekly basis. Payments received from customers are manually recorded in the Cash Book on a daily basis and entered from the Cash Book onto the Nixdorf system at the end of each month prior to generation of end of month reports. Monthly customer account statements are printed, in duplicate, with the original posted to the customer and the copy kept on file in the Accounting office. A sales report is also generated at the end of each monthly accounting period, listing all invoices and credit notes printed from the system. However the system offers no facility to analyse this data. Recently an accounting intern used the printout to generate sales analysis reports (using Excel) of the profitability of major customers and of the total commission payments associated with these sales.

The R&D department is equipped with two powerful workstations used for computer-aided design (CAD) and word-processing. All other departments have stand-alone PCs using Microsoft packages.
AHN Ltd. finds the Microsoft products to be user friendly and is attracted to the free download of upgrades. Many of AHN’s administrative personnel have successfully completed the EDCL course and the company is, therefore, favorably disposed to the use of Microsoft Access and VB.Net for the development of future computer systems.

2.4 The Need for Change: Invitation to Tender

AHN Ltd. wishes to update its existing system (along with the informal paper-based recording that has built up around it) and convert to a more modern, user friendly, flexible, open platform system. The company has issued an ‘Invitation to Tender’, inviting suitably qualified organizations to respond. Tenders are sought for the investigation, design, development and implementation of an Integrated Management Information System (hereafter referred to as an IMIS). The new system should be menu driven, be user friendly with an on-line help facility, and have security and control systems including password levels at Operating System Level and Application Level, to control user access and system audit trails. Live demonstration of a prototype of the proposed system will also be required of all those short-listed in the tendering process. Specifically, contracting organizations must:

- Prepare a full analysis of the requirements of the IMIS Project;
- Outline proposed software development, implementation and project management techniques;
- Submit a profile of proposed project team members and the role(s) they will play;
- Demonstrate an understanding of potential problems and risks involved;
- Attend regular meetings with AHN Ltd. personnel, providing progress reports and demonstrating work done on the prototype to date;
- Present a working prototype solution of their proposed system, including pro forma statements of all paper-based elements;
- Provide supporting documentation, including ‘trouble shooting’ advice and a users’ manual.

3 DISCUSSION AND CONCLUSIONS

In recent years, research has increasingly advocated active learning strategies (Boyle et al., 2001; Biggs, 1994). Active learning strategies will initially challenge most students. However careful introduction can and does offer benefits even to those who were not originally technically oriented. Research on the use of cases as a teaching strategy indicates that the difficulties that can arise relate to the application of active learning methods, rather than with the method in and of itself (Healy and McCutcheon, 2008). Our experience of using the case presented in this paper demonstrates that such difficulties can be overcome, provided lecturers maintain a two-way communication flow with students. The emergent nature of the eventual case ‘solution’ must also be recognised from the outset and viewed as an opportunity for (rather than a constraint to) the on-going development of the eventual prototype solution.

One of the key advantages of the AHN Ltd case is its hybrid orientation towards IT and Business. This dual focus facilitates students in developing a broader range of competences and an increased understanding of two key dimensions of software development: the technical dimension concerned with how to develop systems; and the business relevance dimension concerned with what features to develop. In our use of the business case, staff from the disciplines of Management Accounting and Information Systems are equally involved and play the role of users whose requirements must be met by specific features of the software. The problem-solving approach offered via the AHN Ltd case also generates a multiplier effect which results in students demonstrating a much greater familiarity and expertise with the development platforms they use than the content of formal lectures and tutorial sessions provided to them would otherwise offer. Working on the case study also gives students the occasion to understand the need for hybrid IT-Business expertise. For example in the area of systems development methods, the difference between theoretical knowledge and practical understanding is readily visible, as students re-label existing methods and mix them together to suit the needs of the
project requirements. Some groups are attracted to rapid application development (RAD) type methods and claim to deliver quicker, cheaper software, while others swear allegiance to traditional systems development life cycle (SDLC) and play the purity game. The search for a ‘single right answer’ becomes forgotten as students focus on continuously improving their own solution to the problem posed in the case description. Additionally, for our students, this has led to greater confidence and greater likelihood of attracting potential employers in interview situations.

Classes are divided into teams (with all of the dynamics of a group of developers) and asked to form consultancy companies with the ultimate goal of winning a tender. The teams compete with one another to develop a system for a customer, AHN Ltd., which is a represented by staff teaching on the degree programme utilizing the case to integrate business and IT topics. Therefore learners experience a broad range of tasks and problems inherent in commercial software development (Baskerville, 2008; Yang et al., 2008). The communication that must develop between the developers and their users of the AHN Ltd. prototype is quite realistic and reveals the gaps in understanding that can arise between individuals from different functional backgrounds. This aspect of the case gives students practical experience of the difficulties in discussing and presenting technical concepts to non-technical users. Students are forced to realize, acknowledge and understand the integration of materials taught in distinct stand-alone modules, often by staff from a number of disparate departments. Students are also encouraged to utilize their existing knowledge instead of waiting for the materials to 'come to them' as typically happens in a traditional lecture-based learning environment. The resulting skills-set developed through such activity meets the calls for same in both professional and academic sources (Panko, 2008; Chan and Reich, 2007). In conclusion, the AHN Ltd case described in this paper provides a useful vehicle for enabling lecturers to support students towards bridging disciplinary divides in IT education.

4  TEACHING APPROACH

The case study was designed to facilitate and support understanding and learning of the links between IT and its business applications. The company (AHN) description and tendering process presents students with an opportunity to play the role of practitioners tendering for the contract to develop an IMIS system. Topics covered through the AHN Ltd. teaching case include: analyzing IS requirements; developing the enterprise database; creating interfaces; designing reporting systems and project management. Table 1 outlines the how these topics can be taught and assessed through the AHN Ltd. case. The paragraphs that follow discuss how we have used the case in our own teaching.

At the start of the teaching period, the class is divided into teams of 4-6 members, depending on the size of the group. We ask for volunteer team leaders at the first class meeting; then once each class member has introduced themselves and summarized their academic background in terms of areas of prior studies, the team leaders take it in turns to select team members. Each group is expected to meet at least once a week. Project Teams are asked to record minutes of all meetings and these must be kept for review by the co-ordinators throughout the year (there may be little advance notice of reviews given). Microsoft Project is used to co-ordinate the work effort; thus planning of tasks and workloads well in advance will form an essential element of the overall assignment.

The case is structured to run as a series of workshops (see Table 2) to gradually build knowledge of the subject domains while simultaneously simulating 'real world' situations when the groups are asked to deliver a series of requirements to determine their level of understanding of the business case. At the end of each workshop the goals for the next workshop are set, based on the level of knowledge and understanding demonstrated by students up to that point. When presenting their deliverables at the beginning of each workshop it fast becomes evident if they do not fully understand the content delivered or required. This enables a post-mortem evaluation approach (Kasi et al., 2008). Workshops also provide an opportunity for students to evaluate each others work – in the past, we have
approached this by asking individual groups which aspects of each others systems they would most covet and which they would recommend making changes to. The pace of the work will determine the intervals required between workshops. Typically we hold workshops every three weeks, over the course of a 24-week teaching time.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Through-lines</td>
<td>How to you address programming / business problems?</td>
</tr>
<tr>
<td></td>
<td>How do you identify the IT/Acc requirements of a business?</td>
</tr>
<tr>
<td></td>
<td>How should you address analysis requirements?</td>
</tr>
<tr>
<td></td>
<td>How can you use what you have learned to build a solution?</td>
</tr>
<tr>
<td></td>
<td>How do you use two disciplines like IT and Accounting to solve a business problem?</td>
</tr>
<tr>
<td></td>
<td>How do you use business communication skills?</td>
</tr>
<tr>
<td>Understanding Goals</td>
<td>Students will understand the process of ‘Forming fictitious companies’</td>
</tr>
<tr>
<td></td>
<td>Students will understand how to identify the requirements of a business and to apply a divide and conquer approach to a problem enabled through the provision of a ‘Tender Document’</td>
</tr>
<tr>
<td></td>
<td>Students will understand how to develop a Database with a VB.NET front-end</td>
</tr>
<tr>
<td></td>
<td>Students will learn to tender for a winning contract and compete against their peers through the delivery of a ‘Final Prototype’</td>
</tr>
</tbody>
</table>

**Table 1:** Topics Taught Through the AHN Ltd Case

The final workshop serves as a showcase for the end result of the process (Table 2 – Industry Panel). As the lecturers co-ordinating the game, we do not participate in the discussion and evaluation process during this workshop. However, we ask some of our colleagues to attend this presentation, in lieu of the senior management members of AHN Ltd – thus simulating a real-world tendering process and enhancing the overall learning experience of the students. We also invite relevant professionals from industry to attend and evaluate the IMIS solutions. In all instances, this has served as both recognition of the student inputs throughout the learning process and as a showcase of student capabilities. The final presentation also brings closure to case and a sense of achievement to students and lecturers alike.
<table>
<thead>
<tr>
<th>Content (what)</th>
<th>Methods (how)</th>
<th>Purposes (why)</th>
<th>Forms (communication)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Workshop 1:</strong> Group Dynamics Project Management</td>
<td>Identify the core issues in working &amp; managing groups &amp; in forming a consultancy firm.</td>
<td>To resolve group issues &amp; identify the roles necessary to solve the problem at hand. To engender healthy competition between class groupings.</td>
<td>Groups record meetings using MS Project; actions are planned and communicated. Competition arises via the drive to meet the games requirements.</td>
</tr>
<tr>
<td><strong>Workshop 2:</strong> Accounting principles</td>
<td>Apply principles of Accounting to the ‘Forms’ as required for the AHN Ltd study.</td>
<td>To understand the fundamentals of accounting. To be able to build a useful Accounting system.</td>
<td>Groups present their solutions to the accounting requirements during workshops &amp; consider the programming translation through a user interface &amp; code. Environmental considerations are also addressed through ‘what if’ scenarios are posed during Q&amp;A sessions.</td>
</tr>
<tr>
<td><strong>Workshop 3:</strong> IT Analysis &amp; Development methods</td>
<td>Describe the processes outlined in a problem area &amp; in aligning the proposed solution (Accounting &amp; Programming).</td>
<td>To visualize a problem &amp; determine the requirements of the users. The application &amp; use of ‘best practice’ approaches for building a solution. To understand the integration of IT with the business context.</td>
<td>Groups document the requirements in both a report &amp; through presentations. Each group’s analysis is reviewed by instructors &amp; their peers, offering suggestions for improvements.</td>
</tr>
<tr>
<td><strong>Workshop 4:</strong> Designing accounting Reports</td>
<td>Describe the processes outlined in a problem area &amp; in aligning the proposed solution.</td>
<td>To integrate the disciplinary areas &amp; increase students’ knowledge through the application of the material taught.</td>
<td>Templates are presented to the instructors who as the users of the game provide feedback.</td>
</tr>
<tr>
<td><strong>Workshop 5:</strong> DB Development ERD Techniques</td>
<td>Construct &amp; implement a feasible project plan Create &amp; populate Data set</td>
<td>To use development knowledge in determining the relationships to the identified entities &amp; building a solution. To test the types of relationships &amp; data dictionaries defined.</td>
<td>Students present a series of proposed ERDs editing &amp; updating their solutions as feedback is received from both peers &amp; lecturers. DB are presented &amp; documented with versions amended to the requirements of the different audiences.</td>
</tr>
<tr>
<td><strong>Workshop 6:</strong> VB.NET programming Problem-solving</td>
<td>Differentiate between the possible problem-solving techniques &amp; business solutions.</td>
<td>To apply a ‘divide &amp; conquer’ approach to the problem. To understand the different applications of code (reusing code)</td>
<td>Solutions are presented &amp; discussed. The testing process is also documented as well as ‘what if’ scenarios through error trapping.</td>
</tr>
<tr>
<td><strong>Workshop 7:</strong> Demo of prototype 1.0</td>
<td>Develop working prototypes utilizing the following applications: Access 2003 &amp; VB.NET.</td>
<td>To enable students to adopt a trial &amp; error approach incorporating the end user in the development process.</td>
<td>Prototypes are presented &amp; evaluated by end users emphasizing different functionalities to particular groups of users.</td>
</tr>
<tr>
<td><strong>Workshop 8:</strong> Demo of prototype 2.0 Business communication: presenting, report writing.</td>
<td>Recommend &amp; present a final solution to the business problem.</td>
<td>To simulate a real business scenario. To develop presenting &amp; reporting skills. To understand &amp; value of these skills</td>
<td>Present to internal as well as external representatives of the AHN company to meet the requirements of an audience with different backgrounds &amp; roles within the fictitious company.</td>
</tr>
<tr>
<td><strong>Industry Panel Presentation &amp; Review</strong></td>
<td>Groups presents final prototypes</td>
<td>To simulate a ‘real-world’ tendering process.</td>
<td>Present to practitioners (represented by industrial and academic practitioners).</td>
</tr>
</tbody>
</table>

**Table 2: Workshop Structure**
The IMIS consists of four sub-systems; Purchasing, Manufacturing, Inventory, and Sales/Ordering. Figures 1, 2 and 3 illustrate a sample solution. The focus in this instance is the Sales/Ordering (Figure 1). However this does not preclude selective inclusion of any or all of the other components. Lecturers are therefore enabled to select content which best matches with their overall teaching goals and student learning outcomes as well as the resource constraints of the learning context. Tables 1 and 2 illustrate the capabilities of the case as a pedagogic device - but allow individuals to also tailor their contents as appropriate. For example, different groups within the class could be assigned particular sub-systems to complete.

Table 3 provides a suggested grading structure totalling 200 marks (based on a module worth 10 ECT credits). Each Project Team is assessed not just on their final project submissions but also on performance over each stage and ‘deliverable’. The emphasis is on the learning process throughout the duration of the exercise, feedback is always provided following each scheduled presentation (for post-mortem evaluation). Again, individual instructors can tailor this scheme (or abandon it totally) to meet their own requirements.

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>25% (50 marks)</td>
<td>Programming: VB.Net.</td>
</tr>
<tr>
<td>25% (50 marks)</td>
<td>Database development, DB structures and design issues.</td>
</tr>
<tr>
<td>25% (50 marks)</td>
<td>Accounting Issues.</td>
</tr>
<tr>
<td>12½% (25 marks)</td>
<td>Project Management.</td>
</tr>
<tr>
<td>12½ % (25 marks)</td>
<td>Overall Professionalism / presentation / etc.</td>
</tr>
</tbody>
</table>

**Table 3: The Grading Structure**

In conclusion, we have found the use of the AHN Ltd case to be an enriching, rewarding experience for all participants. Each year, we as co-ordinators learn something new from the process, as well as enabling our students to attain levels of understanding and appreciation for the underlying subject areas that have heretofore been difficult (if not impossible) to achieve in more traditional teaching contexts. Soft skills in terms of presenting, team-work, problem-solving and conflict resolution are embedded into the delivery of the case solution. The unanimous sense of satisfaction and achievement anecdotally reported by our students (despite the initial difficulties they encountered in the learning process) has convinced us of its power and capability to aptly serve as a means of bridging disciplinary divides in IT education.

**References**


Figures

Figure 1: Proposed IMIS
Figure 2:  Sales/Order Process
Figure 3: IMIS Logical Entity Relationship

Proceedings ECIS 2009
THE ROLE OF TRAINING IN DECREASING ANXIETY AMONG EXPERIENCED COMPUTER USERS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0532.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Technology adoption, Skills, Human computer interaction (HCI), Learning</td>
</tr>
</tbody>
</table>
THE ROLE OF TRAINING IN DECREASING ANXIETY AMONG EXPERIENCED COMPUTER USERS

Sokura, Bertta, Helsinki School of Economics, POBOX 1210, FI-00101 Helsinki, Finland, bertta.sokura@hse.fi
Tuunainen, Virpi Kristiina, Helsinki School of Economics, POBOX 1210, FI-00101 Helsinki, Finland, virpi.tuunainen@hse.fi
Öörni, Anssi, Helsinki School of Economics, POBOX 1210, FI-00101 Helsinki, Finland, anssi.oorni@hse.fi

Abstract

Surprisingly, in addition to inexperienced computer users, also those who have used different software applications at least to moderate extent can have feelings of anxiety with their use of computers. This paper examines the role of training in decreasing anxiety among experienced computer users. More specifically, the role of training is studied in relation to computer anxiety, behavioral beliefs and self-efficacy, which are the key mental constructs training may impact. 96 adult and university students, who attended a CDL course on voluntary basis, participated in this longitudinal study. The results suggest that training is an effective tool to decrease anxiety and promote self-efficacy even among experienced users. In addition this study calls in the question of the relation of computer anxiety and ease of use.

Keywords: training, anxiety, Computer Driving License, ease of use, self-efficacy
1 INTRODUCTION

Personal computers have been on the market and on the desktops of employees already for a quarter of a century. Today PCs are also an essential domestic appliance with broadband channel connections. Although the ubiquitous role of information and communication technology (ICT) in people’s everyday lives has grown significantly, average users’ skills to exploit basic computer applications efficiently are still low, and many people feel nothing less than anxiety when having to use computers (Kohrman 2003; Tekinarslan 2008). Computer anxiety can be defined in terms of a psychological response, such as, computer phobia, or in terms of a cognitive reaction, that is, apprehension of using computers (Jancour, Sinclair et al. 1994). As our focus is on more experienced computer users, we will use the latter definition in this study.

Computer anxiety is widely thought to impact early perceptions about the ease of use of a new system (Venkatesh and Davis 2000; Chee Wei, Sutanto et al. 2006). Hence, computer anxiety influences indirectly also the intention to use the system. As computer anxiety exerts such a negative influence on intention to adopt information technology attenuating its effects should remain a high priority. It is commonly argued that training decreases anxiety, yet, empirical evidence on the effect of training is ambiguous. To better decrease levels of computer anxiety among average users, we should possess a clearer picture on the effects of our main tool, training, on anxiety, and that is, precisely, the aim of this study.

In this paper, we report the findings of a longitudinal study on the effect of training on computer anxiety and behavioural beliefs about use of the target system. The study explores remedying the impact anxiety exerts on the perceptions of ease of use and usefulness related constructs. We are especially interested in how training, in our case a CDL-course, influences these perceptions in the presence of anxiety. CDL, or Computer Driving Licence, is a standardized examination developed for providing basic computing skills. It is a hands-on test, including the most commonly used software, such as, word processing, spreadsheet, graphics, databases, and graphical operating systems.

Our main research questions are as follows: 1) How does CDL-training influence user acceptance of information systems among relatively experienced users? 2) Does CDL training decrease computer anxiety? We will also test the relative importance of self-efficacy, as it has been suggested to exert a stronger effect on computer anxiety than that of experience (Wilfong 2006)

The structure of this paper is as follows: In section 2, we introduce the key constructs of anxiety, training, self-efficacy, perceived ease of use, perceived usefulness and intention to use, and introduce our research model. In Section 3, we briefly describe the empirical data collection method, and then present the results of the analyses in Section 4. Summary and conclusions, as well as directions for future research, are discussed in Section 5.

2 ANXIETY AND TRAINING

The rapid development of new computer applications and continuing software upgrading require lifelong learning. Many users remain uncomfortable using new applications, and often revert to using traditional methods to accomplish tasks that could be performed quicker or more efficiently with some software or software features. Because of feelings of uncertainty or anxiety, users remain unaware of many useful computer applications on the market and software features of the applications in use.

In the following, we will review earlier literature on computer anxiety and constructs that have been found to have effect on it, and formulate our research hypotheses based on the earlier research.
2.1 Anxiety (ANX)

We understand computer anxiety in terms of a cognitive reaction that manifests itself as apprehension of using computers (Jancour, Sinclair et al. 1994). In practice, this means that people experience uncertainty, and they are a little uncomfortable working with computers. Anxiety, then, can be viewed resulting from the beliefs an individual holds, rather than as an antecedent to these beliefs (Saade and Kira 2007).

Overall, results of many studies indicate that computer anxiety is an important predictor of the use of technology (Czaja, Charness et al. 2006; Tung and Chang 2007). Although the results of the earlier studies regarding interdependencies between anxiety and behavioral beliefs (discussed in more detail in section 2.4) are inconsistent, anxiety clearly seems to have negative effects on performance of the users.

2.2 Training

It is commonly agreed that training decreases anxiety (Igbaria 1990; Torkzadeh and Angulo 1992; Popedavis and Vispoel 1993; Torkzadeh and Koufteros 1993; Martocchio 1994; Shelley 1998; Beckers and Schmidt 2001). Effects training has on anxiety have been studied both in mandatory and voluntary settings, focusing mainly on particular new technologies, and users with limited experience. Earlier studies have found that increasing computer experience may help reduce computer anxiety (Igbaria 1993; Chang 2005), and that anxiety level decreases also while computer knowledge increases (Kay 2008). On the other hand, it has also been argued that computer self-efficacy beliefs have a stronger impact on computer anxiety than computer experience or use do (Wilfong 2006). Thus, we posit that training will decrease computer anxiety also among experienced users, whose earlier computer training is limited or missing, and we formulate our first hypothesis as follows:

Hypothesis 1 (H1): CDL training will decrease individual computer anxiety level.

2.3 Self-efficacy (SE)

Computer self-efficacy beliefs are expected to have a greater impact on computer anxiety than computer experience and use (Wilfong 2006). Self-efficacy has been defined as “people’s judgments of their capabilities to organize and execute courses of action required to attain designated types of performances” (Bandura 1986, p. 391). This means that self-efficacy is concerned not with the skills one has, but with judgments of what one can do with whatever skills one possesses (Bandura 1986). Thus, self-efficacy is simply the person’s confidence in performing a particular behaviour. In the context of computer use, self-efficacy has been defined as a judgement of one’s ability to use a technology (e.g. computer) to accomplish a particular task (Compeau and Higgins 1995a). A number of researchers agree that computer anxiety and computer self-efficacy are negatively related (Bandura 1977; Henderson, Deane et al. 1995; Igbaria and Ilvari 1995; Compeau and Higgins 1995a; Fagan, Neill et al. 2003; Chang 2005). Hence, we formulate our second hypothesis as follows:

Hypothesis 2 (H2): Reduced computer anxiety will increase computer self-efficacy.

2.4 Perceived ease of use (PEOU) and perceived usefulness (PU)

According to Davis’ technology acceptance model (TAM) (Davis 1989; Davis, Bagozzi et al. 1989) user acceptance of a new technology can be explained by two salient beliefs: perceived ease of use (PEOU) and perceived usefulness (PU). Perceived ease of use has been defined as the degree to which a person believes that using a particular system would be free of effort (Davis 1989). There is a strong negative interdependence between anxiety and PEOU (Venkatesh 2000). However, PEOU’s impact on intention is expected to decrease when experience increases (Davis, Bagozzi et al. 1989).
Computer anxiety has been offered as one of the anchors that determine early perceptions about the ease of use of a new system and it is modeled as an indirect determinant of intention (Venkatesh 2000). In particular, computer anxiety has been found to be an antecedent of perceived ease of use (Venkatesh 2000; Chee Wei, Sutanto et al. 2006). Computer anxiety has been found to have a strong negative effect on perceived usefulness (Igbaria 1993), and also resistance to change is claimed to be a significant determinant of PEOU (Nov and Ye 2008). On the other hand, it has been argued that perceived ease of use and perceived usefulness of microcomputers have a direct inverse effect on computer anxiety (Ferguson 1997).

The nature of the anxiety-PEOU relationship is further complicated by the dispute over whether the relationship is direct or not. System experience has been found to be significantly related to perceived ease of use and the effect of experience fully mediated by computer anxiety (Hackbarth et al. 2003). On the other hand, some studies report that even though anxiety has no mediating role on the impact of computer experience on perceived ease of use, it has some moderating influence on the relationship (Saade and Kira 2007).

While the nature of the anxiety-PEOU relationship remains under dispute, there appears to be a general agreement that anxiety and the salient beliefs of computer use are related. We deem that the weight of the body of literature is in support of anxiety influencing behavioral beliefs. Accordingly, we formulate our third hypothesis as follows:

Hypothesis 3 (H3): Reduced computer anxiety will increase perceived ease of use (PEOU).

2.5 Intention to use (ITU)

Computer anxiety has also been found to have a strong negative effect on behavioral intentions (Igbaria 1993), and only indirect effects on usage, mainly through perceived usefulness (Igbaria and Iivari 1995). In the context of Web-based learning, anxiety has been found to have a significant negative effect on individuals’ continuance intentions of technology usage, (Chiu and Wang 2008).

While some studies have found anxiety’s impact to usage to be insignificant (Compeau and Higgins 1999), many studies have indicated that computer anxiety is indeed an important predictor of the use of technology (Czaja, Charness et al. 2006; Tung and Chang 2007). We thus formulate our fourth hypothesis as follows:

Hypothesis 4 (H4): Reduced computer anxiety will increase intention.

2.6 Proposed model

Based on the above hypotheses, we present our research model in figure 1:

\[ \text{Figure 1. Research model} \]
3 EMPIRICAL STUDY

The empirical data for our study was collected at the Helsinki School of Economics during 2004 – 2007, from altogether 243 students attending eight CDL courses. The duration of each course was seven weeks and the course was voluntary. The data was collected with two sets of questionnaires upon the completion of the first and the last laboratory session. Having a longitudinal set of data made it possible to investigate the possible changes in the level of anxiety and other key constructs.

Our sample consists of fairly experienced users of computer technology. The software applications in question are those included in the widely used Microsoft Office package of software, that is, word processing (MS Word), spreadsheet (MS Excel), graphics (MS PowerPoint), and database applications (MS Access), and the Windows operating system.

All the constructs in our model were validated by previous research (see Appendix 1). The question items were measured using a seven-point Likert scale, with 1 anchored at totally agree and 7 anchored at totally disagree. The detailed measures used can be found in Appendix 1.

4 RESULTS

In this longitudinal study, we explore anxiety’s impact on the perceptions of ease of use and usefulness of computer applications, our particular interest being in how training, in our case a CDL-course, changes these perceptions.

4.1 Descriptive statistics

We analyzed 96 respondents in our first and second data collection. The participants were 50 adult students and 46 university students, of which 25 were males and 71 females. The average age of the respondents was 32 years, and their work experience varied between 0 and 40 years, the average being approximately 9 years. The respondents had on average 10 years of earlier experience with computers (range: 0 - 24 years). The respondents used computers on average for about 3 hours a day. Computer skills prior to the CDL course were acquired mainly (66 %) at work or through self study. Forty percent of the respondents reported having no prior formal computer education (for more details, see Table 1.).

<table>
<thead>
<tr>
<th></th>
<th>Total (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult / University students</td>
<td>50/46</td>
</tr>
<tr>
<td>Men / Women</td>
<td>25 (26 %) / 71 (74 %)</td>
</tr>
<tr>
<td>Age in years (average/ median/ min-max)</td>
<td>32.2 / 29 / 19-60</td>
</tr>
<tr>
<td>Work experience in years (average/median/ min-max)</td>
<td>8.7 / 5 / 0-40</td>
</tr>
<tr>
<td>Computer experience in years (average/median/ min-max)</td>
<td>10 / 10 / 0 - 24</td>
</tr>
<tr>
<td>Computer daily use (average/median/ min-max)</td>
<td>3.3 / 2 / 0-10</td>
</tr>
<tr>
<td>Computer weekly use</td>
<td>7 / 7 / 1-20</td>
</tr>
<tr>
<td>Earlier computer skills (at work or self study/no prior computer education)</td>
<td>66 % / 40 %</td>
</tr>
</tbody>
</table>

Table 1. Descriptive statistics of the respondents.
4.2 Data analysis

We estimated the instrument’s validity in terms of internal consistency, and convergent and discriminant validity (Straub 1989). Internal consistency was tested using Cronbach’s alpha. In both rounds of data collection, all constructs except intention to complete training, displayed an alpha value higher than 0.7 indicating reliability on the common acceptable level (Nunnally 1978) (see Appendix 2).

The instrument’s discriminant and convergent validity was evaluated using a principal component factor analysis of Orthogonal varimax with Kaiser normalization rotation. All measurement items showed high loadings on their respective factors, thus proposing the instrument exhibited satisfactory convergent and discriminant validity (see Appendix 2.)

The summary constructs (mean) of each original construct before and after training are below in Table 2:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Before training</th>
<th>After training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Min.</td>
</tr>
<tr>
<td>ANX</td>
<td>96</td>
<td>1.67</td>
</tr>
<tr>
<td>ITU</td>
<td>96</td>
<td>1.00</td>
</tr>
<tr>
<td>PEOU</td>
<td>96</td>
<td>1.00</td>
</tr>
<tr>
<td>PE</td>
<td>96</td>
<td>1.00</td>
</tr>
<tr>
<td>SE</td>
<td>96</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Valid N (listwise) 96

Table 2. Descriptive statistics before and after training

Before training, computer anxiety (ANX) was not very high (5.6), and after training it was not high at all (6.0). PEOU of basic pc tools changed from being 2.8 before training to being 2.3 after training. The changes of these two construct (ANX and PEOU) were also statistically significant, measured by partial t-test. The other measured constructs did not change significantly. Directions of changes are presented in Table 3 below.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANXc</td>
<td>96</td>
<td>-4.67</td>
<td>1.67</td>
<td>-0.37</td>
<td>1.13</td>
</tr>
<tr>
<td>ITUc</td>
<td>96</td>
<td>-2.50</td>
<td>4.50</td>
<td>0.10</td>
<td>1.12</td>
</tr>
<tr>
<td>PEOUc</td>
<td>96</td>
<td>-2.50</td>
<td>4.50</td>
<td>0.50</td>
<td>1.05</td>
</tr>
<tr>
<td>PUC</td>
<td>96</td>
<td>-4.33</td>
<td>4.33</td>
<td>-0.11</td>
<td>1.16</td>
</tr>
<tr>
<td>SEc</td>
<td>96</td>
<td>-5.92</td>
<td>2.83</td>
<td>-0.04</td>
<td>1.24</td>
</tr>
</tbody>
</table>

Valid N (listwise) 96

Table 3. Direction of change

The results show that computer anxiety decreased with training, as expected. However, a lot of variance can be observed. Perceived ease of use increased the most, while perceived usefulness decreased slightly.
Table 4. Correlations of change constructs

<table>
<thead>
<tr>
<th></th>
<th>ITUc</th>
<th>PEOUc</th>
<th>PUc</th>
<th>SEc</th>
<th>ANXc</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITUc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>1.000</td>
<td>.415**</td>
<td>.596**</td>
<td>.308**</td>
<td>-.295**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.002</td>
<td>.004</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td>PEOUc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.415**</td>
<td>1.000</td>
<td>.215**</td>
<td>.290**</td>
<td>-.174</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.036</td>
<td>.004</td>
<td>.090</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>PUc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.596**</td>
<td>.215**</td>
<td>1.000</td>
<td>.167</td>
<td>-.128</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.036</td>
<td>.0104</td>
<td>.214</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>SEc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.308**</td>
<td>.290**</td>
<td>.167</td>
<td>1.000</td>
<td>-.222**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.002</td>
<td>.004</td>
<td>.104</td>
<td>.030</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>ANXc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>-.295**</td>
<td>-.174</td>
<td>-.128</td>
<td>-.222**</td>
<td>1.000</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.004</td>
<td>.090</td>
<td>.214</td>
<td>.030</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td></td>
</tr>
</tbody>
</table>

Our first hypothesis (H1: CDL training will decrease individual computer anxiety level) was supported. Training significantly influenced anxiety by decreasing it. Anxiety level before training was 5.6 and after training 6.0, and t-test showed the change to be significant. These results confirm the findings of earlier studies (Igbaria 1990; Torkzadeh and Angulo 1992; Popedavis and Vispoel 1993; Torkzadeh and Koufteros 1993; Martocchio 1994; Shelley 1998; Beckers and Schmidt 2001). Also our second hypothesis (H2: Reduced computer anxiety will increase computer self-efficacy.) was supported. The change in the level of self-efficacy before and after training was very low and not statistically significant, but when we look at the correlations of these changes, it can be seen that changes of anxiety and self-efficacy correlate significantly at the 0.05 level (-.222) (see Table 4). The negative correlation between computer anxiety and computer self-efficacy confirm the results of earlier studies (Bandura 1977; Henderson, Deane et al. 1995; Igbaria and Iivari 1995; Compeau and Higgins 1995a; Fagan, Neill et al. 2003; Chang 2005), but our study demonstrates the relationship also in context of longitudinal training.

Our third hypothesis (H3: Reduced computer anxiety will increase perceived ease of use) was, however, not supported. Even though the measurements of anxiety and ease of use showed significant changes after training, the change for ease of use being the strongest, the correlation between the changes of anxiety and ease of use was not significant. A possible explanation to this might be that the users did not experience that the lower anxiety level results in ease of use but instead in use of new features, i.e. the lower the anxiety level, the more prone the users are to learn more advanced functions. This finding is in contrast with earlier studies that have argued a strong relationship between ease of use and anxiety (Venkatesh 2000; Chee Wei, Sutanto et al. 2006; van Raaij and Schepers 2008).

Our fourth hypothesis (H4: Reduced computer anxiety will increase intention) was, again, supported. The negative correlation (-.295) between changes on computer anxiety and intention was significant at the 0.01 level. This means that reduced computer anxiety could increase intention to use, and that anxiety, in general, can have a negative impact on intention to use. There are contradicting findings for this in earlier research (Igbaria 1993; Compeau and Higgins 1999; Kluwin and Noretsky 2005; Chiu and Wang 2008), clearly calling for future research on the issue.
We hence conclude that training decreases computer anxiety and increases perceived ease of use. Training was found to influence user acceptance of information technology, decreasing anxiety and increasing perceived ease of use. There is a negative relationship between computer anxiety and self-efficacy also after training. Contrary to earlier research, there was no relationship between the changes of levels of computer anxiety and ease of use. In addition, anxiety seems to impact intention to use technology.

Our results can be summarized as follows (see Figure 2): Firstly, training, or a CDL course in our study, decreases individual anxiety level (ANX). Secondly, reduced anxiety has a positive impact on self-efficacy (SE). Thirdly, anxiety does not correlate with perceived ease of use (PEOU), which might mean that users do not experience that the lower anxiety level results in ease of use. Fourthly, reduced anxiety increases intention (INT). We have included in our model also the relationships between self-efficacy and intention, self-efficacy and perceived ease of use, perceived ease of use and perceived usefulness (PU), perceived ease of use and intention, and perceived usefulness and intention, but these were not hypothesized, nor tested, in this study, but left for future research.

Our findings suggest that training significantly decreases computer anxiety, and at the same time, increases perceived ease of use among relatively experienced computer users. This means that training has a clear positive influence on user acceptance of these computer applications, Microsoft Office tools in our case. In addition, training seems to strengthen the negative relationship between computer anxiety and self-efficacy. The most surprising outcome was that the relationship between computer anxiety and perceived ease of use was not clear. After training changes in both constructs were statistically significant, perceived ease of use increased and anxiety decreased, but the changes did not correlate. Some scholars argue that anxiety is an anchor of perceived ease of use (Venkatesh 2000; Chee Wei, Sutanto et al. 2006) or vice versa (Ferguson 1997). Furthermore, anxiety has been suggested to be a result of the beliefs an individual has, rather than an antecedent to them (Saade and Kira 2007). While we found that changes at the levels of anxiety and ease of use do not correlate, the
findings of our study indicate that the relationship between these constructs needs deeper investigation. Our results also demonstrated a strong relationship between computer anxiety and intention to use technology.

While our study provided interesting findings about the relationship between anxiety and ease of use the study has a few limitations. Our study examined and reported anxiety’s impact only on perceived ease of use, self-efficacy, and intention. More research is needed to test anxiety’s impact on other user acceptance related beliefs such as perceived usefulness, subjective norm, job-relevance and compatibility.

References


Proceedings ECIS 2009
APPENDIX 1: Measures Used in the Study

**Computer Anxiety (ANX)** (Drawn from the Computer Anxiety Rating Scale (Heinssen et al., 1987):

1. It scares me to think that I could cause the computer to destroy a large amount of information by hitting the wrong key.
2. I feel apprehensive about using computers.
3. I hesitate to use a computer for fear of making mistakes that I cannot correct.

**Computer Self-Efficacy (SE)** (Compeau and Higgins 1995a; Compeau and Higgins 1995b; Compeau and Higgins 1999):

I could complete my job using the technology if….
1. I had seen someone else using it before trying it myself.
2. Someone showed me how to do it first.
3. I could call someone for help if I got stuck.
4. I had used similar packages like this one before to do the job.
5. I had a lot of time to complete the job for which the software was provided.
6. Someone else helped me get started.

**Perceived ease of use (PEOU)** (Davis 1989; Davis, Bagozzi et al. 1989):

1. It would be easy for me to become skillful at using the system.
2. Learning to operate the system is easy for me.
3. I believe that it is easy to get the system to do what I want to do.
4. Overall, I believe that the system is easy to use.

**Perceived usefulness (PU)** (Davis 1989; Davis, Bagozzi et al. 1989):

1. Using basic PC tools makes it easier to do my job.
2. Using basic PC tools increases my productivity.
3. Using basic PC tools enables me to accomplish tasks more quickly.

**Intention to use (ITU)** (Hu and Chau (1999):

1. Whenever possible, I intend to use this basic software in my job/study.
2. I intend to use this software versatility in my job/study.

All items measured using a seven-point Likert scale (1=totally agree, 7=totally disagree).
### APPENDIX 2: SIMPLE STATISTICS

<table>
<thead>
<tr>
<th>Variable</th>
<th>Before training</th>
<th>After training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td><strong>Anxiety (ANX)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANX-1</td>
<td>96</td>
<td>4.85</td>
</tr>
<tr>
<td>ANX-2</td>
<td>96</td>
<td>6.02</td>
</tr>
<tr>
<td>ANX-3</td>
<td>96</td>
<td>6.06</td>
</tr>
<tr>
<td><strong>Intention (INT)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITU-1</td>
<td>96</td>
<td>1.95</td>
</tr>
<tr>
<td>ITU-2</td>
<td>96</td>
<td>2.25</td>
</tr>
<tr>
<td><strong>Perceived ease of use (PEOU)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEOU-1</td>
<td>96</td>
<td>2.57</td>
</tr>
<tr>
<td>PEOU-2</td>
<td>96</td>
<td>2.79</td>
</tr>
<tr>
<td>PEOU-3</td>
<td>96</td>
<td>2.93</td>
</tr>
<tr>
<td>PEOU-4</td>
<td>96</td>
<td>2.83</td>
</tr>
<tr>
<td><strong>Perceived usefulness (PU)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU-1</td>
<td>96</td>
<td>1.57</td>
</tr>
<tr>
<td>PU-2</td>
<td>96</td>
<td>2.16</td>
</tr>
<tr>
<td>PU-3</td>
<td>96</td>
<td>1.82</td>
</tr>
<tr>
<td><strong>Self-efficacy (SE)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE-1</td>
<td>96</td>
<td>3.34</td>
</tr>
<tr>
<td>SE-2</td>
<td>96</td>
<td>2.41</td>
</tr>
<tr>
<td>SE-3</td>
<td>96</td>
<td>2.24</td>
</tr>
<tr>
<td>SE-4</td>
<td>96</td>
<td>2.17</td>
</tr>
<tr>
<td>SE-5</td>
<td>96</td>
<td>2.97</td>
</tr>
<tr>
<td>SE-6</td>
<td>96</td>
<td>2.99</td>
</tr>
</tbody>
</table>
THE DEVELOPMENT AND TEST OF A RELATIONSHIP MODEL ON SYSTEM USE, JOB LEARNING, AND IMPACT

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0555.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Learning, Management control, Customer satisfaction / service, Partial Least Squares</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
THE DEVELOPMENT AND TEST OF A RELATIONSHIP MODEL ON SYSTEM USE, JOB LEARNING, AND IMPACT

Abstract

This exploratory study examined the role of job learning on the relationship between information systems use and impact. Data from 308 end-users were analyzed to evaluate the relationship between systems use, job learning, and technology impact. System use was conceptualized as decision support, work integration, and customer service. Technology impact was conceptualized as effect on management control, task innovation, task productivity, and customer satisfaction. Two sets of hypotheses are presented for these relationships. Results suggest that the pattern of system use significantly and positively influenced job learning. Job learning was found to significantly and positively influence technology impact. We theorize that individuals learned about their job as a result of systems usage. In turn, job learning influenced technology impact. The study findings are discussed.

Keywords: system use, job learning, information technology impact.
THE DEVELOPMENT AND TEST OF A RELATIONSHIP MODEL ON SYSTEM USE, JOB LEARNING, AND IMPACT

1. INTRODUCTION

Information technology continues to play a pivotal role in the structure of work and human productivity. Organisations often attribute their high performance to effective application of information systems. Increasingly, information system executives are required to explain technology expenditures in terms of individual benefits and organisational outcomes. Specifically, organisations are becoming increasingly concerned about what technology use means in the context of the organisation. For example: changes in management control, innovation, productivity, and customer satisfaction. Because of this, the system success paradigm has progressed from an emphasis on ‘suitability for use’ where design features such as content, accuracy, format, and ease of use are considered important, to an emphasis on ‘benefit of use’ where systems’ impact on the individual and the organisation is considered essential (Melone 1990, Torkzadeh and Doll 1999).

Research studies have addressed information technology impact from a variety of perspectives including: new ventures (Fairlie 2006), business performance (Brynjolfsson & Hitt 2000), competitive advantage (Sethi & King 1994), organisational strategy (Mahmood 1991), time management (Sulek & Marucheck 1991), and industry level (Segars & Grover 1994). Increasingly, the relevant question about information technology impact relates to the nature of that impact and how it occurs. Impact through use is one area of investigation suggested (Torkzadeh & Doll 1999). Although it is quite obvious system use would lead to impact, the unanswered question is how. We propose job learning as one possible mechanism in this regard.

As individuals interact with system applications, they learn about their job, and as a result, become more productive. Information technology plays an important role in job learning, and individual productivity. Because of this, information technology plays a critical role in the expansion of knowledge: learning becomes a new form of labor (Zuboff 1988). “Learning is no longer a separate activity that occurs either before one enters the workplace or in remote classroom setting.” “Learning is the heart of productive activity”(Zuboff 1988, p. 395). For the individual, information technology holds promise in terms of job enhancement and the outcome of labor. Information technology impacts how a job is performed and what the outcome might be. Thus, there is a need for better understanding the nature and outcome of the interaction between people and technology in an organisational context.

This exploratory study examines the relationship between system use, job learning, and technology impact. Previously published measures of systems use and technology impact as well as a newly developed measure of job learning are used to collect data from 308 end-users and to examine this relationship. Measures of system use exhibit patterns of use in terms of decision support, work integration, and customer service. Measures of technology impact evaluate perceived outcomes in terms of management control, task innovation, task productivity, and customer satisfaction. Measures of job learning reflect technology influence on the ability to learn and perform job functions as well as to improve the quality of work. In the following section we will review the relevant literature, develop hypotheses, describes the structural model and measures, describe our results, and finally draw conclusions.
2. REVIEW OF LITERATURE AND RESEARCH HYPOTHESES

2.1 System Use

The measurement of information systems success continues to be an important topic for research and practice. At least two perspectives exist in the literature for measuring systems success: the design perspective and the outcome perspective. The design perspective has a strong tradition in the MIS field and involves evaluating systems relative to design specifications or user needs. The outcome perspective calls for performance-related evaluations that focus on outcomes. Measures of user satisfaction (Doll & Torkzadeh 1988) and perceived usefulness (Davis 1989) are widely accepted examples of the design perspective. Measures of technology impact on work (Torkzadeh & Doll 1999) and technology impact on competitive advantage (Sethi & King 1994) are good examples of the outcome perspective.

System use has also been considered as a measure of system success in earlier research studies (Hamilton & Chervany 1981, Ives et al. 1980, Ein-Dor & Segev 1978). It has been viewed as an important construct in conceptualizing information system success (Doll & Torkzadeh 1998, DeLone & McLean 1992). However, other studies argue that the critical success factor in technology investment is not system use in and of itself, but the net benefits to organisations that occurs from that use (Seddon 1997, Szajna 1993). Therefore, while system use is a pivotal link in the ‘system-to-value chain’ from technology adoption to social and economic impact (Doll & Torkzadeh 1991), it is the outcomes of use that reflect system success. In this taxonomy user satisfaction and perceived usefulness are expected to influence system use.

There is a great diversity of system use constructs in information system research (Burton Jones & Straub, 2006). While the emphasis of IS literature on system use is more concerned with the justification for creating and/or utilizing information systems, the social science literature on the nature of work views information technology as being used by individuals in a work context to perform certain organisationally relevant functions (Doll & Torkzadeh, 1998). For example, information technology is used to communicate with subordinates and superiors, to facilitate problem solving, to plan team work, to service customers, and to rationalize decisions, etc.

In order to measure how information technology is actually used by individuals in an organisational context, Doll and Torkzadeh (1998) developed a multidimensional instrument for technology utilisation for the three functions of decision support, work integration, and customer service. The decision support function was defined in terms of ‘problem solving’ (the extent that information technology is used to analyse cause and effect relationships and to make sense out of data) and ‘decision rationalisation’ (the extent that information technology is used to improve the decision making processes or explain/justify the reasons for decisions). Work integration was defined in terms of ‘horizontal integration’ (the extent that information technology is used to coordinate work activities with others in one’s work group) and ‘vertical integration’ (the extent that information technology is used to plan one’s own work, monitor performance, and communicate vertically to coordinate one’s work with superiors and subordinates). Customer service was defined as: the extent that information technology is used to service internal and external customers. These constructs are adopted for the current study.

2.2 Job Learning

Job learning is an important aspect of performance. Although most positions require evidence of capabilities/skills from employees, a large part of learning occurs as the work is being performed. Learning in the workplace has been characterized as the process of seeking technical, referent, and normative information (Morrison 1993). While employees are charged with the responsibility
to exhibit learning behaviour, the organisation needs to provide the opportunities for the employees to learn. Extending work-based learning from the individual, to the group, and to the entire organisation prompted the development for the concept of learning organisation. Implementation of knowledge management systems is one consideration for organisation to provide and/or enhance work-based learning through application of information technology.

Although there is a long history of information systems use at the workplace, research on how this use affects job learning is very limited. As individuals interact with technology to accomplish tasks, they learn more about their job and become more innovative in carrying out responsibilities (Ruiz-Mercader, et al. 2006). Information technology enables employees to deliver more value to the customer (Harvey et al. 1993). The use of information technology is expected to enrich and broaden jobs (Long 1993). Employees use information technology in innovative ways to enhance their customer service. Customer relationship management (CRM) systems are a good example of applications that help employees to develop new and innovative ways of providing customer service. Cross-functional integration and effective data processing provided by CRM applications enable employees to access customer profiles and product information and even predict customer needs (Torkzadeh et al. 2006, Reinartz et al. 2004).

Since the focus of this study is based on system use, we need to examine job learning in the same context. Therefore, we define job learning as a user’s perception of the extent an application enhances learning about the job/task performed. As an employee uses systems for decision support, s/he would likely learn more about the decision variables that need to be included as well as justification for the decision. By using systems to coordinate and communicate with others, the user would see the benefit of the system for learning about the people and work flow related to the task at hand. The enhancement of job learning through system use should eventually produce positive outcomes.

2.3 Impact

Information technology has influenced the nature of work, the process of learning and ways of accomplishing organisationally relevant tasks. Organisations are increasingly interested in the extent and the nature of their IT investment outcomes, and how application development and acceptance benefits their bottom line. While adoption and use of information technology continues to be an important goal of information system executives, there is an increased emphasis on the net benefits that emerge from system use (Seddon 1997). Information system executives are expected to explain the value and contribution of information technology expenditure in terms of increased productivity, quality, and competitiveness (Myers et al. 1997).

Traditional approaches for measuring technology impact emphasize productivity and management control. The extent of information technology use and its influence on productivity and management control has long been the focus of attention (Weick 1990, Zuboff 1988, Braverman 1974). MIS researchers have devoted considerable attention to the impact of information technology on productivity (Sulek & Marucheck 1992, Cooper & Zmud 1990, Kraemer & Danziger 1990, Hirschhorn & Farduhar 1985). More recently that emphasis has included technology impacts on innovation and customer service. In addition to productivity and management control, information technology impacts on innovation and customer satisfaction have also gained increased attention (Filiatrault et al. 1996, Harvey et al. 1993, Davis 1991, Curley & Pyburn 1982).

To help management distinguish between effective and ineffective applications, Torkzadeh and Doll (1999) developed a set of outcome measures in the context of management control, task innovation, task productivity, and customer satisfaction. Management control was defined as: the extent that the application helps to regulate work processes and performance. Task innovation
was defined as: the extent that an application helps users create and try out new ideas in their work. Task productivity is defined as: the extent that an application improves the user’s output per unit of time. Customer satisfaction was defined as: the extent that an application helps the user create value for the firm’s internal or external customers. To the extent that these constructs relate to organisationally relevant outcomes they are appropriate for the context in which this study was carried out. We adopt these concepts of technology impact in this study.

Based on review of the literature presented in this section, two sets of hypotheses are presented that describe the relationship between system use and job learning, and job learning and the impact of technology on work and its components. In this model, job learning is considered the intervening variable between system use and technology impact, as depicted in Figure 1.

H1: Systems use measured in terms of decision support is expected to be positively related to job learning.

H2: Systems use measured in terms of work integration is expected to be positively related to job learning.

H3: Systems use measured in terms of customer service is expected to be positively related to job learning.

H4: Job learning is expected to be positively related to information technology impact in terms of management control.

H5: Job learning is expected to be positively related to information technology impact in terms of task innovation.

H6: Job learning is expected to be positively related to information technology impact in terms of task productivity.

H7: Job learning is expected to be positively related to information technology impact in terms of customer satisfaction.
3. RESEARCH METHODS

3.1 Operationalisation of Constructs

To examine the relationships depicted in Figure 1, a combination of published and newly developed measures were used to collect the data. In this study, the three-factor measurement model developed by Doll and Torkzadeh (1998) was used to operationalise system use. The instrument consists of 13, 11, and 5 items for decision support, work integration, and customer service, respectively. Examples of system use items include: ‘I use this application to control or shape the decision process,’ ‘I use this application to plan my work’, and ‘I use this application to improve the quality of customer service’.

To measure technology impact, the four-factor measurement model of information technology impact developed by Torkzadeh and Doll (1999) was used. Each of the four constructs (task productivity, task innovation, customer satisfaction, and management control) is measured using 3 items. Examples of technology impact items include: ‘This application improves management control’, ‘This application helps me create new ideas’, ‘This application increases my productivity’, and ‘This application improves customer service’. Both system use and technology impact instruments were measured using a 5-point Likert-type scale anchored by (1) ‘Not at all’ and (5) ‘A great deal.’

Job learning was operationalised using eight items that asked respondents how information technology influenced their job learning. Although there is a broad understanding that information technology is a learning tool and that it has the potential to help employees learn more about their jobs and how to better perform, to the best of our knowledge, there are no job learning measures linking technology to job learning. Literature on productivity and organization provide a broad background to the understanding of how technology might influence individuals as they go about learning about their jobs and performing their tasks (Weick 1990, Braverman 1974, Zuboff 1988). That literature suggests that information rich organizations are learning environments with the purpose to be more productive (Zuboff 1988). Job learning in this study...
was conceptualized in terms of how technology assists individuals become more skillful at doing what they are supposed to do as well as help them to better perform their assigned tasks.

Survey items were generated to operationalize technology as it helps the individual “how to do things, rather than what to do or why” (Zuboff 1988, p. 206). This conceptualization also suggests that technology helps the individual to understand the job better as well as to perform the job more effectively. Survey items also intended to assess whether information technology would increase employees’ capabilities to enrich and expand jobs (Lang 1993). Examples of job learning indicators in this study include: ‘This application increases the ability required to do my job’, ‘This application helps me learn how to improve the quality of my work’, ‘This application increases the capabilities required to do my job’, and ‘This application helps me better understand my job’. A Likert-type scale similar to the one used in system use and technology impact measures was used for these questions. The job learning measure was found to be both reliable and valid (reported below).

3.2 Sample

A survey questionnaire comprised of 30 items measuring system use, 8 items measuring job learning, and 12 items measuring technology impact was used to collect data. The survey was also used to collect respondent information, type of application, and the level of use.

The respondents relied heavily on specific applications for completing their job functions. By collecting data from users who relied heavily on the use of a specific application, the researchers were confident that respondents could identify patterns of application use in their organizational context, how the application helped them learn about their job, and how they viewed the impact of job learning on how the technology impacted their work. Demographics revealed a broad industry representation. Respondents worked for government agencies (19.5%), manufacturing (16.2%), health service (14.6%), transportation (12.6%), education (9.3%), finance (8.8%), wholesale and retail (4.9%), and others (14%). Several incomplete responses were discarded and a sample of 308 complete responses to all constructs was used for analysis. Discarded responses were considered too few to suggest a meaningful difference between incomplete and complete responses. Major applications include office automation applications (22.5%), financial applications (20.9%), and accounting applications (13.5%).

4. DATA ANALYSIS AND RESULTS

We used partial least squares (PLS-Graph 3.0) to analyze the proposed relationships. PLS is suitable because the aim of the study is to examine the predictive validity of the ‘system use’ and ‘job learning’ constructs. In addition, PLS relaxes certain distributional assumptions and is considered appropriate for exploratory. Kolmogorov-Smirnov’s test of normality indicates that none of our measurement items are normally distributed. All items are modeled as reflective according to their original design. The measurement and structural model were tested simultaneously. Since PLS does not produce fit statistics, we followed the general criteria of item loadings above .7, path coefficients above .2 (Chin 1998), and t-statistics for item loadings and path coefficients generated from bootstrapping (100 resamples) to evaluate the analysis results.

Because the items for ‘job learning’ were developed for this study, we first ran exploratory factor analysis to examine the factor structure for these items. All eight items loaded on one factor with strong loadings that range from .745 to .854 and explained 66% of the available variance. Cronbach’s alpha for the eight items is .92 and all corrected item-total correlations are above .7, indicating good internal consistency. Thus all eight items for the job learning’ measure were retained.
Results of the PLS measurement model (i.e., item loading, cross-loading, t-statistics, composite reliability, and AVE) are presented in the Appendix. Most item loadings are above .7 and all loadings are significant. Although a few items have marginal loadings, we decided to retain them to be consistent with the original instruments. Although some cross-loadings were observed, all items loaded highest on their respective factors. The composite reliabilities and AVE of all factors are above the accepted .7 and .5 level, respectively. The discriminant validity of the measures is verified by comparing the square root of AVE and cross-construct correlations. As can be seen on Table 1, all correlations are smaller than its respective square-root of AVE. These results provide evidence for convergent and discriminant validity for the measures of ‘job learning’ construct as well as the system use and technology impact constructs.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Make Decisions</th>
<th>Integrate Work</th>
<th>Service</th>
<th>Job Learning</th>
<th>Management Control</th>
<th>Task Innovation</th>
<th>Task Productivity</th>
<th>Customer Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Make Decisions</td>
<td>0.7937</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrate Work</td>
<td>0.7185</td>
<td>0.7183</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service Customer</td>
<td>0.3810</td>
<td>0.4089</td>
<td>0.8408</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job Learning</td>
<td>0.5525</td>
<td>0.5506</td>
<td>0.5133</td>
<td>0.8093</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Management Control</td>
<td>0.5714</td>
<td>0.5785</td>
<td>0.5007</td>
<td>0.6198</td>
<td>0.8809</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task Innovation</td>
<td>0.5488</td>
<td>0.5226</td>
<td>0.3592</td>
<td>0.6758</td>
<td>0.4292</td>
<td>0.9088</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task Productivity</td>
<td>0.4276</td>
<td>0.4035</td>
<td>0.4260</td>
<td>0.6752</td>
<td>0.5006</td>
<td>0.4476</td>
<td>0.8746</td>
<td></td>
</tr>
<tr>
<td>Customer Satisfaction</td>
<td>0.5337</td>
<td>0.3182</td>
<td>0.7769</td>
<td>0.5265</td>
<td>0.5337</td>
<td>0.3321</td>
<td>0.4458</td>
<td>0.9359</td>
</tr>
</tbody>
</table>

Bold = square root of AVE.

The results of the PLS structural model are presented in Figure 2. All system use constructs significantly affected job learning with path coefficients above .2 explaining over 40% ($R^2=.437$) of the variance. Effect sizes were calculated based on the procedure recommended by Chin (1998), and are 0.4, 0.4, and 0.14 for Decision Support, Work Integration, and Customer Service, respectively; indicating small to medium effects. This supports hypotheses 1-3. Job learning also significantly affected each of the ‘impact’ constructs. Path coefficients ranged from .526 to .676, indicating that ‘job learning’ strongly influenced the ‘impact’ constructs. Job learning also explained a significant amount of variance in the impact constructs (.277 to .457). This supports hypotheses 4-7.

![Figure 2: Structural model results](image-url)
5. DISCUSSION AND CONCLUSION

Employees use information technology in an organisational context to accomplish specific tasks and carry out responsibilities. The measurement of information technology use in information systems research has progressed from the traditional focus on the level and frequency of computer use to the conception that incorporates intent and pattern of use. Research interest in this domain has moved from how much technology is used to ways in which technology is used. This latter conceptualization has implications for evaluating technology impact on work; how we evaluate the influence of system use on the nature of work and productivity.

Individuals interact with technology applications to explore ways to improve their job performance, and in that fusion of exploring and doing, they learn and enhance their knowledge about their job. Learning becomes a part of what they need to do in order to do it better. Learning processes occur in the context of work and employees learn as they go about solving problems (Bereiter 2002). In evaluating technology impact, we must go beyond what individuals currently do and examine how prepared they are to do what they need to do next. Formal training is expected to provide the individual with core competency and fundamental knowledge to know how to learn on the job. Employee skills are developed in a learning environment that includes work settings, tools, problems, and co-workers who have common purpose (Lambrecht et al. 2004). Because we are evaluating the impact of technology in organisational contexts we are evaluating the interaction of people and technology in an organisational environment rather than separately evaluating the individual, the technology, or the organisation. This conceptualization is helpful in explaining the perception of a widening gap between the potential of information technology and its actual use, and, represents a major contribution towards work in this important area.

This view of system use and technology impact in an organisational context influenced the researchers during the design and implementation of the current study. Our premise is that when information technology is used by individuals in new ways, that interactive effect has important implications for the nature of work, the need to learn and innovate, and the approach to decision problems. Specifically, our objectives were: (a) to evaluate system use in terms of a ‘function’ that individuals could easily relate to in their work context (e.g., to rationalize decisions, to make sense out of data); (b) to evaluate technology impacts in terms of organisationally relevant outcomes (e.g., improved customer service, improved productivity); (c) to evaluate job learning as a behavior that links system use with the perceived impact of technology; and (d) to extend the conception of technology impact beyond the traditional focus on productivity and management control and to include dimensions of customer satisfaction and task innovation that are relevant to the success and survival of modern organisations.

We encourage confirmatory studies of these findings for specific industries (e.g., service), in specific settings (e.g., in an environment where user participation in system development is strong or where the majority of developmental activities are offshored), and for specific technologies (e.g., customer relationship management). Studies that are more focused on an industry, environment, or technology would demonstrate the potential benefits for research and practice in these specific settings. In these follow up studies, part or all of the ‘system use’ and ‘technology impact’ constructs may be appropriate.
References


## Appendix

**Factor loadings, cross loadings, reliabilities and t-statistics**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Decision Support</th>
<th>Work Integration</th>
<th>Customer Service</th>
<th>Job Learning</th>
<th>Mgmt. Control</th>
<th>Task Innovation</th>
<th>Task Productivity</th>
<th>Customer Satisfaction</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decision Support</strong> (Composite Reliability = 0.957, AVE = 0.630)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS1</td>
<td>0.872</td>
<td>0.639</td>
<td>0.258</td>
<td>0.450</td>
<td>0.465</td>
<td>0.528</td>
<td>0.334</td>
<td>0.251</td>
<td>30.505</td>
</tr>
<tr>
<td>DS2</td>
<td>0.854</td>
<td>0.616</td>
<td>0.347</td>
<td>0.458</td>
<td>0.490</td>
<td>0.478</td>
<td>0.372</td>
<td>0.333</td>
<td>25.261</td>
</tr>
<tr>
<td>DS3</td>
<td>0.822</td>
<td>0.576</td>
<td>0.256</td>
<td>0.429</td>
<td>0.494</td>
<td>0.482</td>
<td>0.377</td>
<td>0.259</td>
<td>21.707</td>
</tr>
<tr>
<td>DS4</td>
<td>0.815</td>
<td>0.575</td>
<td>0.365</td>
<td>0.420</td>
<td>0.509</td>
<td>0.366</td>
<td>0.309</td>
<td>0.360</td>
<td>19.385</td>
</tr>
<tr>
<td>DS5</td>
<td>0.812</td>
<td>0.595</td>
<td>0.369</td>
<td>0.454</td>
<td>0.447</td>
<td>0.475</td>
<td>0.270</td>
<td>0.358</td>
<td>19.405</td>
</tr>
<tr>
<td>DS6</td>
<td>0.804</td>
<td>0.606</td>
<td>0.350</td>
<td>0.465</td>
<td>0.458</td>
<td>0.421</td>
<td>0.407</td>
<td>0.324</td>
<td>19.636</td>
</tr>
<tr>
<td>DS7</td>
<td>0.802</td>
<td>0.570</td>
<td>0.294</td>
<td>0.355</td>
<td>0.421</td>
<td>0.398</td>
<td>0.260</td>
<td>0.254</td>
<td>16.875</td>
</tr>
<tr>
<td><strong>CS1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Task Productivity</strong> (Composite Reliability = 0.907, AVE = 0.765)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TM1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CS2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CS3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Customer Service</strong> (Composite Reliability = 0.923, AVE = 0.707)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CS1</strong></td>
<td>0.322</td>
<td>0.319</td>
<td>0.884</td>
<td>0.439</td>
<td>0.456</td>
<td>0.307</td>
<td>0.359</td>
<td>0.776</td>
<td>26.644</td>
</tr>
<tr>
<td><strong>CS2</strong></td>
<td>0.299</td>
<td>0.294</td>
<td>0.880</td>
<td>0.408</td>
<td>0.388</td>
<td>0.232</td>
<td>0.375</td>
<td>0.680</td>
<td>26.070</td>
</tr>
<tr>
<td><strong>CS3</strong></td>
<td>0.329</td>
<td>0.337</td>
<td>0.871</td>
<td>0.499</td>
<td>0.462</td>
<td>0.373</td>
<td>0.358</td>
<td>0.746</td>
<td>28.731</td>
</tr>
<tr>
<td><strong>CS4</strong></td>
<td>0.377</td>
<td>0.371</td>
<td>0.817</td>
<td>0.419</td>
<td>0.406</td>
<td>0.274</td>
<td>0.343</td>
<td>0.560</td>
<td>15.488</td>
</tr>
<tr>
<td><strong>CS5</strong></td>
<td>0.272</td>
<td>0.409</td>
<td>0.746</td>
<td>0.379</td>
<td>0.383</td>
<td>0.311</td>
<td>0.361</td>
<td>0.471</td>
<td>12.341</td>
</tr>
<tr>
<td><strong>Job Learning</strong> (Composite Reliability = 0.938, AVE = 0.655)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TM1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TM2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TM3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Management Control</strong> (Composite Reliability = 0.912, AVE = 0.776)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MC1</strong></td>
<td>0.550</td>
<td>0.530</td>
<td>0.450</td>
<td>0.593</td>
<td>0.897</td>
<td>0.410</td>
<td>0.481</td>
<td>0.516</td>
<td>39.314</td>
</tr>
<tr>
<td><strong>MC2</strong></td>
<td>0.474</td>
<td>0.530</td>
<td>0.479</td>
<td>0.553</td>
<td>0.885</td>
<td>0.402</td>
<td>0.460</td>
<td>0.458</td>
<td>25.440</td>
</tr>
<tr>
<td><strong>MC3</strong></td>
<td>0.482</td>
<td>0.464</td>
<td>0.389</td>
<td>0.484</td>
<td>0.863</td>
<td>0.314</td>
<td>0.372</td>
<td>0.430</td>
<td>18.189</td>
</tr>
<tr>
<td><strong>Task Innovation</strong> (Composite Reliability = 0.935, AVE = 0.826)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TI1</strong></td>
<td>0.489</td>
<td>0.479</td>
<td>0.304</td>
<td>0.591</td>
<td>0.372</td>
<td>0.926</td>
<td>0.420</td>
<td>0.278</td>
<td>44.889</td>
</tr>
<tr>
<td><strong>TI2</strong></td>
<td>0.512</td>
<td>0.491</td>
<td>0.352</td>
<td>0.601</td>
<td>0.440</td>
<td>0.910</td>
<td>0.385</td>
<td>0.291</td>
<td>38.825</td>
</tr>
<tr>
<td><strong>TI3</strong></td>
<td>0.495</td>
<td>0.456</td>
<td>0.323</td>
<td>0.646</td>
<td>0.360</td>
<td>0.891</td>
<td>0.415</td>
<td>0.333</td>
<td>32.243</td>
</tr>
<tr>
<td><strong>Task Productivity</strong> (Composite Reliability = 0.907, AVE = 0.765)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TP1</strong></td>
<td>0.386</td>
<td>0.376</td>
<td>0.378</td>
<td>0.634</td>
<td>0.472</td>
<td>0.385</td>
<td>0.898</td>
<td>0.388</td>
<td>26.315</td>
</tr>
<tr>
<td><strong>TP2</strong></td>
<td>0.400</td>
<td>0.360</td>
<td>0.373</td>
<td>0.632</td>
<td>0.424</td>
<td>0.439</td>
<td>0.880</td>
<td>0.406</td>
<td>34.000</td>
</tr>
<tr>
<td><strong>TP3</strong></td>
<td>0.329</td>
<td>0.317</td>
<td>0.368</td>
<td>0.484</td>
<td>0.415</td>
<td>0.341</td>
<td>0.846</td>
<td>0.375</td>
<td>15.565</td>
</tr>
<tr>
<td><strong>Customer Satisfaction</strong> (Composite Reliability = 0.954, AVE = 0.876)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CS1</strong></td>
<td>0.337</td>
<td>0.309</td>
<td>0.755</td>
<td>0.460</td>
<td>0.520</td>
<td>0.290</td>
<td>0.405</td>
<td>0.942</td>
<td>48.564</td>
</tr>
<tr>
<td><strong>CS2</strong></td>
<td>0.300</td>
<td>0.288</td>
<td>0.710</td>
<td>0.495</td>
<td>0.476</td>
<td>0.303</td>
<td>0.404</td>
<td>0.933</td>
<td>42.828</td>
</tr>
<tr>
<td><strong>CS3</strong></td>
<td>0.370</td>
<td>0.296</td>
<td>0.715</td>
<td>0.518</td>
<td>0.502</td>
<td>0.336</td>
<td>0.438</td>
<td>0.928</td>
<td>41.017</td>
</tr>
</tbody>
</table>
# THE ENTERPRISES SIMULATION IN SECOND LIFE. THE CASE OF PERTING LTD

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0188.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Virtual world, Simulation, E-learning, Knowledge Management Systems</td>
</tr>
</tbody>
</table>
THE ENTERPRISES SIMULATION IN SECOND LIFE. THE CASE OF PERTING LTD

Tampieri, Laura, Bologna University - Forlì Faculty of Economics, Piazzale della Vittoria 15, 47100 Forlì, Italy, laura.tampieri@unibo.it

Abstract

The diffusion of ICT applications in virtual worlds as Second Life (SL), Forterra, There and Multiverse created many interesting research opportunities, particularly in knowledge management and e-learning. Virtual worlds are strongly immersive and produce entrepreneurial and social initiatives in which personal, professional and business relationships move quickly and could be transferred into the real environment.

Enterprises, Municipalities, Universities and Chambers of Commerce are increasing the use of SL enforcing the methodological relevance of networking, a new approach that bases its impact, for the development of business, on the creation and maintenance of net links.

This paper will consider the case of Perting Ltd, the simulated enterprise created in 2001 by Forlì Faculty of Economics, that established in February 2008 its premises in SL and started managing businesses in different sectors.

The main subject of this analysis is the use of virtual reality in management not only for business but also in research and didactics. In managerial theories the research would be an opportunity for a new perspective to know the symmetries among ICT business platforms and to understand, in a better way, how the learning process could support organizations in their approach to the market of internet relationships.

Keywords: virtual world, simulation, e-learning, knowledge management systems
1 FROM INFORMATION SYSTEMS TO VIRTUAL WORLDS

It is widely accepted that ICT offers increasing opportunities for implementing innovative structures and processes in organizations. ICT is connected not only to the low cost processing of information, but opens new perspectives for networking\(^1\), interacting and communicating across the world overcoming physical distances (Ndou, Troshani 2008).

In recent years the increasing interaction between ICT and organizational systems moved the evolution of information systems towards the search of a major integration and coordination among the different business functions / area / departments and external relationships with customers and suppliers (Di Donato 2003).

The use of virtual reality by enterprises and public administrations showed the relevance of visual and perceptual impact on organizational behaviour and contributed to the implementation of BPR through the introduction of subjective factors and intangible assets as key – factors in the management of e-business (Carignani 2005).

In a few years the targets of Internet usage changed from information distribution to the creation of business web sites including knowledge platform linked to products, services and enterprises, based on interactive processes as forums and blogs.

The same purpose derives from the creation and management of virtual worlds (3D Web) as NORRATH (The first virtual space which introduced a virtual currency convertible in the market), MMORPGs (Massively Multiplayer Online Role-Playing Game), METAVERSE (Fully immersive 3D virtual spaces as SL), MMOLEs (Massively Multiplayer Online Learning Environment), FORTERRA, THERE and MULTIVERSE, enable the business development in virtual environment (Kish 2007).

The 3D creates the illusion of the real business world with Avatars which materialize the image of the decision makers, users and visitors moving in SL. One of the main features of the virtual world is the interactivity among Avatars that interact on the basis of an informatics platform appearing as landscape, buildings, roads, etc.

The organizational studies undertaken on management and technology (De Marco 2000) stressed the relevance of Practice Community (Brown, Duguid 2002) and the impact on individual behaviour and trust (Giddens 1990). Moreover it is emphasized their influence on the process of professional and entrepreneurial culture building, together with their dynamism and differentiation, opening the possibilities of a deeper specialization (Bertolotti et al. 2003).

Also the outsourcing of knowledge process (Virtuani 2005) coming from the network through data mining activities support business flows with a new approach based on links among the virtual world actors that represented a huge innovation comparing to the internal and traditional way of information management (Ciborra, Hanseth 2006).

Many Scholars pointed out the difficulties to define the “Virtual World” and the “Virtual Community” (Schoereth, Schrott 2001) owing to the multidisciplinary attitudes needed to approach the subject (Krcmar, Leimeister 2005) and considering the variety of languages used to this purpose by academic and unacademic Authors (Preece 2000).

Choi (2007) analyses the threshold to cross from the informatics and technology to business and network dynamics, defining the Virtual World as an environment in which the participants,\(^1\) Although it is connected to the concept of Networking approach used by Authors to define a way of creating new organizations by stabilizing and implementing links with other entities, networking is used in the meaning of activity oriented to establish links among different individuals on a common knowledge platform, Real, Simulated and Virtual (RSV).
represented by Avatars, coordinate their business actions with an unknown speediness and overcoming the normal space and time limits (Table 1).

<table>
<thead>
<tr>
<th>Armstrong and Hagel (1997)</th>
<th>Virtual communities are computer-mediated spaces where there is a potential for an integration of content and communication with an emphasis on member-generated content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee, Vogel, Limayen (2002)</td>
<td>A virtual community is a technology-supported cyberspace, centred upon communication and interaction of participants, resulting in a relationship being built up</td>
</tr>
<tr>
<td>Porter (2004)</td>
<td>A virtual community is defined as an aggregation of individuals or business partners who interact around a shared interest, where the interaction is at least partially supported and/or mediated by technology and guided by some protocols or norms.</td>
</tr>
<tr>
<td>Krčmar, Leimeister (2004)</td>
<td>A virtual community consists of people who interact together socially on a technical platform; The community is built on a common interest, a common problem or a common task of its members that is pursued on the basis of implicit and explicit codes of behaviour</td>
</tr>
<tr>
<td>MacInnes and Hu (2005)</td>
<td>A virtual world is a persistent synthetic environment where people communicate with each other using a virtual person, Avatar</td>
</tr>
<tr>
<td>Jackson (2007)</td>
<td>A virtual world is a technical platforms; computer-generated worlds where people participate using Avatars; an outgrowth of online gaming; evolving into rich ecosystems of online communities</td>
</tr>
<tr>
<td>Thomas and Brown (2007)</td>
<td>Virtual worlds are persistent, Avatar-based social spaces that provide players or participants with the ability to engage in long-term, joint coordinated action</td>
</tr>
</tbody>
</table>

Table 1. Some relevant definition of “Virtual World” and “Virtual Community”

In SL, the most diffused virtual world (Table 2), the language is very specialized (Thomas, Peters 2007). By the way Land is defined as a place in which it is built the store/shop; Group is a community of residents sharing common interests and needs and People is the residents’ nickname.

<table>
<thead>
<tr>
<th>Year</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total hours used by all Residents</td>
<td>1,968,905</td>
<td>7,337,424</td>
<td>25,646,287</td>
<td>31,990,070</td>
</tr>
<tr>
<td>Size of Landscape (Km²)</td>
<td>293,65</td>
<td>963,66</td>
<td>1,429,65</td>
<td></td>
</tr>
<tr>
<td>Balance sheet of Linden Residents</td>
<td>520,703,616</td>
<td>1,418,354,523</td>
<td>4,183,573,385</td>
<td>5,173,515,776</td>
</tr>
<tr>
<td>L$ exchanged</td>
<td>2,031,208,854</td>
<td>2,507,966,464</td>
<td>2,507,966,464</td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>99,623</td>
<td>2,267,092</td>
<td>11,704,934</td>
<td>13,830,008</td>
</tr>
<tr>
<td>Residents Premium</td>
<td>12,433</td>
<td>49,776</td>
<td>93,219</td>
<td>88,585</td>
</tr>
</tbody>
</table>

---

3 Second Life is a 3D virtual world launched in 2003 and developed by Linden Lab. The users, named “residents”, interact each other through “Avatars” providing an advanced level of a social network service. Residents can participate to individual and group activities, create and trade items (virtual property) and services with one another. Second Life uses an internal currency called the “Linden Dollar” that are usually obtained by changing real money. http://www.secondlife.com
4 Linden Labs defines residents as the accounts created in Second Life. www.secondlife.com
5 Population is defined as the total of enrolled users
6 It is defined as the complex of users that have subscribed an account Premium and have bought Linden with real currency. www.secondlife.com
<table>
<thead>
<tr>
<th>Score of Italy in the classification of Countries for the number of active users</th>
<th>7^</th>
<th>Active users in the range of age 25-34</th>
<th>35%</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of use (in hours) by male</td>
<td>55.14</td>
<td>58.58</td>
<td>59.32</td>
</tr>
<tr>
<td>% of use (in hours) by female</td>
<td>44.86</td>
<td>41.42</td>
<td>40.68</td>
</tr>
</tbody>
</table>

Table 2. The general profile of activities in SL^7

The use of Internet stressed by SL moves from the management input-output processes to the coordination of transactions focused on actors, relationships, operative activities, integration mechanisms and technologies (Virili 2008).

To this purpose the diffusion of different ICT tri-dimensional premises^8 is performing a strong impact on the structures and processes of trade, mainly in term of knowledge management ^9, assuming an increasing function of support and assistance in managing changes (Barabasi 2004) from the direct product-driven to the indirect one (Buttle 2004). This translation of perspectives is based on the personalization and on the maintenance of business links through an adequate knowledge management completely based on specialized software typed on CRM.

Together with the 3D relational marketing we have to consider, in addition, the opportunities of networking among operators that from the traditional concept of “fair” and “business meetings” or “business tours”, have been developed in events, forum and blog building^10 structured as instant media.

2 SECOND LIFE AND THE ENTERPRISE SIMULATION

Many public and private organizations set a place inside SL, exploring the way in which virtual worlds may change their business management. In recent times SL is considered as an effective entrepreneurial environment. For many individuals the first business experience into SL is represented by the creation of a store for marketing and branding targets but other key business areas have been explored to this purpose (Table 3).

The usage of SL in the enterprise simulation has the main aim of allowing the residents to interact directly with the teachers and students during lectures and seminars enforcing an interactive approach that can be considered the added value achieved by this methodology rather than the traditional e-learning platforms by which students can only download the didactical materials.

The attitude of SL in enterprise simulation is particularly useful for the development of entrepreneurship in transition countries, characterized by scarce entrepreneurial initiatives and an environment not favourable to entrepreneurship, in which the virtual reality could be considered as a

^7 For 2005, 2006 and 2007 the data refer to 31th December, for 2008 at 31th May.
^8 In each activity of support and transformation of supply chain we can consider different operative applications of Internet: for the infrastructure management, human resources, technologies development and purchase, the ICT tools are intranet, knowledge management and community; the e-commerce refers to logistics in and out; the conversion process uses CAD (Computer-Aided Design) and CAM (Computer-Aided Manufacturing); while marketing and selling uses the e-franchising, email and web site. The service post sale utilizes the mailing list, forum and extranet (Benassi 2005).
^9 In the era of e-commerce the problem of channel management is linked to the conflicts management of different and new channels of communication (Castaldo 2001).
^10 In the fashion sector, www.fashionblog.it; www.fashionteen.it/blog/.

Proceedings ECIS 2009
way for the first expression of entrepreneurial behaviour and to attract talents for the development of local systems that could be imaged and designed in a very innovative look.

<table>
<thead>
<tr>
<th>Area</th>
<th>Mission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education and Training</td>
<td>Realization of lectures and videoconferences based on interactive approach.</td>
</tr>
<tr>
<td>Communication and Collaboration</td>
<td>Organization of meetings and conferences</td>
</tr>
<tr>
<td>Sales</td>
<td>Branding, Marketing, Customer Relationship Management, Improving Customer Loyalty</td>
</tr>
<tr>
<td>Innovation</td>
<td>Testing and developing products and services</td>
</tr>
</tbody>
</table>

Table 3. Emerging key business areas in SL

Another area of interest refers to the operative marketing with the target of increasing the revenue by products selling according to CRM strategies implementation to know and fidelize the customers. At the same time SL could work as test laboratory to develop new products and services, that can be realized in the real world.

In the education field the framework of enterprise simulation proposed opportunities in the: 1) Definition of the educational process: each participant rotates among the different departments, carrying out tasks connected to the job and to the organizational goals achievement; 2) Development of behavioral capacities, both internal to the enterprise simulation, and external with other foreign companies; 3) Providing adequate and consistent responses to the needs of real companies; 4) Learning to manage a role with the process of know-how transfer; 5) Increasing the basic motivations through greater participation in the processes of learning and cooperative environment; 6) Problem solving activities with customers and suppliers; 7) How to learn and work in team.

The applied methodology is the learning by doing based on the synergy between the knowledge acquired during the studies and the practical skills achieved through the simulated activities.

Enterprise simulation program faces the ICT challenges in using Internet as the main tool to manage relationships with customers and suppliers, together with the communications within internal departments. Last implementation in SL increased the enterprise simulation innovation for didactic and research purposes.

Particularly the ICT sector on the use of virtual reality is resumed by Carpenter, Harrison and List (2005) through researches on boundaries and networking establishing new benchmarking for private and public organizations. On this purpose more than 300 Universities and educational organizations have set up platforms in virtual worlds and are exploring ways and concepts for distance learning purposes.

Virtual Worlds enable new ways of communication and collaboration over the Internet by applying 3D environments and VoIP technologies. Through the combined use of innovative Internet technologies and due to their immersive features, virtual worlds offer new possibilities for the computer-mediated communication and co-operation. Hence, virtual worlds provide innovative learning arrangements and are particularly suitable to transfer experimental knowledge. Students can take part in a distance learning environment to develop a real sense of community in which they cooperate with each other and there is a regular sense of classroom interaction.

Furthermore, Virtual Worlds offer new opportunities for enterprises to interact and understand their existing and potential customers’ needs and desires (Lang, Fetscherin, Lattemann, 2008).
3 THE CASE OF PERTING LTD

Most of the key business areas in SL, above mentioned, are implemented in Forlì Laboratory for enterprise simulation. Forlì Laboratory reproduces for didactical and research purposes the structures and the activities of an enterprise simulating the business in connection to an international network of other simulated enterprises named Europen.

This experimental activity is managed by students under the mentoring and supervising of the teacher and tutors. The sector chosen in Forlì is the business services and IT products trade, under the brand of Perting ltd, a simulated enterprise with its general meeting and board of directors. In the network this enterprise makes business with other firms of the Europen Network operating as suppliers and customers (Table 4).

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>N. of Teachers</th>
<th>N. of Tutor</th>
<th>N. of Students</th>
<th>Hours of Activity</th>
<th>N. of Customers</th>
<th>N. of Suppliers</th>
<th>N. of Visitors in SL</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001-2002</td>
<td>1</td>
<td>1</td>
<td>51</td>
<td>50</td>
<td>14</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>2002-2003</td>
<td>1</td>
<td>2</td>
<td>52</td>
<td>50</td>
<td>11</td>
<td>9</td>
<td>-</td>
</tr>
<tr>
<td>2003-2004</td>
<td>1</td>
<td>2</td>
<td>48</td>
<td>50</td>
<td>48</td>
<td>13</td>
<td>-</td>
</tr>
<tr>
<td>2004-2005</td>
<td>1</td>
<td>2</td>
<td>57</td>
<td>50</td>
<td>13</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>2005-2006</td>
<td>1</td>
<td>2</td>
<td>63</td>
<td>50</td>
<td>121</td>
<td>51</td>
<td>-</td>
</tr>
<tr>
<td>2006-2007</td>
<td>1</td>
<td>2</td>
<td>54</td>
<td>50</td>
<td>92</td>
<td>28</td>
<td>-</td>
</tr>
<tr>
<td>2007-2008</td>
<td>1</td>
<td>2</td>
<td>55</td>
<td>50</td>
<td>113</td>
<td>45</td>
<td>-</td>
</tr>
<tr>
<td>2008-2009</td>
<td>2</td>
<td>7</td>
<td>76</td>
<td>50</td>
<td>15</td>
<td>7</td>
<td>172</td>
</tr>
<tr>
<td>2009-2010</td>
<td>3</td>
<td>3</td>
<td>44</td>
<td>50</td>
<td>55</td>
<td>21</td>
<td>1613</td>
</tr>
</tbody>
</table>

Table 4. The profile of activities carried out by Perting

The corporate purpose of Perting Ltd is the consulting in management, accounting and internet business. Moreover Perting is committed in the trade of informatics equipments (PC, printers, photo cameras, modems, software and so on).

From the academic year 2008-2009 the Laboratory used SL as innovative tool to realize lectures, video conferences and interactive e-learning platforms improving the didactic and research activities. By using SL the number of contacts activated by Perting highly increased from 172 to 1613 enforcing the networking approach creating a network of individuals and organizations.

---

11 The Laboratory is managed by University of Bologna – Forlì Faculty of Economics.
12 The data linked to the number of teachers, tutors, students and hours of activity refer to the first semester. The data related to the number of customers, supplier and visitors in SL refer to 31th October 2008.
13 The data linked to the number of teachers, tutors, students and hours of activity refer to the second semester. The data related to the number of customers, supplier and visitors in SL refer to 31th March 2009.
Summarizing the main relevant steps of Perting Ltd:

- In October 2001 the creation of Perting Ltd was the first milestone of the enterprise simulation Laboratory\textsuperscript{14} connected with the course of Business Management\textsuperscript{15} and supported by Fondazione Cassa dei Risparmi of Forlì. Perting Ltd operated in organizational consulting, business networking and merchandising of ICT products as the first unit of Enterprise simulation established by an Italian University\textsuperscript{16}.

- In 2004 Perting participated to its first international project for the “Formation and technical assistance for SMEs development in Durazzo harbor district”, financed by the Ministry of Foreign Affairs. To this purpose the first Simulated Unit supported by Perting was created in Shkoder (Albania): the KK Personal Robe, operating in textile sector.

- During A.Y. 2005-2006 Enterprise simulation Laboratory of Forlì contributed actively to the start up of the Unit “NoRisk” at Economics’ Faculty of Parma University, a simulated company offering insurance services and trade of products for the safety at the workplace.

- In 27th February 2008 Perting opened its headquarters in SL, becoming the first enterprise simulation created by a University and seat in this virtual world (Figure 1).

![Figure 1. The headquarters of Perting Ltd in SL\textsuperscript{17}](http://slurl.com/secondlife/Kouhun/246/248/54)

The simulation is very realistic because all the managerial processes, from the business plan to the product/service conception, to the delivering of orders and related invoices are carried out as in the reality. At the end of each year on the basis of cost and revenues a balance sheet is prepared and submitted to the Europen Center that certifies the result.

This activity allowed the students to realize a real internship in a concrete job while applying of knowledge received in theoretical courses. This learning is connected to the building of the sense in entrepreneurship and self achievement. For this reason the Laboratory was involved in many international projects for the development of small and medium business in transition countries. As it could reproduce entrepreneurial and marketing practices Perting ltd was used in researches on enterprise start up and on the management of relationships with suppliers and customers.

\textsuperscript{14} Managed by Daniele Gualdi, Professor in charge of « Simulimpresa » course.
\textsuperscript{15} Massimo Bianchi, Full Professor of Business Management in Forlì Faculty of Economics.
\textsuperscript{16} Certified by Network Europen
\textsuperscript{17} www.secondlife.com; http://slurl.com/secondlife/Kouhun/246/248/54
In the premises of Perting in SL the reception desk, the departments (Figure 2) and the conference room are located. In the reception the residents can ask information and join the Perting group. There are the offices in which, with the mentoring of the tutors, the students can receive visitors carrying out the operative activities and the conference room in which the participants organize meetings. The Forum Area is addressed to perform lectures and video conferences and the international projects exhibition that shows the initiatives in which the simulated enterprise participated (Figure 3).

Figure 2. The departments  
Figure 3. The forum area

The new challenges of Perting are connected to the start up of a new business represented by New Fashion Perspectives (NFP), a micro enterprise in fashion sector, with a Laboratory in which the operative tools as clippers and stapler can be used by Avatars to simulate the real activities reproducing the true premises of a micro business undertaken by a mentor supporting didactics from the real market (Figure 4).

Figure 4. The laboratory of NFP in SL

The reason of this support, although the impressive harvest of contributions published on line (Barabasi 2004), is the main source of practical knowledge on the rapidly evolving world of micro enterprises, particularly in fashion sector, could be individuated only in the real market.

18 The mentor or « impresa madrina » is a real enterprise owned by Etienne Meite, supporting the simulation with its practical and continuous up to date know how. The relevance of this support is stated by the basic simulation methodology (Gualdi 2001)

19 www.secondlife.com
4 EARLY EMPIRICAL STATEMENTS

At the moment we would focus on some hypothesis about Perting case and the experimental materials gathered from the impact of virtual reality on previous seven years of experiences in enterprise simulation:

1. The relations among real, simulated and virtual different levels of representation, operate within a multi dimensional symmetry.

2. The real, simulated and virtual business platforms, although working on the same entrepreneurial subject, could apply different strategies and operative methods.

3. The Virtual world is mainly based on exploring strategies while the exploiting refers to the simulated environment.

4. The level of strength and weakness in relationships among organizational entities changes significantly moving from Real to Virtual world.

The research focused on the relationships expressed by the number of contacts with suppliers and customers in real, simulated and virtual world. In particular the research sample composed by New Fashion Perspective (Virtual) Perting (Simulated) and the Mentor Micro Enterprise, evidenced the increasing level of networking when previous stages of simulation started to move in SL.

One of the first problems was to understand the role of the three levels of real, simulated and virtual in making business.

A basic topic is concerning the symmetrical relationships among three steps in business development as a) Business in Real World b) Business Simulation and its transfer to the Virtual platform c) Business purely virtual among actors and organizations living in 3D.

The Figure 5 summarizes these Symmetries on a two axes basis as a framework to manage the business simulation in the virtual world underlining the role of interpersonal links overcoming difficulties in perceiving the indirect real, simulated and virtual connections.

![Figure 5](image)

Figure 5. The multiple symmetry among real, simulated and virtual world

These links enhance the central position of trust processes deepened in recent years on the mechanisms requested for individual and business relationships (Sitkin, Weingart 1995) and on the indirect approach among actors or decision makers. The combination among real, simulated and virtual worlds underlines the range of business opportunities together with the need of selecting strong and personal relationships not only on the basis of trust but on the selection of most reliable links too so as to limit the enormous and uncontrollable mass of contacts and connected information. In these terms the approach to face the start up of new businesses and particularly of new enterprises operating...
in real, simulated and virtual worlds can be linked to the balance between the trust towards new relationships according with an exploring strategy and, on the other side, the trust on existing links as the result of an exploiting approach. The first one is addressed to a higher renewal of system elements and so to weak links among the organizations that belong to the environment. The exploiting strategy is mainly addressed to a low renewal and so is adequate to strong organizations that have overcome the start up phase (Dittrich, Duysters 2007).

On this purpose weak and strong links coming from respectively by the exploring and exploiting strategies produce a low and high level of commitment, measured by the scarce and high frequency of collaborations in the partnership.

In particular the exploiting strategy aims to widen the existing technological capacities, to enhance strong links due to the high trust in the commitment and so to establish equal agreements.

The exploring is characterized by the search for new technological capacities producing, in the start up period, weak links with the adoption of opportunistic behaviour and of low commitment, that allow the organizations to establish not equal agreements.

In the relationships between real, simulated and virtual worlds is possible to survey an operative symmetry, as agreed by many Researchers on the setting up (Mei Alves de Oliveira 1994) but also in the analysis of organizational borders (Bianchi 2003).

The Enterprise Simulation is based on the reproduction of interactive internal and external processes in order to improve the performance of participants. On this purpose the falsification of symmetry in the reproduction of organizations can be recognized as a criterion for performance evaluation (Bianchi, Tampieri 2008). Moreover the increasing interaction among real, simulated and virtual environments can be expressed by the e-commerce and the virtual platforms built by real enterprises for branding and marketing targets. Concentrating the attention on R-V comparison, the analysis of organizational relations (Table 5) could be defined as a set of contacts in which the enterprises are inserted.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Buyer (Real / Virtual) (R-V)</th>
<th>Seller (Real / Virtual) (R-V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Store building in SL</td>
<td>R</td>
<td>V</td>
</tr>
<tr>
<td>e-commerce</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>e-commerce</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>Customer loyalty process</td>
<td>V</td>
<td>R</td>
</tr>
</tbody>
</table>

*Table 5. The network structure*

Considering the processes of structure building, the Network Dynamics approach faces the competitive and environmental complexity analyzing the inter organizational network (Reagans, McEvily 2003) as the passage from atomistic structures and processes with separated entities to the network and the metaorganization (Antonelli 2004), in which, in a restricted perspective, the hyperlink (Bianchi 2001) that involves structure and process elements, allows to define the phenomena in a synthetic way (Bruni, Perrotta 2007).

The different structures of network (Vicari 2001) aim to develop a variety of coordination forms able to create an adequate dialogue among organizational systems. To this purpose the fundamental elements of network organization (Grandori 1998, 1999) in virtual environment are: the entities, the relational objective of network, the links form and the institutional mechanisms of coordination as social and contractual ones (Lomi 1997; Mercurio, Testa 2000). In particular the links network can assume different forms according to the modalities through which the entities interact each others. In recent years these new innovative business models belonged to b-web (business-web) (Tapscott et al
In this perspective the issues related to real, simulated and virtual world will be the target of further studies as one of the main mission for didactics and business management.

References


The b-web is defined as a system of suppliers, merchandisers, commercial societies, customers that use Internet as the main channel of communications and transactions.
WHAT HABBO GOERS DO IN PRACTICE? DECOMPOSING ATTITUINAL BELIEFS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0465.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Behavioural theories, Social Networking (e.g. Facebook, second life), Structural Equation Modeling, Virtual world</td>
</tr>
</tbody>
</table>
WHAT HABBO GOERS DO IN PRACTICE? DECOMPOSING ATTITUDINAL BELIEFS

Jani Merikivi, Turku Centre for Computer Science, 20520 Turku, Finland, jani.merikivi@tse.fi

The popularity of social virtual worlds (SVWs) stems from the proficiency of designing appealing activities. In a volitional use context, hedonic outcomes, such as pleasure and enjoyment, along with social interactivity are fundamental attitudinal beliefs fostering the success of SVW. As such, the attitudinal beliefs affecting attitude toward using SVWs with multiple functions is worth studying. The practitioners should however focus on the actual behavioural success factors beyond using SVWs. Using Habbo as an example, this study paper develops a research framework and examines how attitude toward using SVWs mediates Habbo goers’ attitudinal beliefs on the actual behavioural incentives. Based on a review of prior literature a decomposed theory of planned behaviour suggested by Taylor and Todd (1995) is employed. The research model is tested with data collected from 1225 active Habbo goers. The main findings of the study suggest that while the Habbo goers desire for social interaction within Habbo the construct of attitude toward using the service fails to reflect it. This indicates that following the omission of discovering the proper attitudinal beliefs behind the actual behavioural factors investments may well be lost.

Keywords: Social Virtual world, Decomposed Theory of Planned Behaviour, Continuous Use, Attitude
1 INTRODUCTION AND BACKGROUND

Social virtual worlds, such as Habbo, Club Penguin, Teen Second Life, Precious Girls Club, and Poptropica, are analytically separable from game worlds which offer the users an opportunity to contend with each other adherence to a set of rules (cf. Bartle, 2004). While possibly including games for pastime SVWs may be described as simulating the physical space of everyday existence, not entirely conjured up by one’s imagination. SVWs are practically persistent environments with no specific narrative goal structures congruent with games. These social environments are attracting an enormous group of people sharing somewhat common interest, that is, a need for social interaction (see Fetscherin & Lattemann, 2008). In comparison to virtual communities which are largely designed for communication, SVWs, for example, serve as a technological tool to make videos or design clothes, and contain graphic elements such as avatars in either two or three dimensional virtual environment in order to create an immersive experience.

From users’ perspective, successful SVWs aimed particularly at the young perform multiple functions. What these SVWs provide is, however, as individual as the uniqueness of the young themselves. But for nearly every young person they generally are social interaction and gaming. Habbo, for example, is a popular and influential commercial social virtual world among those targeted for preteens and teens aged 10 to 18. It provides a free access to over 30 local portals with several public facilities such as virtual parks, and cafés, and millions of user-generated private virtual rooms. In Habbo, the young Habbo goers are able to communicate with one another and play various non-violent online games. To express themselves the Habbo goers may customise their walking, talking, shouting, and dancing avatars which are representations of themselves or their alter egos, and buy Habbo credits to create and furnish their very own personal virtual rooms. The positive attitude toward participating in free events, running distinct groups, or activities, and creating the content, for example, to honour celebrities has extended beyond the boundaries of the service, motivating them to establish an infinite number of fan sites on the internet (Global Habbo Youth Survey, 2006). Especially the live visits of famous real-life athletes and artists from the music and film industry have endeared Habbo to the young (Global Habbo Youth Survey, 2008). Not building on access fees but commercials, and voluntary premium services Habbo has so far succeeded in holding the critical mass of the young and translated their loyalty directly into monetary value.

The technology acceptance and post-adoption behaviour have attracted prominent and extensive research coverage within IS discipline (Agarwal & Prasad, 1997; Bhattacherjee, 2001; Hsu & Lin, 2008; Karahanna, Straub & Chervany, 1999; Venkatesh & Brown, 2001). Furthermore, prior studies have focused on virtual communities and their unique characteristics through various methods (Bagozzi & Dholakia, 2002; Wellman & Gulia, 1999). Little information is, however, available on expanding the understanding of continued use intention of using SVW with multiple purposes. Consequently, SVWs offer opportunities for testing theories and models in a new context. Hence, for thoroughly grasping the phenomenon regarding the behavioural incentives behind using SVWs calls research in several disciplines such as sociology, economics, and psychology, to mention only a few.

To better complement these existing approaches and to focus more explicitly on examining the conceptualised role of attitudes and the accessible behavioural beliefs behind them, this study exemplifies the theory of planned behaviour (TBP) (Ajzen, 2005). Moreover, agreeing with Shimp and Kavas (1984), who suggest that cognitive constructs of belief cannot be combined to single conceptual determinant, the study expands upon decomposed TPB (Taylor & Todd, 1995; Hsieh, Rai & Keil, 2008) as it provides a comprehensive framework for analysing the information on the antecedents of attitude toward using SVWs. It should, however, be worth noting that behavioural models, such as the theory of reasoned action (TRA) (Fishbein & Ajzen, 1975), the technology acceptance model (TAM) (Davis, Bagozzi & Warshaw, 1989; Davis, 1989) and TPB (Ajzen, 2005)
are constrained to analysing only one behaviour of interest. Applying these theories is rather challenging when examining the use of any technological innovation serving multiple purposes. Consequently, this study is confined here to dealing only with factors regarding attitude toward using SVWs, for which the decomposed TPB is assumed to provide a valuable framework.

In this study, the focus is on Habbo, which is favoured by many active young people because they use it for playing games, trading in virtual furniture, familiarising themselves and communicating with online and offline friends (Global Habbo Youth Survey, 2006; 2008). In other words, meeting with new and old friends and creating, sharing, and consuming the interesting content influenced by norms, values, and offline life is perceived important within such pleasuring and enjoyable virtual environment (Bagozzi & Dholakia, 2002; Hagel, 1999). The SVWs have clearly moved beyond mere entertainment into social realm. Therefore, attention is particularly paid on hedonic and social outcomes. Developing the construct of attitude toward use is of major importance, as it may increase the risk of losing information instead of identifying the actual evaluative variables to distinguish in details why exactly SVWs are being used and how the accessible beliefs are linked to them.

As a result, the purpose of this study paper is to develop a framework and examine how the attitude toward using SVWs mediates the impact of Habbo goers’ attitudinal beliefs on the underlying actual behavioural interests. This study paper has three primary goals. First, by identifying the actual factors behind using SVWs it should enhance theoretical knowledge of developing behavioural frameworks utilised in other online services as well, and, second, provide better managerially relevant and readily discernible in-depth understanding of true behavioural incentives that are beneficial in designing features and services for SVWs. Thirdly, by examining particularly the existing Habbo goers it accumulates the post-adoption user behaviour research (Limayem & Cheung, 2008; Premkumar & Bhattacherjee, 2008).

2 RESEARCH MODEL

Individuals use information technology from both extrinsic and intrinsic motives (Davis, Bagozzi & Warshaw, 1992; Vallerand, 1997), also articulated in consumer behaviour (Venkatraman & MacInnis, 1985) and IS literature on technology acceptance (Hsieh et al. 2008; van der Heijden, 2004; Venkatesh & Brown, 2001). Extrinsic incentive propels individuals for achieving a specific goal, whereas intrinsic, in other words, hedonic is the degree to which users experience pleasure, joy, or satisfaction derived from a specific behaviour (Deci & Ryan, 1980). Teo, Lim, and Lai (1999) have empirically shown that intrinsic incentive has a significant effect on using computer and the internet. In a post-adoption context, users may obtain, at least in part, pleasure through appealing environmental characteristics but more salient is the need for maintaining the concentration and curiosity, as well as satisfying the entertainment purposes (cf. e.g. Ahn, Ryu & Han, 2007; Igarria, Schiffrman & Wieckowski, 1994). For example, the Habbo goers look for a playful job or task inside Habbo in a form of a role play and work as fashion models walking on virtual catwalks. As a result, perceived enjoyment (ENJ) is an emotionally fulfilling activity, and playfulness (PLA), which represents the user’s short-term cognitive involvement, are believed to influence attitude toward using SVWs (Woszczynski, Roth & Segars, 2002). It is, thus, postulated that:

\[ H1: \text{Enjoyment (ENJ) positively affects the attitude toward using Habbo.} \]

\[ H2: \text{Playfulness (PLA) positively affects the attitude toward using Habbo.} \]

It is assumed that leisure-oriented post-adoption user behaviour is not driven by strong productivity but rather instrumental value as to effectively communicate and develop social identity (Bagozzi & Dholakia, 2002). In SVWs, the interaction occurs with one another, not between users and the system as mentioned by Heijden (van der Heijden, 2004). The behaviour may still be extrinsic when the purpose of exchanging information or resources through computers is to accomplish tasks, for instance, to express uniqueness. Unfortunately, it may, however, be extremely difficult to differentiate extrinsic from intrinsic behaviour since interaction may as well carry positive or negative emotions not
necessary in achieving external objectives (Yuan & Gay, 2006). Therefore, within this study paper, interaction with one another implies to social outcomes such as social connectedness (CON), self-exploration (UNI), and status (STA), indicating that these three concepts should be distinguished from intrinsic and extrinsic, and discussed separately.

The Habbo goers have an access to a virtual network that facilitates the exchange of social closeness and status. The need for belongingness reflects subjective awareness of interpersonal closeness (Lee & Robbins, 1995; Lee & Robbins, 1998). Therefore, social connectedness particularly relates to self-fulfilling emotional outcomes such as affiliation and support in bringing individuals together both in offline and online (Baumeister & Leary, 1995; Chiu, Hsu & Wang, 2006; Rheingold, 2000; Wellman & Gulia, 1999). As a result, it is hypothesised, that:

**H3:** Connectedness (CON) positively affects the attitude toward using Habbo.

Social status pertains to the probable instrumental value and is, therefore, conferred on individuals as a result of adopting a technological innovation (Moore & Benbasat, 1991; Venkatesh & Brown, 2001; Venkatesh & Davis, 2000). In SVWs, through participation individuals are volitionally pursuing to gain the esteem, in which others hold them. The relative position of individuals in SVWs is closely linked to the need for uniqueness in differentiating themselves from others through consumption of virtual items (cf. e.g. Ruivo, 2008; Snyder & Fromkin, 1977; Tian, 2001). In Habbo, to appeal the users’ desire for uniqueness Habbo goers are able to express their true or idealised identity via acquiring and possessing products such as avatars and virtual furniture that are not recognised as being outside of the norm. To encourage the Habbo goers to purchase and possess new and vintage virtual furniture it is designed to convey product-scarcity, uniqueness, and nonconformity. On the other hand, the Habbo goers exploit the opportunity to use avatars reflecting their offline appearance, expressing sometimes provoking messages, or releasing themselves from offline social norms (Vasalou, Joinson, Bänziger, Goldie & Pitt, 2008). For empowering the Habbo goers to seek status and uniqueness it is hypothesised that:

**H4:** Self-exploration (UNI) positively affects the attitude toward using Habbo.

**H5:** Status (STA) positively affects the attitude toward using Habbo.

In SVWs, where individuals are geographically separated and it is literally impossible to touch one another, they share the illusion of being physically in the same virtual environment in the guise of fictional characters conveying and embodying real-time socio-emotional behaviour such as movements and facial expressions (Bente, 2008; Bailenson, 2005). The feeling of closeness of another individual refers to the concept of social presence (PRE) in a mediated interaction (Biocca, Harms & Burgoon, 2003; Rice & Love, 1987; Short, 1976). If the shared virtual environment and the existence of avatars enable the users to experience non-verbal cues, they both facilitate the experience of social presence in virtual encounters and strengthen the positive evaluation of using Habbo. Therefore, it is postulated that:

**H6:** Social presence (PRE) affects positively toward using Habbo.

Development of the scales to measure the evaluation summary of attitude toward behaviour of interest is generated based on the suggestions of Ajzen (2005). The construct of attitude is the degree to which using Habbo is valued. In order to investigate the suitability of the construct of attitude used in prior IS literature, the construct of attitude should thus reflect the evaluation of performing the actual behaviour of interest since it would be challenging to specify every action beyond using technologies with multiple functions. The attitude of construct must therefore be accordingly determined. As a result, the following hypotheses are proposed:
H7: Attitude toward using Habbo mediates the actual willingness to social interaction with one another.

H8: Attitude toward using Habbo mediates the actual willingness to collect virtual furniture.

H9: Attitude toward using Habbo mediates the actual willingness to play games within Habbo.

Figure 1. Research model

3 EMPIRICAL RESEARCH

3.1 Research design and data collection

First, to explain the key factors why Habbo is actually being used the attitudinal beliefs were drawn from prior literature. Second, the research model was developed on the basis of TPB and the constructs as represented in IS research with one exception: those constructs founded on the key factors indicating the actual behaviour beyond using SVWs were formed using exploratory factor analysis (EFA). After the key factors were found, pilot study was conducted prior to the primary data collection. Finally, the goal of the empirical part of the study was to validate and test the research model (Figure 1) with confirmatory factor analysis (CFA), and investigate causal relationships between the latent variables.

The actual online survey instrument was built, and, in July 2008, the quantitative data from the users of Habbo Finland was collected. The survey was available approximately for a week and it contained worded items which respondents had to evaluate on a 7-point Likert-scale adapted from existing measures. To receive only one response from a single respondent a reward of any kind was not offered to the respondents (O’Neill & Penrod, 2001). In addition, multiple submissions were software-disallowed by excluding any responses from the same computer. The results reported here, are the responses of those willing to participate. That is to say, those Habbo goers who were the focus of the study were able to reach out.

A total of 3265 usable responses were received. To ensure the best possible quality of the responses, only fully completed responses were totally included in the analysis. After excluding cases with missing or incomplete responses 1225 fully completed and usable cases were retained for the analysis. Large drop of valuable responses was due to a long survey. 94 % of the respondents were between 10 and 18 years old.

To test the applicability of the research model AMOS 16.0.1, the maximum likelihood estimation method was applied here since it has been considered robust against the moderate violation of
multivariate normality with sample sizes exceeding 100 cases (Muthén & Kaplan, 1985; Steenkamp & van Trijp, 1991). The distortion of chi-squares and standard errors is likely if most of the skewness and/or kurtosis are larger in absolute value than 2.0, and the correlations are 0.5 or higher. According to Bollen (1989), all normal distribution estimations methods, including maximum likelihood, are consistent even if the assumption of normality is moderately violated. The test for normality showed that the skewness of only a few observed variables was higher than 2.0, thus, implying that the data is moderately non-normal, and therefore maximum likelihood would be an appropriate and robust estimation method. Asymptotical distribution free is suggested as superior to maximum likelihood only when the observed variables have an average univariate kurtosis larger than three and the sample size greater than 400 (Hoogland & Boomsma, 1998). The assessment of normality showed that every univariate kurtosis was below three.

3.2 Measurement Model and Results

The adapted constructs of the research model were evaluated by examining internal consistency and convergent validity. To do this, the item-construct-loading, composite reliability, and average value extracted (AVE) were assessed. As regards item reliability, item loadings exceed 0.707 (Agarwal & Karahanna, 2000), the composite reliabilities 0.707 (Nunnally, 1978) and AVE values 0.50 (Fornell & Larcker, 1981). The information on the loadings of the measures used in the research model, descriptive statistics, and the reliability of the constructs are represented in table 1. The results showed that all reflective measures fulfilled the recommended levels of composite reliability and average variance extracted.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Loading</th>
<th>Composite Reliability</th>
<th>AVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENJ1</td>
<td>0.909</td>
<td>0.953</td>
<td>0.723</td>
</tr>
<tr>
<td>ENJ2</td>
<td>0.958</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENJ3</td>
<td>0.933</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLA1</td>
<td>0.869</td>
<td>0.842</td>
<td>0.593</td>
</tr>
<tr>
<td>PLA2</td>
<td>0.837</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CON1</td>
<td>0.813</td>
<td>0.841</td>
<td>0.592</td>
</tr>
<tr>
<td>CON2</td>
<td>0.889</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNI1</td>
<td>0.907</td>
<td>0.953</td>
<td>0.769</td>
</tr>
<tr>
<td>UNI2</td>
<td>0.923</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNI3</td>
<td>0.902</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNI4</td>
<td>0.921</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STA1</td>
<td>0.888</td>
<td>0.893</td>
<td>0.618</td>
</tr>
<tr>
<td>STA2</td>
<td>0.909</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRE1</td>
<td>0.872</td>
<td>0.859</td>
<td>0.601</td>
</tr>
<tr>
<td>PRE2</td>
<td>0.864</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATT1</td>
<td>0.915</td>
<td>0.936</td>
<td>0.713</td>
</tr>
<tr>
<td>ATT2</td>
<td>0.906</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATT3</td>
<td>0.910</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOC1</td>
<td>0.818</td>
<td>0.851</td>
<td>0.597</td>
</tr>
<tr>
<td>SOC2</td>
<td>0.902</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRA1</td>
<td>0.935</td>
<td>0.943</td>
<td>0.641</td>
</tr>
<tr>
<td>TRA2</td>
<td>0.954</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAM1</td>
<td>0.778</td>
<td>0.785</td>
<td>0.564</td>
</tr>
<tr>
<td>GAM2</td>
<td>0.829</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Loadings, Composite Reliability, and AVE
Discriminant validity was investigated by examining whether the AVE was higher than the squared correlation for each construct (Fornell & Larcker, 1981). The correlations provided clear evidence that the relationship among all constructs were below the AVE in question (Table 2). Model fit indices address (GFI=.921, TLI=.954; NFI=.953, CFI=.962, and RMSEA=.058) that the proposed model was acceptable with the exception that AGFI (0.896) went slightly below the recommended threshold (see (Gefen, Straub & Boudreau, 2000).

<table>
<thead>
<tr>
<th></th>
<th>GAM</th>
<th>SOC</th>
<th>TRA</th>
<th>ATT</th>
<th>ENJ</th>
<th>PLA</th>
<th>CON</th>
<th>STA</th>
<th>PRE</th>
<th>UNI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAM</td>
<td>0.564</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOC</td>
<td>0.047</td>
<td>0.597</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRA</td>
<td>0.036</td>
<td>0.036</td>
<td>0.641</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATT</td>
<td>0.224</td>
<td>0.209</td>
<td>0.160</td>
<td>0.713</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENJ</td>
<td>0.113</td>
<td>0.105</td>
<td>0.081</td>
<td>0.503</td>
<td>0.723</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLA</td>
<td>0.088</td>
<td>0.082</td>
<td>0.063</td>
<td>0.393</td>
<td>0.637</td>
<td>0.593</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CON</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.010</td>
<td>0.002</td>
<td>0.592</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STA</td>
<td>0.034</td>
<td>0.032</td>
<td>0.025</td>
<td>0.154</td>
<td>0.206</td>
<td>0.416</td>
<td>0.098</td>
<td>0.618</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRE</td>
<td>0.076</td>
<td>0.070</td>
<td>0.054</td>
<td>0.336</td>
<td>0.504</td>
<td>0.490</td>
<td>0.011</td>
<td>0.274</td>
<td>0.601</td>
<td></td>
</tr>
<tr>
<td>UNI</td>
<td>0.017</td>
<td>0.016</td>
<td>0.012</td>
<td>0.077</td>
<td>0.103</td>
<td>0.205</td>
<td>0.061</td>
<td>0.249</td>
<td>0.158</td>
<td>0.769</td>
</tr>
</tbody>
</table>

Table 2. Squared pairwise correlations and assessment of discrimination validity

In total, the results support six of the nine hypotheses. Contrary to initial postulations, connectedness (CON), self-exploration (UNI), and status (STA) had no significant effect on attitude (ATT). Playing games (PLA) (0.13) alongside social presence (PRE) (0.12), a noteworthy exception of those measures associated with social outcomes, had only a moderate impact on ATT, whereas ENJ (0.52) had the strongest influence, a finding consistent with past research. On the other hand, the path coefficients from ATT to becoming friends (SOC) (0.46), trading in furniture (TRA) (0.40), and GAM (0.47) were statistically significant and interpretable if only somewhat surprising, given that the strength of paths from CON, UNI, and STA were statistically insignificant. The attitudinal constructs examined within this study paper explained 52 percent of the variance in the Habbo goers’ evaluation of using the service. ATT, for one, explained 21 percent (SOC), 16 percent (TRA), and 22 percent (GAM) of the variance in Habbo goers’ actual behavioural incentives. Figure 2 represents the path coefficients and the squared multiple correlations.
4 DISCUSSION

4.1 Findings

The results showed that 52 percent of the variance in the Habbo goers’ evaluation of using Habbo attitude (ATT) was shaped by enjoyment (ENJ), playfulness (PLA), and social presence (PRE), indicating that those social outcomes, such as connectedness (CON), self-exploration (UNI), and status (STA), had no significant influence on attitude. However, a post hoc analysis revealed that when paths from connectedness, self-exploration, and status were directly connected to becoming friends (SOC), trading in furniture (TRA), and playing games (GAM), the actual behavioural incentives beyond use were better explained. Thus, in the context of SVWs with multiple functions aimed particularly at the young, the users are not only craving for pleasure but social interaction with one another. These findings are in line with empirical evidence presented in prior literature examining virtual communities (Bagozzi & Dholakia, 2002; Rheingold, 2000; Wellman & Gulia, 1999) and leisure oriented technologies (van der Heijden, 2004). Taken as a whole, the users who volitionally have engaged with SVWs are especially driven by enjoyable social activities, such as playing games, collecting virtual items, and becoming friends with other users.

4.2 Implications for theory and practice

Empirical evidence on how both hedonic and social outcomes can be included in a research model adapted to virtual worlds, has been somewhat an open question. This study paper examines this issue by demonstrating, that among the experienced young users, hedonic and social aspects are pivotal factors for successful hedonic information systems. Contrary to prior studies that focus on social influence (i.e. social pressure to perform behaviour), this study has introduced the lack of a sophisticated attitudinal construct that measures the need for social interaction (cf. e.g. the distinct between playing video games alone or with other individuals). Therefore, the represented research model expands upon the current continued use model by underscoring the user’s need for enjoyable experiences and social interaction, postulating a direct path between them and the attitude toward continued use of SVWs. First, by identifying the actual key factors behind using SVWs the study enhances theoretical knowledge of developing behavioural frameworks built upon the IS research. Second, since practitioners are responsible for designing attracting virtual environments, the study has focused on providing better managerially relevant in-depth understanding of true behavioural incentives that contributes to the building of user commitment. The results suggest that managers should not concentrate only on the technology use but the actual reasons why it is being used. When adding or removing features it is important to be aware how attitudinal beliefs affect the actual behavioural incentives. Thirdly, by particularly examining the experienced Habbo goers the study accumulates the post-adoption user behaviour research (e.g. Limayem & Cheung, 2008; Premkumar & Bhattacherjee, 2008).

To develop a robust research model for examining the use behaviour, it is likely that the items related to the attitude toward using SVWs must cover the social aspect as well. Since only a few, if any, studies that focus on developing behavioural constructs to measure the use of SVWs exist, this study paper offers a theoretical implication for IS research.

4.3 Limitations and future research

Employing IS theories is rather challenging when examining the use of any technological innovation serving multiple purposes. Consequently investigating SVWs is confined to dealing only with factors regarding attitude toward using SVWs, determined by the individual’s assessment of the outcomes associated with the behaviour (Ajzen, 2005). Determinants such as subjective norms and perceived behavioural control have been deliberately omitted. The future research is, thus, required to examine how subjective norms, and perceived behavioural control, in addition to attitude, affect continued use intention in the context of SVWs. Furthermore, large drop of valuable responses due to a long survey
may have limited the accuracy of the results. Also, generalising these results into context other than SVWs should be handled with care. It is proposed to take into account that the results depend on respondents’ subjective assessment (see Straub & Limayem, 1995) and their socio-cultural backgrounds that may have a significant impact on usage behaviour. This is of great importance especially when the service is globally available. Finally, this study represents a cross-section of behavioural incentives beyond using SVWs. When the appropriate measures are discovered longitudinal study examining post adoption behaviour may provide better understanding of the users’ behaviour, their motivation, and their intentions.

5 CONCLUSIONS

Providers of SVWs should focus on the actual behavioural incentives beyond using the technology. Creating interactive and hedonic services to the young is a fundamental goal which is virtually as central as the technical characteristics of SVWs. The present study paper exemplifies a decomposed theory of planned behaviour in order to represent an important step toward examining and developing the construct of attitude toward using SVWs with multiple functions aimed at the young. Most importantly, using Habbo as an example, study paper empirically demonstrates how TPB-grounded research model explains a significant amount of variance in the Habbo goers’ evaluation of using SVWs but, at the same time, relegates the social aspects into the background. The findings indicate that, in the context of SVWs, the construct of attitude toward using the technology, must include items related to social outcomes together with hedonic outcomes.

ACKNOWLEDGEMENTS

I would like to state my gratitude to Finnish Youth Research Society, Foundation for Economic Education, and Sulake Corporation for assistance in conducting this study, as well as Jussi Puhakainen, Tommi Hoikkala, Mikko Ketokivi, and Matti Määttymäki.

References


**APPENDIX 1**

<table>
<thead>
<tr>
<th>Construct</th>
<th>Measurement Item (on a 7-point Likert-scale)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATT1</td>
<td>Extremely displeasing...pleasing</td>
<td>Karahanna et al., 1999; Ajzen, 2005)</td>
</tr>
<tr>
<td>ATT2</td>
<td>Extremely frustrating...easy; a great pleasure to use</td>
<td></td>
</tr>
<tr>
<td>ATT3</td>
<td>Extremely terrible...delightful</td>
<td></td>
</tr>
<tr>
<td>SOC1</td>
<td>I like getting to know new people</td>
<td>Global Habbo Youth Survey 2006; 2008</td>
</tr>
<tr>
<td>SOC2</td>
<td>I like making friends with other Habbo users</td>
<td></td>
</tr>
<tr>
<td>TRA1</td>
<td>I like collecting rare furniture</td>
<td></td>
</tr>
<tr>
<td>TRA2</td>
<td>I like collecting valuable furniture</td>
<td></td>
</tr>
<tr>
<td>GAM1</td>
<td>I like organising and playing games as well as participating in competitions</td>
<td></td>
</tr>
<tr>
<td>GAM2</td>
<td>I like playing games</td>
<td>Ruvio, 2008; Snyder &amp; Fromkin, 1977; Tian, 2001</td>
</tr>
<tr>
<td>UNI1</td>
<td>I actively seek to develop my personal uniqueness by using special products or brands</td>
<td>Ruvio, 2008; Snyder &amp; Fromkin, 1977; Tian, 2001</td>
</tr>
<tr>
<td>UNI2</td>
<td>Having an eye for products that are interesting and unusual assist me in establishing a distinctive image</td>
<td></td>
</tr>
<tr>
<td>UNI3</td>
<td>The products and the brand I like the best are the ones that express my individuality</td>
<td></td>
</tr>
<tr>
<td>UNI4</td>
<td>I am often on the lookout for new products or brands that will add to my personal uniqueness</td>
<td>Ruvio, 2008; Snyder &amp; Fromkin, 1977; Tian, 2001</td>
</tr>
<tr>
<td>ENJ1</td>
<td>It is fun to use Habbo</td>
<td>Venkatesh &amp; Brown, 2001; Davis et al. 1992; van der Heijden, 2004</td>
</tr>
<tr>
<td>ENJ2</td>
<td>It is entertaining to use Habbo</td>
<td></td>
</tr>
<tr>
<td>ENJ3</td>
<td>It is pleasant to use Habbo</td>
<td></td>
</tr>
<tr>
<td>PRE1</td>
<td>There is a sense of sociability in Habbo (users are companionable)</td>
<td>Short, 1976</td>
</tr>
<tr>
<td>PRE2</td>
<td>There is a sense of human warmth in Habbo</td>
<td></td>
</tr>
<tr>
<td>STA1</td>
<td>Using Habbo improves my status among those who are richest and smartest</td>
<td>Venkatesh &amp; Brown, 2001</td>
</tr>
<tr>
<td>STA2</td>
<td>Using Habbo improves my status among those who are the most meaningful to me</td>
<td>Venkatesh &amp; Brown, 2001</td>
</tr>
<tr>
<td>CON1</td>
<td>I feel so distant from others in Habbo</td>
<td>Lee &amp; Robbins, 1995; Lee &amp; Robbins, 1998</td>
</tr>
<tr>
<td>CON2</td>
<td>I have no feeling of togetherness with others in Habbo</td>
<td></td>
</tr>
<tr>
<td>PLA1</td>
<td>Using Habbo increases my interest in exploring things</td>
<td>Ghani &amp; Deshpande, 1994; Koufaris, Kambil &amp; Labarbera, 2001; Koufaris, 2002</td>
</tr>
<tr>
<td>PLA2</td>
<td>Using Habbo arouses my imagination</td>
<td></td>
</tr>
</tbody>
</table>
On the evolution of online tourism communities - Network Battle or Long Tail Niches?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0728.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Virtual community, Empirical study, Web 2.0, Tourism</td>
</tr>
</tbody>
</table>
ON THE EVOLUTION OF ONLINE TOURISM COMMUNITIES

Network Battle or Long Tail Niches?

Felix-Robinson Aschoff, Department of Informatics, University of Zurich, Binzmühlestrasse 14, CH-8050 Zürich, aschoff@ifi.unizh.ch

Gerhard Schwabe, Department of Informatics, University of Zurich, Binzmühlestrasse 14, CH-8050 Zürich, schwabe@ifi.unizh.ch

Abstract

Even though the emergence or respectively the construction of online communities is of great interest for scientists and community engineers, only few empirical data has been presented on community growth. This article starts with a reflection on possible growth curves of virtual communities. It contrasts a network externality perspective that produces clear winners and losers in a market with a long tail perspective that also allows small niche products to be successful. These considerations are empirically tested with a sample of 74 travel communities whose numbers of registered members were recorded at two measure points. The results show that online travel communities develop into an archetypical long tail. A very small number of communities with exceedingly high numbers of members are accompanied by a vast amount of communities with only few members. An analysis of the long tail, however, reveals that the community tail is not dead but is populated by a large number of especially regional communities that show considerable growth rates.

Keywords: evolution, online communities, growth rates, network effects, long tail
THE EVOLUTION OF ONLINE COMMUNITIES

Online communities have been the object of extensive research within the last 25 years (for an overview see Preece 2000; for a multi-disciplinary topology Porter 2004). Within these decades of research the phenomenon has been highlighted from different perspectives. One commonly cited source is Rheingold (1993) who describes online communities as technological enablers for communication among community members. The social interaction is the focus of the author’s attention and the mutual support (e.g. in case of illnesses) is vividly described. Hagel & Armstrong (1997) in contrast, focus on the economic potential online communities offer to companies. The authors see them as platforms to redefine and strengthen a company’s relationship to customers, suppliers, and competitors. From the customer’s point of view, Hagel & Armstrong (1997) conceive communities as the option to be a more informed bargainer than in traditional markets. For the domain of tourism see Wang et al. (2002) who take the more social perspective by Preece as well as the economic perspective by Hagel & Armstrong to discuss implications for marketing and design of travel communities.

These two perspectives are reflected by approaches for the construction of communities as well. Preece (2000) proposes a community-centered approach which is related to a member or user-centered approach. In accordance with a participatory-design approach Preece claims user needs and social aspects to be crucial pillars for the development of a community (another practical guide on community development is provided by Kim (2000)). Hagel & Armstrong (1997), again, conceive community members more as potential customers. They claim that communities fulfil four basic needs: interest, relationship, fantasy and the need for transaction and bartering (like e.g. purchasing goods). In accordance with this potential-customer perspective their development strategies focus on binding the members to the community. Consequently, their strategy to reach a community’s critical mass point comprises three stages: generating user traffic, concentrating traffic and locking in traffic.

Just like Hagel & Armstrong (1997) and Hummel & Lechner (2002), we also conceive Online Communities as phenomena that realize network effects. This implies that the benefit users receive from using a product or a virtual community is influenced by the number of persons who also use this service or product. For most communication products within networks like traditional telephone networks it is the case that the personal benefit increases with a rising number of other persons who also use the network (cf. Shapiro & Varian 1999). Crucial aspects of a networks’ evolution is the fact that they can evoke positive as well as negative feedback loops and that they follow an S-shaped curve in three phases: a) flat during launch, b) a steep rise during takeoff as positive feedback kicks in and c) leveling off as saturation is reached (ibid.). Figure 1 depicts the assumed growth curve of virtual communities based on this network effect assumption.

![Figure 1: Assumed growth curve of online communities due to network effects (cf. Shapiro & Varian 1999)](image)
From a practitioners point of view it is relevant to know at which number of users the accelerated growth curve starts for a specific community. If a common point or interval could be estimated for a specific community domain, it would be possible to predict the growth of different communities within the near future.

In accordance with this idea of a network market, virtual communities within one domain (e.g. English-speaking forums covering travel-related questions) cannot be considered to be independent from each other. Traditional off-line communities are to a strong degree bound to geographical aspects as well as aspects of affiliation. This means that originally one did not choose a community but one was born into one. Geographical boundaries as well as barriers between social classes were difficult to transgress. Even though this has changed to some degree since the industrial revolution, off-line communities are still rather a fate than a choice. This situation is considerably altered by the net situation: It provides the real choice to become a member of one community or the other.

This, however, implies that forum communities within one confined domain can be said to compete for users like products on a market compete for customers. From a neutral users’ point of view it should always be more beneficial to join a community that has at least reached the take-off phase (for evidence that higher user activity indeed leads to better information quality within a community see Prestipino et al. 2006, Aschoff et al. 2007). If one community within a domain reaches this point, this would on principle be a good precondition for a monopolistic development. Wikipedia seems to be an interesting example of this situation. Only one main Wikipedia platform has developed up to now, even though some of the described topics on this platform are highly controversial. The positive network effects are so strong that users do not seem to be willing to launch competing knowledge sharing platforms. Figure 2 shows this division into winners and looser in a network market. Winners are said to enter a virtuous cycle while loser are said to enter a vicious cycle (Shapiro & Varian, 1999).

Thus, a market with strong network effect tends to produce clear winners that are able to gain big revenues as well as clear losers that are not able to reach the positive feedback zone. This creates a situation in which a company either has to gain considerable amounts of the market share or will seize to exist.

In contrast to this situation, we see the conception of a long tail market as described by Anderson (2006). In these markets, products can be profitable in very small niches due to the decreased costs of warehousing and distribution in the digital economy (for a description of a long tail market for the business of online bookselling see Brynjolfsson et al. 2003). In these markets, the aggregated revenue from all the products that sell only in small numbers can equal or exceed the few hit products that sell in vast amounts. Thus, from an economic perspective it pays off to offer niche products for a small customer segment as long as the distribution costs are minuscule.
A market can show network effects as well as a long tail distribution. The characteristics of this market will, however, depend on how dominant either of these influences are. A market that is dominated by network effects will, as described above, favour one winner and one or more losers. This is especially given when systems or products are incompatible with each other, i.e. the user of one system cannot communicate with the user of another system. This can eventual lead to format wars (e.g. the recent competition between the HD DVD and the Blue-Ray format; for numerous additional examples see Katz & Shapiro 1994). In contrast to this, a long tail market allows products with far smaller consumer than the hit products to be “economically sustainable”. For a related controversy that contrasts a long-tail assumption that favours “underdogs” with a head-of-the tail assumption that favours “superstars” refer to Elberse & Oberholzer-Gee (2008). Applied to the market of online communities these perspectives lead to different predictions:

a) The network effect perspective would predict a winner-takes-it-all market, i.e. few large online communities with considerable growth rates are accompanied by a number of small communities with stagnating or declining growth.

b) The long tail perspective would predict small communities in the tail to be successful as well, i.e. they would show user activity and growth rates.

2. Empirical Method

To empirically test these two assumptions, we assembled a comprehensive collection of online communities that dealt with travel-related questions. This search started in February 2008 and we only selected platforms that met the following characteristics:

a) In this research, we focussed on classical online forum platforms. These forums are characterized by a communicational structure in which one community member starts a thread by sending an initial post (like a question) and possibly receives a number of replies. In contrast to wikis, product recommendation sites or newer social networking platforms (like e.g. www.facebook.com), these online forums have been developing on the Internet for more than 20 years. By now they cover a wide area of topics including cooking, health-related issues, computer problems or mobile phone technology. For an early account of the online forum community “The Well” founded in 1985 refer to Rheingold (1993). To obtain a more homogeneous sample, we confined our sample to browser-based forums.

b) The main purpose of the forum had to be travel-related information exchange. Prominent examples are the Thorn Tree forum1 run by the Lonely Planet Publisher Group or the Virtual Tourist platform2. The forum members had to discuss typical travel related topics (like e.g. How do I get to a certain place? How can I solve travel related problems? What are suggestions for attractive travel routes or locations? etc.).

c) The language used by the community member to communicate had to be English.

The following search engines were used for this research: Google, Yahoo!, Altavista, Live Search by Microsoft as well as the meta search engine Mamma. In addition Boardreader was used which is a search engine specialized on finding online forums. The used key words comprised: “Travel Forum”, “Traveler Forum”, “Travel Board”, “Independent Traveler Forum”, “Independent Traveler” and “vbulletin travel”.

The number of registered members was registered at two measure points. The first measure point was between February and July 2008 and the second measure point was in November 2008. The number of

1 http://www.lonelyplanet.com/thorntree/index.jspa
2 http://forum.virtualtourist.com/
registered users was measured by reading out the respective number from the websites of the communities.

3. Results

The described search procedure resulted in 120 travel-related online discussion forums. The first measure point was between February and July 2008 and the second measure point was in November 2008. The number of the community members is not displayed by all communities and some communities seized to exist or were temporarily offline at one of the two measure points. In addition to this, two communities were so massively spammed at the time of the second measure point that travel related communication was hardly taking place anymore. We excluded these forums from the sample. Finally, we were able to measure 74 online communities at two points in time. Figure 3 shows the 74 communities with their respective numbers of member at the second measure point in November 2008.

![Figure 3: Number of registered members at the second measure point (Nov. 2008) for each forum](image)

The distribution of the absolute numbers of registered members shows an archetypical long tail for this community sample. Based on this distribution we can roughly distinguish three groups of communities that are separated from each other by about one order of magnitude. We have one community which claims to have more than 100,000 registered members. The second group consists of three communities between 100,000 and 300,000 members. The third group starts with communities with about 50,000 registered members and entails 70 communities whereas the smallest communities only have between 10 and 20 members. The data shows that those forums above 50,000 constitute only about 5% of the forums in the sample. One the other hand, note that the forum with the highest acclaimed number of members already has more members than all the remaining forum communities.

We calculated the absolute growth rates of registered members as well as the growth rates relative to the absolute number of registered members at the first measure point for each community per month. Overall, communities in our sample grew by an absolute average of 393 members per month which relates to a relative monthly growth of 17.7%. The absolute growth rate of the largest community with about 100,000 members amounted to 8085 members per month (0.78%). The second group between 100,000 and 300,000 members had two forums with only marginal or even negative monthly growth (924.5 members (0.76%) and -1155.2 members (-0.66%). The third forum in this group, however, showed a considerable growth rate with 14267.56 additional members per month (11.16%).
Figure 4 shows the absolute growth rates per month of communities below 30,000 registered members. The Loess fit line (Epanechnikov-Kernel with 50% of points to fit, entire sample included) indicates a steady increase in the absolute number of new members per month with considerable derivations. Overall, the absolute increase of members correlates significantly with the size of the community at the first measure point (Pearson = .336). This part of the sample consists of 70 communities out of which 62 communities show a positive growth level while 8 communities show zero growth or a negative growth level. The average growth level of these communities amounts to 17.66% (7.56% without one extremely fast growing community).

Thus, while this tendency is hard to prove due to the small number of communities at the head of the long tail, this data may indicate a certain saturation point. Overall we see a clear tendency that the biggest relative growth does not happen in the head of the curve but in the tail. Figure 5a shows these relative growth values for the range between 0 and 5,000 members. Figure 5b shows the range between 5,000 and 30,000 members.
Again, the data show a considerable variance among the different forums. Most of the forums grow between 1 and 20%. A notable high number of forums show increased growth rates in the interval between 0 and 500 members. To gain insights into the considerably high growth level at the end of the long tail we categorized our community sample into different subsamples.

**What grows in the long tail?**

As mentioned we see an archetypical long tail with very few communities with exceedingly high numbers alongside a vast amount of communities with only few members. Figure 6 shows the number of communities at the very end of the long tail only depicting communities with less than 10 000 members.

![Frequency histogram for communities with less than 10 000 members](image)

*Figure 6. Frequency histogram for communities with less than 10 000 members*

Note that 53 communities fall into the range between 0 and 10 000 community members. This is already 72% of the entire community sample whereas the largest communities have more than 100 000 members. In addition, the interval with the most communities within this range lies between 0 and 500. The network effect approach as well as the long tail approach would expect a diversification of the communities at the end of the tail. From a user perspective a community in the tail should provide some special good that cannot be offered by the big communities. Only based on the probability to obtain good answers to posted questions or the chance to meet many new people, the larger communities should always be in advantage of gaining new users. To test this assumption we divided our sample into three groups. The first group of communities are communities that on principal cover all regions of the globe. The second group are communities that are specialized on certain regions like specific continents or countries. The third group, finally, focusses on special interest groups like mounting biking, diving, or travelling by motorcycle.

Figure 7 shows the end of the long tail with communities below 10 000 members divided by these three groups. As could have been expected the ratio between worldwide and regional forums changes if the sample is divided by this threshold. Above 10 000 the sample contains 10 worldwide forums and 7 regional forums, and below this threshold the sample contains 10 worldwide forums and 37 regional forums.
Thus, we can conclude that there is a vast number of small communities that grow at a considerable rate. We assumed that regional forums in the end of the long tail would grow faster than worldwide. To diminish the effect of the extreme outliers within the sample we transformed the relative growth variable into a ranking scale and performed a non-parametric Kruskal Wallis Test. The mean ranks for the global forums as well as for the regional forums proved to be almost identical (36.47 and 37.42) whereas the specialized forums had lower mean ranks of 24.39. The test failed to reach significance with p=.212. Thus, the data indicate that small global communities grow at a similar rate as small regional communities.

Limitations
The research is yet limited due to the fact that the number of registered members in a forum is only a very indirect measurement of the activity within a forum. Future research has to substantiate our findings by additional measures like e.g. the number of written posts per month recorded over a longer period of time.

4. Discussion of the results

With respect to the absolut numbers of registered members our sample shows an archetypical long tail. Very few travel communities with an exceedingly high number of members (ca. 100 000 members and more) are accompanied by a large amount of small communities (between ca. 10 members and ca. 30 000 members). With respect to growth rates, however, the data show that this long tail is not dead but shows considerable average growth rates. Hence, the long tail does not seem to be populated by losers but rather by communities that are developing in a promising way. This leads to the question why end user would join a smaller travel community if they also have the chance to join a “market leader”. This can have a number of reasons:

a) Differentiation. Some small communities might offer a special service or cover a special aspect that is not provided by the communities at the head of the tail. This might be related to special travel-related interests (e.g. diving, hiking) or special region-related coverage (e.g. communities specializing only on Thailand). However, our current data does not support this claim. There are more specialized communities at the long end of the tail compared to communities with a worldwide coverage. However, these worldwide communities show growth rates that are similar to the growth rates of the more specialized communities. Further analysis considering additional characteristics (like e.g. degree of commercialization etc.) has to show whether variables can be found that separate the long tail communities from the
communities in the head of the tail. Another possibility is that the size of the community is a differentiating characteristic by itself. A number of Internet users might prefer smaller communities for different reasons: the risk of information overload is decreased, it is easier to gain a higher social status and the feeling of social belonging might be stronger in a smaller community.

b) Intransparency of the online community landscape. The idea of the network effect driven community market assumes that Internet end users have a choice to become a member of any arbitrary community. Obviously, this implies that Internet user have an overview over a more or less transparent market situation. Since central repositories for online (travel) communities are not very widely spread, most users probably do not even know what kind of choices they have. This might lead to a situation in which a user engages into the first online community she finds more or less suitable without being aware of other options. Further research especially in the field of information behaviour can yield fruitful insight as to how Internet users actually search for virtual communities and how they decide which one to join.

These results also underline the fact that online communities assumably do not only have positive feedback loops but also negative feedback loops as they increase in size. This negative feedback loop is caused by information overflow that can be caused if too many persons ask questions (this means for example that each question gets less attention) or too many replies are given. In addition, the transaction costs are increasing for a single user if she has to manage relationships to many points within the social network of the community. This aspect sets online forums apart from cooperatively constructed online encyclopedia like Wikipedia. At least for the vast amount of passive consumers of Wikipedia, an increased user community only has the positive effect of increased quality without the negative effects of information overload or increased transaction costs. This might have led to the monopolistic position wikipedia is holding today.

This argumentation points to a relationship between community size and perceived value of an online forum that is rather characterized by a U-shaped function opening downward than by a steady increase as depicted in Figure 1. Further research should focus on the lower as well as on the upper boundaries of this value curve. This is also related to the question whether online forums need a minimum number of members for a sustained development. Our current data shows considerable relative growth rates for communities between 10 and 500 members. Additional future measure points of our sample will reveal whether these developments are persistent over time. Practical implications of this research arise especially for the field of online community development. Approaches in this domain up-to-now have often been of a prescriptive or design-oriented nature. These rules of community development, however, have to be complimented by quantitative models of online community evolution that are empirically well founded.

Acknowledgements
We thank Johnny Nia, Salvatore Bonaccorso and Mary-Anne Kockel for supporting the data collection. This research was partly funded by the Richard Büchner-Stiftung as well as by the Gebert Rüf Stiftung.

References


IT-CONTROLLING IN FEDERAL ORGANIZATIONS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0553.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>IT governance, IT/IS management, IT regulation, IT Business Alignment / Value</td>
</tr>
</tbody>
</table>
Abstract

This paper investigates the characteristics of federal and modular organizations and elicits conclusions on their requirements for IT controlling through a literature review. The literature review showed that different organizational structures create specific conditions concerning IT and IT controlling. Although experience in the regulation and controlling of IT in large and complex organizations has been reported, the characteristics of these specific organizational conditions and the resulting requirements for the design of an IT controlling concept have not been extensivively researched. Creating the missing link between the characteristics of federal and modular organizations and their requirements regarding IT controlling may serve as a foundation for future research and the development of a comprehensive IT controlling concept which encompasses the characteristics and key drivers of this specific organizational form.

Keywords: IT controlling, IT governance, federal organizations, modular organizations.


1 INTRODUCTION

Alignment between business strategy and information technology (IT) strategy is regarded as a key driver for realizing value from IT investments (Henderson and Venkatraman 1999, Luftman 2006). This high level of importance may result from the fact that IT investments constitute a major part of IT costs but the benefits of IT usage are not always obvious and therefore difficult to justify. This phenomenon is often described as ‘productivity paradox’ (e.g. Brynjolfsson 1998, Carr 2003). Nevertheless some organizations manage to specify accountabilities for IT-related business outcomes better than others because of more effective IT governance (Weill and Ross 2004). Weill and Ross have researched forms of IT governance and classified underlying structures by the location where IT decisions are made. According to this classification, six types of decisions from ‘central decisions’ to ‘decentralized decisions’ are to be distinguished:

- Business monarchy: Decisions taken by a member of the management or a group of managers;
- IT monarchy: Decisions taken by the IT director or a group of IT directors;
- Federalism: Decisions taken by executives of the middle management of all operative divisions and the integration of the IT direction is also considered;
- IT duopoly: Decisions taken by IT direction and a group of members of the management;
- Feudalism: Decisions taken autonomously by respective divisions;
- Anarchy: Decisions taken autonomously by a user or a group of users.

This classification is not only relevant for decision making but also to inform decision makers about the origin of input (Weill and Ross 2005). A study conducted on 197 mainly ‘Global 1000’ companies showed that firms with a federal IT organization had a significantly higher IT/business alignment maturity than others (Luftman and Kempaiah 2007). A cluster analysis of 40 companies by Gordon and Gordon (2002) showed similar results. Nevertheless a federal IT organization is “no silver bullet” (Luftman and Kempaiah 2007) and federal IT governance is described as demanding a great deal of management attention (Weill 2004). For example, in large and complex multunit organizations with interdependent information resources where federal IT is particularly useful, conflicts and coordination difficulties are likely to occur. Tsai (2002) describes the phenomenon of “coopetition” in which subunits of large multunit organizations which are supposed to cooperate become instead competitors when it comes to using internal resources and are therefore likely to reject information sharing.

How can a federal organizational form for structuring IT be attained while avoiding its disadvantages? One solution is using a shared controlling concept across units to enable federal organizations to make use of the advantages of a federation and to steer clear of intra-organizational competition (Wenninger-Zeman 2003). Current research, however, has not considered specific IT controlling styles as they relate to the organizational and governance perspectives that characterize organizations. The available literature offers a broad range of tools and concepts for controlling IT (e.g. Krcmar 2005, Weill and Ross 2005) and has developed various approaches for avoiding intra-organizational competition in multunit organizations (e.g. Brass et al. 2004, Schaefer 2008). Still, there is a lack of research evidence which supports combining IT controlling concepts with the special requirements of federal organizations.

The objective of this paper is to understand the reasons for the contradictions that exist between the theories and practice of IT controlling in federal organizations. On the one side, many IT controlling concepts and elaborated general organization types do exist. On the other side, the interrelation between a specific organizational form and the need for IT controlling is missing and thus many organizations are not capable of effectively controlling their IT. This paper uses organizations governed by federalism as an example and describes the key drivers of federally governed organizations and their specific needs for IT controlling. Further, existing IT controlling concepts are compared and their suitability for federal organizations is evaluated. The following research questions are addressed in this paper:

1. What are the constitutive elements of federal organizations and which key drivers characterize their specific (IT) controlling needs?
2. Which experiences, approaches and implementations for structuring and designing IT controlling already exist?

3. What are possible appropriate approaches and concepts for successful IT controlling in federal organizations?

2. LITERATURE REVIEW

2.1 Parameters of Federal and Modular Organizations

The word federalism is derived from Latin ‘foedus’ (confederation, confederacy, treaty, alliance) (Rudolf 1981). Frantz (1962) called federalism the leading principle for the social, governmental and international organization. According to him, the structure of a state has to be federative to achieve political freedom. Such a structure is characterized by districts and provinces having their own legislation which they advance autonomously (Frantz 1949). The larger a state is, the stronger the central power needs to be. Federalism described from the political perception includes larger autonomous political entities formed by the union of smaller political units who maintain their autonomy as well as the existence of coequal statehood of the whole state and the member states (Thöni 2005). A pure political perception of the term is not sufficient (Kinsky 2004); federalism as aggregation of uniformity and diversity can rather be a model for a great number of societal structures even beyond the state, for example in companies, associations, clubs or unions.

Although, the term federalism is rarely used in a context outside of societal structures, the underlying principles of autonomy, cooperation, solidarity, contractual or consensual conflict resolution, two-way control and distribution of power, subsidiarity and participation, are the same. Autonomy is based on self-determination of the particular members of a federal structure as well as the voluntary collaboration within the federal organization. Cooperation means that conflicts between units and the federal organization are not being solved by power, but based on specified authorizations. The specific units operate in solidarity. Compromises are often the conclusion of conflict resolution. Two-way control of federal units is realized by equal distribution of power between the units. Decisions are, according to the principle of subsidiarity, made where they occur. By contrast, competencies have to be transferred to the headquarters where reasonable.

Handy (1995b) expanded the established understanding of federalism to non-governmental environments and described federal organizations independent from purpose and scope of the organization. While the headquarters of a typical organization may be the center of decision-making, it is characteristic for federal organizations that initiative and dynamics result mainly from the subunits. According to Handy, the emergence of federalism in organizations is not conscious but emerges rather because the core of the organization cannot cope with all the information that is being provided by the decentralized units. As many organizations downsize their headquarters, they stop information overload and stop centralized control of the organization. That is when, as stated by Handy, decentralization turns into federalism.

The headquarters of federal organizations only define long-term objectives and leave the implementation of the objectives to the subunits. However when making decisions, headquarters must consider the opinions of the subunits. This is described as a place where persuasion has to be achieved and discussions lead to consensus (Handy 1995a). Constraints will be accepted on a subunit level if the acceptance of constraints benefits the superordinate unit. Picot, Reichwald and Wigand (2003) describe this type of organization as modular characterized by being split in legally autonomous units. The relatively small headquarters takes over coordinating tasks whereas the subunits are capable of acting legally autonomously and handle the more operational tasks (Picot et al. 2003). With few staff, management develops long-term strategy and coordinates cross-sectional activities. Following the creation of units in the modular organization, management must keep the number of interaction dependencies as low as possible (Weber 2001). Small units are characterized by flat hierarchies,
simple structures, and low division of work, which, in combination with personal responsibilities and integration of functions, leads to long range autonomy (Weber 2001). The strengths of both centralized and decentralized units have to be recognized and utilized accordingly. The advantages of specialization are either in the specificity of processes of customers, in the specificity of overall organizational infrastructures, or cross-specific functions (Picot et al. 2003). Therefore, tasks of the first group, where knowledge about specific customer-oriented workflow for problem solving is important, should be handled in the decentralized departments. In contrast, tasks with a high impact of overall methodical and technical aspect for problem-solving should be undertaken by centralized departments (Picot et al. 2003).

Table 1 summarizes the characteristics of federalism and federal and modular organizations leading to the requirements of IT and IT controlling in federal organizations. The structure of an organization is important to the thesis of this paper in terms of analyzing interrelations between the organizational form of federalism and IT controlling. Governance principles are taken into account as they influence controlling decisions (Weill and Ross 2004). The principles of cooperation are important as they are a main source of conflict (Tsai 2002).

<table>
<thead>
<tr>
<th></th>
<th>Federal public administration</th>
<th>Federal organizations</th>
<th>Modular organizations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organizational structure</strong></td>
<td>Association of smaller political units who maintain their autonomy to larger, autonomous political entities;</td>
<td>Aggregation of uniformity and diversity; centralization of strategic decisions; decentralization of operational decisions;</td>
<td>Operational activities: Subunits capable of entrepreneurial acting, legally autonomous Management, long-term planning and coordination of cross-sectional functions: centralized;</td>
</tr>
<tr>
<td><strong>Governance principles</strong></td>
<td>Coequality of superordinate and subordinate units;</td>
<td>Two-way control and distribution of power (checks and balances);</td>
<td>Coordinative function of headquarters;</td>
</tr>
<tr>
<td><strong>Principles of cooperation</strong></td>
<td>Authorizations and laws;</td>
<td>Contractual or consensual conflict resolution: large amount of information handling cannot be centralized;</td>
<td>Split-up the organization in legally autonomous units e.g. by core competencies, business division or region;</td>
</tr>
<tr>
<td><strong>Attributes</strong></td>
<td>Autonomy; Independence, being part of two institutions at the same time;</td>
<td>Autonomy, cooperation, solidarity, subsidiarity, participation, initiative and dynamics, subunits – retention of the headquarters, culture of discussions and consensus, being part of two institutions at the same time;</td>
<td>Responsibility of subunits, few interaction dependencies to resign a voluminous interface management; flat decentralized hierarchies, simple structures and low division of work, autonomy, profit responsibility;</td>
</tr>
</tbody>
</table>

Table 1: Characteristics of federalism, federal and modular organizations

The structure of any federal administration, company, or modular organization is crucial for organizational embedding of IT. Thus, the integration of the value-added chain requires organization-wide coordinated IT systems which support the coordination of autonomous units and guarantee the supply of information for each unit (Picot et al. 2003). Subunits responsible for the handling of a special task can be connected via IT infrastructure. To guarantee access to essential data at any time and to guarantee problem-oriented handling of data, a continuous integration and networking of all operational information systems is required (Picot et al. 2003). Coordination and cooperation of the particular units is realized by the means of IT through common and shared information databases and knowledge databases (Picot et al. 2003).

In the context of governance of federal organizations, the main aim is the localization of IT and IT controlling decisions (Weill and Ross 2004). Due to changing market conditions, the localization of decisions might be subject to change and different types of federalism can emerge over time. Depending on current governance structures, different requirements for IT and consequently for IT controlling arise. Although headquarters might delegate IT-related responsibilities to subunits, headquarters must retain control of IT in terms of being informed about operations performed in the
organizations to monitor and if necessary take corrective action on IT matters (Weber 2001). Weber (2001) proposes to provide the responsible divisions with a criterion for performance measurement and to communicate at what point headquarters is expected to intervene. To design the process of control comprehensively, individual agreement on the objectives for the unit and the documentation and review of compliance with these objectives is required.

The delegation of service activities is one trait of decentralization that impacts on the functions of controlling in an organization (Horváth 2006). It is assumed that a high level of delegation at the formation of a (controlling) system leads to a higher differentiation of the created system. In addition to spacious and technical characteristics, the organizational aspect of centralization and decentralization of information systems has to be taken into account (Lehner et al. 1991). This aspect specifies the degree of decentralization in planning, implementing and maintaining systems. A central solution has the advantage of a simpler construction of integrated solutions with coordinated data and being able to meet the information demand of management. Furthermore, centralized IT reduces the risk of redundant work and incompatibilities and facilitates the operation of organization-wide application systems. Similarly, creation, implementation and application of tools and standards are simplified and calculating load is optimized using a central IT organization. In contrast, in an organization with autonomous divisions, modifications in IT must be made promptly and units must be flexible in order to satisfy the needs of the decentralized units. Usually, the IT staff in decentralized units is more experienced than staff in centralized IT units in dealing with the problems of a particular division. Table 2 summarizes the requirements regarding IT and IT controlling that result from the characteristics and parameters of federal organizations as found in the literature review and described above.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Requirements regarding IT</th>
<th>Requirements regarding IT controlling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizational structure</td>
<td>Integration of decentralized organizational tasks; Coverage of management’s information demand; Central provision of organization wide data;</td>
<td>Provision of decentralized information for management; Centralization of controlling or decentralization with central administration;</td>
</tr>
<tr>
<td>Governance principles</td>
<td>Coordination of organizational units; Flexibility for prompt reaction on decentralized demands;</td>
<td>Knowledge about decentralized processes for supporting complex problems; Provision of decentralized information for management; Documentation of compliance with strategic objectives by decentralized units; Provision of criterion for performance measurement; Controlling spin-off as a service task;</td>
</tr>
<tr>
<td>Principles of cooperation</td>
<td>Decentralized storage and availability of data for special decentralized activities and processes; Coverage of management’s information demand;</td>
<td>Agreement on objectives for units; Definition of points of intervention of headquarters; Assignment of responsibilities for operational processes;</td>
</tr>
<tr>
<td>Attributes</td>
<td>Decentralized storage and availability of data for special decentralized activities and processes; Provision of information for decentralized organizational units; Coverage of management’s information demand;</td>
<td>Provision of decentralized information for management; Monitoring of compliance with operational objectives in decentralized units; Documentation of compliance with operational objectives by decentralized units; Enabling of objectives correction; Enabling of incentive and sanction scheme; Centralization of controlling or decentralization with central administration; Differentiation of the controlling system by the demands of the decentralized units; Provision of a reporting system; Assignment of responsibilities for operational processes;</td>
</tr>
</tbody>
</table>

Table 2: Requirements regarding IT and IT controlling
2.2 IT Controlling

IT controlling is, according to Krcmar (2005), the control of IT-related operations in the organization. The goal is to ensure efficiency and effectiveness of IT operations while providing, quality, functionality and compliance to deadlines in information processing. IT controlling has a monitoring function as well as a coordination function for the management of information.

An institutional and a functional view of controlling can be distinguished (Britzelmaier 1999). A multiplicity of controlling conceptions is discussed in the literature with differing emphasis depending on the application field. Vöhringer (2004) for instance, differentiates between profit-oriented, reporting system-oriented or key figure-oriented, and coordination-oriented controlling conceptions. The Anglo-American research area rarely uses the term controlling at all (Schauer 2006); it is being replaced by the associated contentual questions. This is why there is a differentiation made between IT/IS (Information Systems) (investment) evaluation, IT/IS (performance) measurement, and measurement of IT/IS costs, and benefits. Table 3 presents IT controlling concepts published within the last five years. The selection of the concepts follows criteria proposed by Schauer (2006): The presented approaches are to give an overview about functions and methods of IT controlling and not to be limited to some aspects. Moreover, the concepts should not be older than five years. They are presented in alphabetical order of the authors.

The controlling concepts presented in Table 3 serve as a basis to research possible criteria and starting points for organization-specific adjustments regarding controlling frameworks for federal organizations. The implementation of controlling in federal organizations should be a combination of central and decentralized controlling. Whereas central IT controlling deals with strategic planning, decentralized controlling is concerned with the implementation of the controlling concept in a particular division. Central strategy development, planning, controlling and regulating allow the longer-term alignment of an IT landscape to the corporate strategy in the subunits of federal organizations. The aim of strategy development is the definition of a nominal condition and to derive options and needs for action. On the basis of options and needs for action, agreements on objectives are made with the subordinated units and the objectives are connected with corresponding indicators (operating figures). An essential part of the IT strategy is the longer-term alignment of IT on decentralized operational processes. The planning of IT intentions and IT projects is, therefore, necessary. Decisions are not usually made at the operational unit level in federal organizations. By analyzing all possible interdependencies, the strategic relevance and effectiveness of the IT portfolio of the complete organization can be guaranteed (Krcmar 2000).
<p>| Objectives | Strategic objective: Effectivity; Operational objective: Efficiency; Quantitative &amp; qualitative objectives | Alignment of IT support on organizational objectives; | Formal objectives: Efficiency, Effectivity; Real objectives: Quality, Functionality, Compliance of deadlines; | Trade-off between supply and demand of IT performance; Consideration of goods and services and utilization; Utilization of goods and services, divisions; Utilization of goods and services, organization wide; Classification as per objects; Support of divisional IT controlling, optimization of organization wide IT controlling in the foreground; | Strategic information systems planning; | Basis for planning in IT; Means for decision making, cost reduction, performance assurance; Motivation for employees; Profitability; |
| Functional View | Coordination; Evaluation; | Coordination; Process orientation; Object orientation; | Coordination; Process orientation; Object orientation; | Task-oriented; | Profit orientation; Product orientation; |
| Institutional View | Derived from organizational structure/strategy; | Derived from organizational objectives; | Controlling of IT in the organization; | Controlling of IT in the organization; | Controlling of IT in the organization; |
| Objects | Strategy, Projects Operating Applications; IT-Infrastructure; Cost-performance-management; Organization of IT-division; | Strategy/Projects; Operating Applications; | Portfolio Controlling; Project Controlling; Product Controlling; Infrastructure Controlling; | Project; System; Process; Service; | Acquisition and processing of information; Human resources; Technical infrastructure; Applications; | IT product controlling; Controlling of IT resources; Project controlling; |</p>
<table>
<thead>
<tr>
<th>Functions</th>
<th>Planning; Organization; Service management;</th>
<th>Evaluation of the strategic relevance of IT; Strengths/Weaknesses, Opportunities/Threads (IT degree of maturity) Process oriented planning; Multi-project</th>
<th>Compliance of strategic relevance; Compliance of profitability; Planning; Evaluation &amp; selection of projects; Compliance of</th>
<th>Portfolio controlling; Preparation of and compliance to SLA; Evaluation, selection, initiation &amp; realization of projects; Regulation of resource management;</th>
<th>Strategic task for maintenance and protection of the organization in terms of reactivity and adaptability by use of information technology; Administrative tasks for coordination of planning, regulation and information tasks;</th>
<th>Cost and activity accounting; Allocation of costs;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>management; Project controlling; Relationship management/ service provider &amp; service receiver;</td>
<td>quality, functionality; Monitoring of Product lifecycle; Regulation and advancement of infrastructure;</td>
<td>Operative tasks for monitoring the organization and its environment; Counteractive measures in the sense of an early warning system;</td>
<td>Portfolio analysis; Organization in profit centers; Accounting for services; Make-or-buy-decisions; Appointment of quota of fixed costs and overhead costs;</td>
<td>Portfolio arrangement; Reporting management ratio; Benchmarking;</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Survey of introduced IT controlling concepts
There as a complete controlling concept comprises a holistic view on controlling, the core of a controlling concept is its applied methods. One of the widely-used methods is the Balanced Scorecard (BSC) which was first introduced by Kaplan/Norton (Kaplan and Norton 1992) and represents the central method of IT controlling (Rehäuser 1999, Rehäuser and Krcmar 1995). The BSC is the ideal controlling method for federal organizations because it is not limited to the presentation of the entire organization but can be used for the controlling of particular organizational domains, divisions or projects (Heilmann 2001). In the range of federal organizations, ratio systems are best used for organization-wide controlling that reflects the objectives of particular subunits. Such a controlling ratio system can be derived from the BSC. Its adoption does not occur with regard to its content, but rather its structure. Activity and cost data of IT have to be represented in management ratios in a way that using benchmarking, comparisons between the subunits of federal organizations can be made to increase transparency. Process oriented cost accounting offers the possibility to measure costs where they emerge (Aurenz 1990). A prerequisite for the strategic use of results of IT controlling in federal organizations is an established reporting system of the decentralized units among each other as well as between the subunit and headquarters. Because recipients of the reports make decisions in different areas and vary in their need for information, it is appropriate to consider, recipient, form, and date of the report when creating reports or planning the reporting system (Tiemeyer 2005).

Table 4 compares the requirements of IT controlling in federal organizations to possible approaches of existing IT controlling concepts structured by methods and instruments. Both columns result from a comparison of the literature reviews on IT controlling and federal organizations depicted above, conducted by the authors.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Requirements regarding IT controlling</th>
<th>Possible IT controlling solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizational structure</td>
<td>Provision of decentralized information for management; Centralization of controlling or decentralization with central administration;</td>
<td>Methods and instruments: Profitability analysis, benefit evaluation, ratio system, reporting systems, BSC, cost accounting, portfolio analysis, build-up experience database, Benchmarking, determination of fixed costs and overhead costs quota</td>
</tr>
<tr>
<td>Governance principles</td>
<td>Knowledge of decentralized processes for solving complex problems; Provision of decentralized information for management; Documentation of compliance with strategic objectives by decentralized units; Provision of criterion for performance measurement; Controlling spin-off as a service task;</td>
<td>Methods and instruments: Profitability analysis, benefit evaluation, ratio systems, reporting systems, BSC, cost accounting, portfolio analysis, build-up experience-database, benchmarking, organization profit centre, make-or-buy-decision, determination of fixed costs and overhead costs quota</td>
</tr>
<tr>
<td>Principles of cooperation</td>
<td>Definition of objectives for units; Definition of intervention time by headquarters; Assignment of responsibilities for operational processes;</td>
<td>Methods and instruments: Profitability analysis, benefit evaluation, ratio systems, reporting systems, BSC, cost accounting, portfolio analysis, build-up experience-database, benchmarking, organization profit centre, make-or-buy-decision</td>
</tr>
<tr>
<td>Attributes</td>
<td>Provision of decentralized information for management; Monitoring of compliance with operational objectives in decentralized units; Documentation of compliance with operational objectives by decentralized units;</td>
<td>Methods and instruments: Profitability analysis, benefit evaluation, ratio systems, Reporting systems, BSC, cost accounting, Portfolio analysis, Build-up Experience-DB, Benchmarking,</td>
</tr>
</tbody>
</table>
Enabling of objectives correction;  
Enabling of incentive and sanction scheme;  
Centralization of controlling or decentralization with central administration;  
Differentiation of controlling systems depending on needs of decentralized units;  
 Provision of a reporting system;  
Assignment of responsibilities for operational processes;  
Organization Profit centre, Make-or-buy-decision, determination of fixed costs and overhead costs quota

| Table 4: Requirements of federal organizations and possible solutions |

3 CONCLUSION

In summary, the major challenges for IT controlling in federal organizations lie in the provision of information about the decentralized, operational units for the centralized, strategic management. In addition, the decentralized units have to document their compliance with strategic objectives and their performance must be measureable. The literature review shows that existing controlling concepts use different methods and instruments to meet the IT controlling requirements of federal and modular organizations. The unique characteristics of decentralized units in federal and modular organizations make it difficult to implement one particular concept. This literature review focuses on German publications. Future reviews should include international literature. Research in this area should explore the application of existing instruments and methods of IT controlling in federal organizations in order to enable the transfer of information by controlling. Different theories, such as principal-agent theory or contingency theory, might provide an explanation as to how to improve the relationship between centralized and decentralized units in federal organizations. The resulting explanations could be further investigated in practice, for example by conducting case studies in a real world federal organization.
References


HOW GERMAN HOSPITALS GOVERN IT – AN EMPIRICAL EXPLORATION

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0565.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>IT governance, IT Project Management, E-health, Empirical study</td>
</tr>
</tbody>
</table>
HOW GERMAN HOSPITALS GOVERN IT – AN EMPIRICAL EXPLORATION

Köbler, Felix, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany, felix.koebler@in.tum.de
Fähling, Jens, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany, jens.faehling@in.tum.de
Leimeister, Jan Marco, Universität Kassel, Nora Platietl Str. 4., 34127 Kassel, Germany, leimeister@uni-kassel.de
Krcmar, Helmut, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany, krcmar@in.tum.de

Abstract

Health care services in German hospitals are causing immense expenses. Successful IT Governance might help to support specific challenges for every organization with an adequate use of IT. The market structure of hospitals in Germany is very heterogeneous, e.g. in size and sponsorship. This paper analyses the state of the art of IT Governance based on a survey among 220 IT executives in German hospitals. The quantitative analyses of collected survey data reveal that hospitals govern their IT differently according to size and sponsorship. In addition, our analyses show that decision-making authority for the IT budget rises with hospital size and is positively correlated with the fraction of IT projects in the overall IT budget. We also show that the investments in innovative IT projects increase with hospital size. Our study revealed that a high number of private and larger hospitals lack a systematic IT Governance approach within the decision domain on IT projects. This study is the first to shed light into the empirical situation of IT Governance in German hospitals.

Keywords: IT Governance, IT Management, German hospitals, eHealth.

1 INTRODUCTION

The expenses for healthcare services in the German healthcare sector were estimated at 239 billion euros in the year 2003, which constitutes an 11.1% share of the gross domestic product of the Federal Republic of Germany (Destatis 2006). Based on this background Germany is listed in the Top 10 of OECD countries for per capita expenses on health care services (Anderson et al. 2006) and according to the Federal Healthcare report of the Federal Statistical Office (2008b), about 59 billion Euros are spent in the hospital sector.

In many industrial branches innovative information technology (IT) and its use are key drivers for increasing effectiveness and efficiency in production processes for goods, services and successful business processes (Leimeister et al. 2009). The effect of IT usage in healthcare is found in medical service provisions (Jähn and Nagel 2004, Schweiger et al. 2007) and administrative support processes (Haas 2005, Lehmann 2005). Hacker and Schommer (2004) report on increased effectiveness and efficiency in examination, treatment and administrative processes in hospitals. IT usage can be a driver for diversification in competition and the creation of innovative strategic competitive advantages in hospitals and the health care sector (Piccoli and Ives 2005). Multiple studies report success of IT-driven improvements in administrative and business processes, e.g. information systems for the input and integration of treatment data, reminder functions, medicament management and medication...
(Crane and Raymond 2003, Leimeister et al. 2005, Raymond and Dold 2002). Since the early 1990s, studies have shown that cost reductions are not only generated by the automation of information accumulation and processing but additionally by contemporary and optimized information allocation for decision makers (Borzekowski 2002). This fact verifiably induces improved treatment quality (Apkon and Sighaviranon 2001, Hacker and Schommer 2004) and fostering of patients (Leimeister et al. 2008b, Nahm and Poston 2000).

Despite the economic importance of the health care sector and important role of hospitals, Information Systems research has not been able to generate empirically collected data and declarative and conceptual models on IT management and information allocation in hospitals. A number of recent studies are either driven conceptually or are missing hypotheses testing on an empirical basis (Sachs 2005). Other studies focusing on special tasks of IT management, e.g. investment and budget decision making (Bernnat 2006), are missing the application of statistical analysis methods (Riedel 2006) or lack statistical significant results because of low participation numbers (Irving and Nevo 2005).

To close the gap in academic research and based on the above-mentioned immense expenses for health care services in Germany, a unique and empirically broad state of the art descriptive study on IT Governance in German hospitals was conducted.

1.1 Motivation for the study

In conducting this study, we were inspired by the approach taken by Weill and Ross (2004), who describe two sources of motivation when conducting a study on IT Governance in corporations. Weill and Ross were motivated by the stock-market premiums given to firms with excellent corporate governance and therefore “suspected a similar premium existed for excellent IT governance” (Weill 2004) and the “fact that relatively sophisticated financial governance in most enterprises could provide a good model for IT Governance” (Weill 2004). In addition, we identified a previously described gap in academic research on IT management and specifically IT Governance in German hospitals. Our motivation derives from the assumption that differences in IT Governance structures and operationalization exist across hospital size and sponsorship. The foundation of our motivation derives from the results of survey studies conducted in the scope of the Krankenhaus-Barometer (Blum et al. 2007), which indicates differences in types of sponsorship and hospital size.

<table>
<thead>
<tr>
<th>Future Objectives</th>
<th>Hospital Sponsorship</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>public mean</td>
<td>rank</td>
<td>non-profit mean</td>
<td>rank</td>
</tr>
<tr>
<td>patient satisfaction</td>
<td>4.6</td>
<td>1</td>
<td>4.6</td>
<td>1</td>
</tr>
<tr>
<td>high process quality</td>
<td>4.5</td>
<td>2</td>
<td>4.5</td>
<td>2</td>
</tr>
<tr>
<td>good image of the hospital</td>
<td>4.5</td>
<td>3</td>
<td>4.4</td>
<td>3</td>
</tr>
<tr>
<td>intensification of the relationship between hospital and medical practices</td>
<td>4.2</td>
<td>4</td>
<td>4.1</td>
<td>4</td>
</tr>
<tr>
<td>revenue increase</td>
<td>3.6</td>
<td>10</td>
<td>3.4</td>
<td>11</td>
</tr>
<tr>
<td>realization of profit</td>
<td>3.4</td>
<td>11</td>
<td>3.4</td>
<td>12</td>
</tr>
<tr>
<td>achieving the highest possible return on investment</td>
<td>2.7</td>
<td>12</td>
<td>2.8</td>
<td>13</td>
</tr>
</tbody>
</table>

*Table 1. Importance of future objectives across hospital sponsorship (extract from (Blum et al. 2007))*

Table 1 displays the importance of hospital future objectives based on arithmetic mean and ranks results according to the type of hospital sponsorship. Although, the primary future objectives are identical across the three types of hospital sponsorship, we also see major differences. In all types of sponsorship, patient satisfaction, high process quality, the good image of the hospital and intensification of the relationship between hospital and medical practices, are the most fundamental objectives. However, private hospitals focus more on economic objectives than public (e.g. university)
and non-profit hospitals (e.g. comparable by type of sponsorship to organizations like the Red Cross). Increasing revenue, realizing profit and achieving the highest possible return on investment are ranked higher by survey participants from privately sponsored hospitals. We reflect that the realization of the mentioned objectives is achievable through a value proposition of IT utilization based on findings by Fähling et al. (2009), who found that IT utilization traces back to decisions and that these decisions are primarily influenced by the IT Governance. To consolidate our assumption of the influence of IT Governance, we hereby derive (1) revenue increase by the utilization of administrative information systems that support process cost controlling and increase the rate of return on capital employed through the allocation of (information) resources, (2) realization of profit by the utilization of medical information systems that enable new business models and treatment methods and (3) highest possible return on investment by optimized and more efficient treatment methods supported by medical information systems.

Table 2 displays composite results of a survey conducted by the Deutsches Krankenhausinstitut (Blum et al. 2007) which describes the differences in hospital size (measured by the number of beds) on the basis of three survey items. We assume that the chosen items: cooperations between hospitals, mergers between hospitals and new allocation of tasks between medical and non-medical staff, call for different IT Governance cultures and models.

<table>
<thead>
<tr>
<th>Survey item (frequency of occurrence)</th>
<th>50 to 299 beds</th>
<th>300 to 599 beds</th>
<th>&gt; 600 beds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperations between hospitals since 2004</td>
<td>43.9 %</td>
<td>53.3 %</td>
<td>56.5 %</td>
</tr>
<tr>
<td>Mergers between hospitals since 2004</td>
<td>7.5 %</td>
<td>11.1 %</td>
<td>16.1 %</td>
</tr>
<tr>
<td>New allocation of tasks between medical and non-medical staff</td>
<td>14.0 %</td>
<td>23.0 %</td>
<td>38.7 %</td>
</tr>
</tbody>
</table>

Table 2. Importance of objectives across hospital size (extract from (Blum et al. 2007))

Cooperations and mergers between hospitals require a flexible and interoperable IT landscape, which should have consequences for the IT Governance — especially in large hospitals. To enable and support many reallocations of tasks between medical and non-medical staff, IT managers should have a high degree of decision-making authority, and must understand the requirements of their medical and non-medical clients.

2 IT GOVERNANCE IN HOSPITALS

Weill and Ross (2004) state that “the difference between management and governance is like the difference between a soccer team running harder and practicing longer and the stepping back to analyze its composition and game strategy”. The result of an analysis may be that a team needs to introduce new coaches or different playing positions or provide diverse decision making responsibilities. Therefore a company and respectively a hospital organization needs to involve different people in “IT decisions, designing new ways or making IT-related decisions, or developing new techniques for implementing IT decisions” (Weill and Ross 2004) to achieve more value from IT artifacts. Along with other definitions of IT Governance (IT-Governance-Institute 2000, Krcmar 2005, Van Grembergen 2003), Weill defines IT Governance as “specifying the framework for decision rights and accountabilities to encourage desirable behavior in the use of IT” (Weill 2004). The definition implies a strict separation between management and governance. Whereas management talks about the specific decisions that are made, governance “is about systematically determining who makes each type of decision (a decision right), who has input to a decision (an input right), and how these people (or groups) are held accountable for their role” (Weill 2004). Consequently, the IT Governance framework proposes five major decision domains and six exclusive governance archetypes for making IT decisions. These were adopted and linguistically adjusted and translated into German to fit common expressions used in praxis and the German health care sector. Table 3 displays information on the
decision domains according to Weill (2004) and the corresponding terminology used in the underlying study and analyses (see Figure 3).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IT Principles</td>
<td>IT Strategy</td>
<td>High-level statements about how IT is used in the business</td>
</tr>
<tr>
<td>IT Architecture</td>
<td>IT Standards</td>
<td>Standardization of technical capabilities that should be standardized enterprise-wide to support IT efficiencies and facilitate process standardization and integration. Activities that must be standardized enterprise-wide to support data integration</td>
</tr>
<tr>
<td>IT Applications</td>
<td>IT Applications</td>
<td>An integrated set of technical choices to guide the organization in satisfying business needs. The architecture is a set of policies and rules for the use of IT and plots a migration path to the way business will be done (includes data, technology, and applications)</td>
</tr>
<tr>
<td>IT Infrastructure Strategies</td>
<td>IT Infrastructure</td>
<td>Strategies for the base foundation of budgeted IT capability (both technical and human), shared throughout the firm as reliable services, and centrally coordinated (e.g., network, help desk, shared data)</td>
</tr>
<tr>
<td>Business Application Needs</td>
<td>IT Projects</td>
<td>Specifying the business need for purchased or internally developed IT applications</td>
</tr>
<tr>
<td>IT Investment and Prioritization</td>
<td>IT Investments / IT Budget</td>
<td>Decisions about how much and where to invest in IT including project approvals and justification techniques</td>
</tr>
</tbody>
</table>

Table 3. Decision domains according to Weill (2004)

Based on considerations of corporate governance, state governance and information politics, Weill and Ross identified six IT governance archetypes; namely, Business Monarchy, IT Monarchy, feudal, federal, IT Duopoly and anarchy, through a logical combination of above mentioned decision maker types. Table 4 displays the IT Governance archetypes which describe the combination of people who have either decision rights or input in IT decisions (Weill and Ross 2004) and short descriptions outlining each archetype.

<table>
<thead>
<tr>
<th>Governance Archetype</th>
<th>Corporation decision makers</th>
<th>CxO Level Executives</th>
<th>Hospital decision makers</th>
<th>Hospital and/or polyclinical1/ clinical directors</th>
<th>Corp. IT and/or Business Unit IT</th>
<th>Central hospital IT and/or clinic IT</th>
<th>Business Unit Leaders or Process Owners</th>
<th>Clinic directors or head of department/ assistant med. director</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Monarchy</td>
<td>A group of, or individual, business executives (i.e., CxOs). Includes committees comprised of senior business executives (may include CIO). Excludes IT executives acting independently.</td>
<td>x</td>
<td>Hospital and/or polyclinical1/ clinical directors</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT Monarchy</td>
<td>Individuals or groups of IT executives.</td>
<td></td>
<td>Hospital and/or polyclinical1/ clinical directors</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feudal</td>
<td>Business unit leaders, key process owners or their delegates.</td>
<td></td>
<td>Hospital and/or polyclinical1/ clinical directors</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Polyclinic is a clinic for ambulant therapy.
Federal: C level executives and at least one other business group (e.g., CxO and business unit leaders)—IT executives may be an additional participant. Equivalent to a country and its states working together.

IT Duopoly: IT executives and one other group (e.g., CxO or business unit leaders).

Anarchy: Each individual user

Table 4: Six IT Governance archetypes (Source: according to Weill (2004) and Weill and Ross (2004))

Additionally, Table 4 shows the results of the organizational and hierarchical entity mapping between corporation and hospital decision makers. The aim of the underlying study is to deliver first insights regarding the empirical situation of IT Governance in German Hospitals by conducting a descriptive analysis of collected data.

3 RESEARCH DESIGN

3.1 Research method

In order to explore the IT Governance in German hospitals, we conducted 12 expert interviews on IT Governance in German hospitals to structure the research objective and questionnaire. The data was collected through a standardized online questionnaire. The questionnaire was adjusted linguistically to fit the professional domain terminologies of IT executives in German hospitals. The duration of the survey was three months, from March to June 2008. The questionnaire forms were pre-tested among ten experts and adjusted in advance where required. Address data was collected from commercially available domain-specific address data collections and internally compiled address data sets. After consolidation of address data sources and validation of doubles, 2391 different hospitals and medical institutions could be identified as potential participants. The numbers show that multiple contact persons in the same professional domain per institution/hospital were contacted. All mailings included an individual code to avoid multiple participations. The average response time for the questionnaire was approximately 30 minutes. The data collection included two follow-up calls for participation through email and postal mailings. Additionally, two articles describing the study, including a call for participation, were published in professional journals: Krankenhaus-IT Journal and Management & Krankenhaus. The results presented below are based on descriptive analysis of relative frequencies, t-tests and correlations with the use of the correlation coefficient by Spearman.

3.2 Structure

After two researchers independently and iteratively conducted a data cleaning process, 206 data sets (11 anonymous) were collected from IT executives through the online questionnaire. More than two thirds of the IT executives (70%) hold an academic certification, whereas 30% graduated with an apprenticeship certificate.
Table 5. Sample structure (n=200)

For further analysis, the data set was segmented based on two segmentation attributes: type of sponsorship and size of the hospital as discussed previously. The size of the hospital was measured in number of beds. We adopted the classification of the hospital size from Leimeister et al. (2008a). Hospitals were classified into three categories “under 200 beds”, “200 to 799 beds” and “800 and more beds”. This segmentation is based on an interview series among experts from the German medical and hospital environment. The type of sponsorship is partitioned to three categories: “public”, “private” and “non-profit” and corresponds with the German hospital market structure.

<table>
<thead>
<tr>
<th>Sponsorship</th>
<th>Public</th>
<th>Private</th>
<th>Non-profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federal Statistical Office</td>
<td>36.01%</td>
<td>38.37%</td>
<td>25.62%</td>
</tr>
<tr>
<td>Sample</td>
<td>40.50%</td>
<td>19.00%</td>
<td>40.50%</td>
</tr>
<tr>
<td>Size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small (1-199 beds)</td>
<td>55.54%</td>
<td>40.54%</td>
<td>3.92%</td>
</tr>
<tr>
<td>Medium (200-799 beds)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Big (800+ beds)</td>
<td>20.00%</td>
<td>63.50%</td>
<td>16.50%</td>
</tr>
</tbody>
</table>

Table 6. Comparison between German hospital market structure and our sample (according to Destatis 2008a)

When compared with data from the Federal Statistical Office, our sample is not consistent with the statistical data about German hospital market structure (see Table 6). While our sample consists of 40.5% non-profit hospitals, only 25.62% of the hospitals in Germany are under control of a non-profit sponsor. In contrast, our sample contains only 19% of private hospitals, whereas 38.37% of hospitals in Germany are held privately. We also see differences in hospital size between our sample and the German hospital market structure. Within our sample only 20% of all hospitals are small compared with 55.54% in the German hospital market. By contrast large hospitals are overrepresented in our sample, with a share of 16.5% compared to 3.92% in the German hospital market. Based on the structure of the sample displayed in Table 5 and Table 6, we accept that the informative value of conducted t-tests might be limited and its application restricted.

4 EMPIRICAL DATA

4.1 IT-budget

In a first step, we analyzed the distribution of the IT budget. The respondents could assign 100% of the IT budget to three different categories. The first category Operations contains all costs for the operations of current IT systems in the hospital. The second category Projects includes all expenses for projects beyond operating tasks. This part of the IT budget is invested to develop new and innovative IT solutions within the hospital. The last category Organization involves all personnel expenses.

The three left bars in Figure 1 show the comparison between the hospitals by size (n=203). The rightmost bar represents the average values across all hospitals (n=203). The IT budget for the IT organization in large hospitals with 800 beds and more is slightly lower (33.9%) compared to 35.1% in medium and 36.6% in small hospitals. On the other hand, large hospitals invest more in IT projects (22.1%), compared to 21.8% in medium and 18.1% in small hospitals.
Bars number four to six from the left give an overview over the IT budget distribution across different sponsorships (n=200). Here we could identify differences in all budgeting areas. First, the part for IT operations represents almost half of the whole IT budget in private hospitals (47.3%) compared to 43.8% in public and 41.9% in non-profit hospitals. Second, the organizational part on the IT budget is much lower in private hospitals (30.1%) compared to 34.5% in public and 38.2% in non-profit hospitals. Finally, private hospitals invest a bigger share in IT projects (22.6%) compared to 21.7% from public and 19.9% from non-profit hospitals. T-tests could not reveal any significant differences in hospital size or sponsorship. This might result from the structure of the sample.

4.2 IT decision-making authority

The second question was related to the decision-making authority. Three types of decision-making authority were formulated in the questionnaire: the IT manager can only prepare but not make decisions, the IT manager is allowed to make decisions within a specific amount of budget and the IT manager may decide over the full amount of the IT budget (see Figure 2).
the full amount of the IT budget by their own. The percentage of IT managers which only prepare
decisions is almost the same across all hospitals (from 22.2% in public to 23.7% in private to 25.9% in non-public hospitals). Across hospital size, IT managers from large hospitals are more likely to be able to make decisions by their own (30.3% in large, 10.8% in medium and 5.0% in small hospitals) than in medium and small hospitals. This fact is also underpinned by a correlation between the freedom of IT budget decisions (decision-making authority index) and the number of beds (r=0.20; p=0.004; n=203; see Table 7). The decision making authority index is based on a three point scale: "prepare IT decisions" was mapped on a value of 1, "make decisions within a specific amount of budget" was mapped on a value of 2 and "make decisions within whole IT-budget" was mapped on a value of 3 (higher values represent higher decision authorities).

Another analysis of decision-making authority index and the different areas of IT budget revealed a positive correlation between the decision-making authority and IT projects (r=0.17; p=0.017; n=203). The more decision-making authority the higher the part of IT projects on the whole IT budget. One explanation is that IT managers with low budget responsibility have difficulties to convince other stakeholders of their ideas for IT projects. Another explanation is that IT managers which execute more IT projects receive more trust from the other stakeholders and therewith more authority for decision-making.

Another analysis of decision-making authority index and the different areas of IT budget revealed a positive correlation between the decision-making authority and IT projects (r=0.17; p=0.017; n=203). The more decision-making authority the higher the part of IT projects on the whole IT budget. One explanation is that IT managers with low budget responsibility have difficulties to convince other stakeholders of their ideas for IT projects. Another explanation is that IT managers which execute more IT projects receive more trust from the other stakeholders and therewith more authority for decision-making.

![Table 7. Correlations between decision making authority index (Source: sample data)](attachment:image)

Results of the t-test between hospital sponsorship across hospital size (see Table 8) disclose that differences in decision-making authority exist between big private hospitals and public (level of significance: 0.026) as well as non-profit (level of significance: 0.048) hospitals. The results need to be interpreted carefully as large public and non-profit hospitals are underweighted in the sample.

![Table 8. T-tests between hospital sponsorship across hospital size about decision-making authority (Source: sample data)](attachment:image)

4.3 IT decision domains

As previously described, we analyzed all IT decision domains across hospitals by classifying our sample in hospital sponsorship and size. The first IT decision domain we focused on is IT Strategy. Figure 3 shows that the fraction of the federal archetype increases and the fraction of IT duopoly decreases with hospital size. This can be explained by an increasing number of departments at larger hospitals. Decision processes for the IT strategy more often follow the archetype of an IT Monarchy in non-profit hospitals (34.2%) than in public (24.7%) and private (19.8%) hospitals. In public hospitals the most frequent archetype is federal, possibly because the share of large hospitals in public hospitals is higher than in private and non-profit ones, according to our sample data (Figure 3 and Table 5). Results for the decision domain IT Standards show that the most common archetype for IT standards is IT Monarchy (from 57.5% in small to 69.7% in big hospitals). This is not surprising, considering that decisions surrounding IT standards are often technically-driven, and do not primarily affect business process issues. In any case, almost one-fifth of the hospitals use a federal archetype in decisions on IT standards. The biggest share of the federal archetype can be found in private hospitals.
(23.5%), the smallest in non-profit hospitals (5.3%). Noticeable is the high share of IT duopoly in non-profit hospitals (15.8%) which is almost twice as high as in public hospitals (8.6%).

Next, we analyzed the decision domain IT Applications. In this decision domain, the archetypes are distributed very differently. IT Monarchy plays an unimportant role. The most mentioned archetype is federal. The fraction grows with the hospital size (42.5% in small, 58.5% in medium and 63.6% in large hospitals). This result shows that the clinical departments which use IT applications are mostly involved in decisions about these applications. In non-profit hospitals we identified a percentage of 18.4% and in private hospitals a percentage of 16% in which decisions on IT applications are made according to a Business Monarchy archetype — that means, without the involvement of IT decision makers and departments. The lower percentage of the federal archetype in non-profit hospitals is
derived from the high share of small and medium sized hospitals for this sponsorship type (see Table 5). In addition, the share of IT Monarchy in non-profit hospitals is much higher than in the other types of hospitals. One explanation might be that IT managers in non-profit hospitals place themselves near to medical and operating departments so they get more trust in selecting the best application for the requirements of their internal customers. The decision domain IT Infrastructure is dominated by the IT Monarchy archetype. The second important archetype is IT Duopoly. The federal archetype only plays a role in small hospitals (22.5%). The distribution of archetypes is highly similar across all sponsorship types. More than 50% of the decisions are made within a IT Monarchy governance type. These results demonstrate that IT infrastructure issues as well as IT standards are highly specified decisions of IT specialists in German hospitals.

The four main governance archetypes within the IT Investment decision area in German hospitals are IT Monarchy, Business Monarchy, IT Duopoly and Federal archetypes, with the last two archetypes being more dominant according to sponsorship and hospital size. The share of the IT Duopoly archetype is especially high in medium-sized (31.5%) and non-profit (31.6%) hospitals, whereas the share of the federal archetype is high in large (39.4%) and public (37.0%) hospitals. The results might be influenced by the share of large public hospitals in our sample data (24.7% compared to 15.8% of large private and 8.7% of large non-profit hospitals).

Our descriptive analysis on the IT Projects decision domain reveals that decisions on projects are mostly conducted in an IT Duopoly governance type. Just above 39.0% of the participants from larger hospitals confirm that an IT Duopoly governance is used for project decisions. A comparable picture describes the numbers derived from the survey data on small and medium sized hospitals where 32.5% and respectively 36.9% of participants indicate an IT Duopoly archetype for decisions on IT projects. Nevertheless, we see a relatively high fraction of the answer item “not applicable” (n/a), especially in medium (10.8%) and large (12.1%) as well as public (9.9%) and private (12.3%) hospitals. This indicates that the decision domain on IT projects does not exist in the mentioned sponsorship types and sizes of hospitals.

5 DISCUSSION

Based on different objectives and attributes between small and large as well as between private and non-private (public and non-profit) hospitals, we analyzed the IT governance approaches between hospital sizes and sponsorships. Overall, we concentrated our analyses on three topics: IT budget, IT decision-making authority and IT decision domains.

The relative distribution of IT budgets across German hospitals is dominated by IT budgets concerning IT operations. One explanation for this distribution could be that German hospitals deal with many legacy systems. In contrast, IT project-related budgeting which is invested to develop new and innovative IT solutions within the hospital, covers the smallest share. Although the partitioning of the IT budget differs between hospital size and sponsorship, these differences are not statistically significant.

In a next step, we identified a significant correlation between decision-making authority and the percentage of IT projects on the overall IT budget. We assume that IT managers invest mainly in IT projects. In other words, the more decision-making authority on IT budget, the more IT improvement through innovative IT project deliverables can be expected. Furthermore, with the help of t-tests we proved statistically significant differences between large and medium, as well as small hospitals in decision-making authority. In other words, the larger the hospital size measured in number of beds, the higher the decision-making authority of IT managers. We could also show that these significant differences only exist between private and non-private (public and non-profit) sponsorships.
We assume that private hospitals are more dedicated to increasing revenue and optimizing return on investment than hospitals under different sponsorship types based on findings of described similar studies. Our study revealed that IT managers from private hospitals try to support these objectives with a higher fraction of IT projects on the overall IT budget.

The third topic of the underlying study investigated the structures of IT decision domains in German hospitals. Our study revealed that in most hospitals a project-oriented specification of business needs for purchased or internally developed IT applications is not conducted in a systematic IT Governance approach. The high absence rate of a systematic IT Governance execution, especially in private and large hospitals in the decision domain IT Projects is very surprising as IT project-related budgeting is invested to develop novel IT solutions and drive innovation.

6 LIMITATIONS AND FURTHER RESEARCH

This study is not without its limitations. First, the study is limited by its data collection process and data set structure. The collected data represents a snapshot of reality and therefore conclusions on dynamics and timely progression cannot be derived. In the future, multiple and frequent data collection processes could lead to interesting findings over time. The data set structure is not representative along sponsorship and size segmentation. Second, the study only considers the perspective of chief information officers and IT managers in German hospitals, which might generate a bias on question items. Further analysis might be coupled with additional qualitative data collection, in order to create a deep and coherent understanding of these preliminary findings. In combination with further in-depth statistical analysis, the researchers plan to generate and test structural equation models to expose cause and affect chains in IT Governance in German hospitals and to identify the role of contingency factors in choosing the optimal IT Governance archetype. The generated findings could be used to discover similarities and differences to industry-related results and patterns discovered in comparable studies.

7 REFERENCES


### TOWARDS A SERVICE GOVERNANCE FRAMEWORK FOR THE INTERNET OF SERVICES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0736.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Services operation and management, Global service organisations, IT governance, IT/IS management</td>
</tr>
</tbody>
</table>
TOWARDS A SERVICE GOVERNANCE FRAMEWORK FOR THE INTERNET OF SERVICES

Janiesch, Christian, SAP Research CEC Brisbane, SAP Australia Pty Ltd, Building A4, Level 7, 52 Merivale Street, South Brisbane QLD 4101, Australia, c.janiesch@sap.com

Niemann, Michael, KOM – Multimedia Communications Lab, Technische Universität Darmstadt, Merckstr. 25, 64283 Darmstadt, Germany, niemann@kom.tu-darmstadt.de

Repp, Nicolas, KOM – Multimedia Communications Lab, Technische Universität Darmstadt, Merckstr. 25, 64283 Darmstadt, Germany, repp@kom.tu-darmstadt.de

Abstract

The paradigm of the Internet of Services envisions trade on a global service-enabled internet. Companies, which participate in this new world of services, face the challenges of changing market conditions, new competitive threats, and new legal regulations. Service-oriented Architectures (SOA) provide a promising way to address some of these challenges at the level of the company’s IT infrastructure. In order to guideline an enterprise’s organization and IT and ensure smooth operations, governance frameworks have been established. More specifically, IT Governance and recently SOA Governance have been introduced. The basic structure of IT Governance frameworks is applicable to an SOA. However, they lack functionality or applicability concerning SOA-specific challenges. Current approaches, which focus on mere SOA Governance, lack framework scope and are mostly driven by individual companies. This issue aggravates taking into account the shift to an Internet of Services. We identify key issues and provide initial insights on building blocks for a Service Governance Framework which enables operations for companies in a moderated service network. We discuss service life cycle phases, stakeholder roles, and management processes taking into consideration existing frameworks such as ITIL and COBIT as well as industry-specific approaches from companies such as SAP, Oracle, and HP.

Keywords: Service Oriented Architecture, Framework, Service Governance, Governance Management Processes, Governance Roles, Internet of Services.
1 INTRODUCTION

The Internet of Services is thought to enable agile enterprises to reach out to a global market and focus on core competencies but also create global competition. It is extending today's internet to become service-enabled, i.e. facilitating the trade as well as execution of services. For businesses it is supposed to be the underlying global infrastructure which allows forming flexible and agile service networks to provide value-added services (Heuser, 2007). It is “a multitude of connected IT services, which are offered, bought, sold, used, repurposed, and composed by a worldwide network of service providers, consumers, aggregators, and brokers resulting in a new way of offering, using, and organizing IT supported functionality” (Villasante, 2009). Trends like this make companies face constantly changing market conditions, new competitive threats, and new legal regulations which have an impact on their IT (Barros & Dumas, 2006). The Service-oriented Architecture (SOA) paradigm provides a promising way to address these challenges at the level of the company’s IT infrastructure.

Diligent governance has been recognized in recent years as a major requirement for successful adaptation and operation IT, especially for large systems. Governance in general, be it political governance, Corporate or IT Governance, provides guidance for the definition of expectations and responsibilities as well as directions to assess the performance of organizations or projects. Governance elaborates guidelines and rules that need to be adopted and realized by the affected management processes and stakeholders. Service Governance must provide means to effectively exploit the capabilities of SOA in an Internet of Services.

For IT Governance, a number of existing frameworks provide structures, action scope, guidelines, and best practices. However, while the basic structure of IT Governance frameworks exceeds the needs of an SOA, they lack applicability concerning SOA-specific challenges, e.g. cross-company service deployment and third party service management. Current approaches, which centre on mere SOA Governance, lack framework scope and are mostly driven by individual companies. Hence, in order to meet governance requirements, existing frameworks need to be extended and/or refocused.

We analyze existing approaches with respect to their applicability to SOA scenarios and the Internet of Services in particular. In a first step, we blend both, COBIT and ITIL, to get a basic set of governance processes, tasks, as well as a related set of key performance indicators and controls based on mature frameworks. Then, we discuss the major elements of a Service Governance approach. We provide the theoretical fundamentals and initial keystones of a Service Governance Framework covering the service life cycle, involved roles, and management processes.

This research belongs to the design science paradigm (Hevner et al., 2004, March & Smith, 1995). It strives for developing a practically relevant artifact in form of a method framework for Service Governance in the Internet of Services. According to March and Smith (1995) and Hevner et al. (2004), (IT) artifacts are of four types: constructs, models, methods, and instantiations. Constructs are the vocabulary of a domain, a specialized language, and shared knowledge of a discipline or subdiscipline. Models are a set of propositions or statements expressing relationships among constructs. Methods are goal directed plans for manipulating constructs so that the solution statement model is realized. Instantiations (also implementations) operationalize constructs, models, and methods resulting in specific products.

The paper is organized as follows: In the subsequent section we elaborate on the fundamentals of an SOA as well as IT and SOA Governance. In Section 3, we outline necessary requirements for Service Governance by reviewing industry practice and provide building blocks for a Service Governance Framework. They include a frame for the service life cycle phase, stakeholder roles, and a management process cycle. The paper closes with a summarizing analysis of Service Governance for an Internet of Services realized through service marketplaces.
2 RELATED WORK

2.1 Service and Service-Oriented Architecture (SOA)

The SOA paradigm is a holistic approach towards the execution of business processes consisting of services within or across enterprise architectures. Structured adaptation is crucial to the success of a company’s SOA. As an architectural paradigm, SOA defines a number of mechanisms, principles, and conditions: All functions (e.g. business functions) are defined as services. Services can be regarded as “not storable and intangible goods which are constructed in cooperation with an external factor (usually the service consumer). Construction and consumption traditionally occur at the same time (uno-actu principle). Electronic services differ insofar as they are storable in a sense and their consumption, i.e. execution, does not necessarily involve concurrence” (Janiesch et al., 2008a). Services are designed to support reusability in different scenarios. In order to reuse, only a new parameterization is necessary without any change in its implementation. Service functionalities can be automatically discovered via service brokers or registries. They are centrally registered at a database which provides information about the services upon request. Services are self-describing. While interacting, services are loosely coupled. This means a mutual association via messages; dependencies are minimized to mere awareness. It facilitates a number of operations, e.g. their replacement by other services during runtime. Service operations always involve several parties. Services therefore adhere to a communications contract, a Service Level Agreement (SLA), defined by one or more service descriptions and related documents in order to regulate and control service execution. Services are autonomous concerning their logic. They are independent of other services, e.g. software modules, as well as resources such as databases. The realization of services follows the information hiding principle. Services are stateless. They minimize stored information regarding activities, i.e. state or context information concerning either execution or communication with other services is not saved. Services are completely independent of any platform, programming language or operating system. Their technical realization is transparent for service requestors and brokers. Services can be accessed by an invokable interface without any knowledge of its location. An important characteristic of services is their combinability. Services are designed to be assembled to form composite services, each consisting of several single or further composite services (Erl, 2005, Huhns & Singh, 2005).

Accordingly, an SOA is “an application architecture within which all functions are defined as independent services with well-defined invokable interfaces which can be called in defined sequences to form business processes” (Channabasavaiah et al., 2003).

In order to support and facilitate coordination and cooperation between service providers and service consumers, service brokers are established. These are registries known to all eligible providers and consumers. Providers register services by providing meta information such as name, functionalities, interfaces, etc. Consumers query a broker for a service needed – and if a service is found, the consumer and provider exchange interface specifications. If both sides agree to cooperate, a cooperation is established and concluded by an SLA (Huhns & Singh, 2005). Instead of merely being a broker, the platform host can evolve into a moderator who also supervises the delivery of services. In order to realize this vision and achieve this goal, a diligent governance concept is needed from the very beginning, to support the realization of these benefits.

2.2 IT and SOA Governance

The SOA paradigm offers advantages compared to common enterprise architectures such as an increased response rate to changing conditions or interoperability. However, these advantages entail new challenges such as the need for permanent monitoring and control of services – service guidance is required. By following specific guidelines in a top-down approach, an SOA is adopted, operated,
and continuously monitored and checked for adherence to regulations. Governance also ensures compliance, i.e., compliance to intra-company, normative, or legal standards (e.g., the Sarbanes Oxley Act, Basel II etc.).

For IT Governance, numerous frameworks have been specified, e.g., COBIT, ITIL, ValIT, ISO 20000, ISO 17799, etc. Basically, each of them focuses on a different aspect of a company’s IT. While the IT Infrastructure Library (ITIL), e.g., mainly deals with management and support process definitions (Office of Governance Commerce, 2007), the ISO 17799 standard primarily targets security management (International Organization for Standardization, 2006). When these approaches are compared, we see that they do not exclude but rather complement each other. In comparison, COBIT is a high-level governance and control framework, more tightly aligned with the business objectives of the organization than with operational issues (IT Governance Institute, 2007). As a matter of fact, all other frameworks class into COBIT. It has, so far, become a de facto standard for IT control globally, and its implementation increasingly gains interest.

Concerning SOA Governance, many software companies introduced their own definitions in whitepapers, often on behalf of own market interests (cf. Section 3.1). This resulted in the definition of a number of different approaches. Similarly, a number of definitions for SOA Governance can be found: “SOA Governance is a management structure including creational and administrative elements” (Fabini, 2007) or “SOA Governance is a set of solutions, policies and practices which enable companies to implement and manage an enterprise SOA” (Brauer & Kline, 2005).

It is difficult to give a common definition for SOA Governance as a number of different descriptions and understandings exist. The main objective is to define and introduce company-wide policies for the adoption and operation of an SOA as well as to introduce mechanisms controlling their enforcement (Fabini, 2007, Windley, 2006). According to several authors, the basic difference between SOA and common former enterprise architectures is the service life cycle (Bea Systems, 2006, Brauer & Kline, 2005, Software AG, 2005). There is a paradigm shift concerning software development (Johannsen & Goeken, 2007). Hence, a central issue of Service Governance is considered to be the service life cycle governance.

In most cases, an SOA requires the restructuring or adaptation of the company’s organizational structure. A SOA Center of Excellence as well as accompanying boards, consisting of management members and members from the different departments is recommended (Bieberstein et al., 2005b). In case of big companies, one central governance structure seems problematic – decentralized and coordinating as well as hierarchically structured governance positions are combined (Fabini, 2007). The deployment of an SOA Maturity Model is also recommended (Afshar, 2007, Johannsen & Goeken, 2007, Progress Software Corporation, 2006), measuring the maturation of a company’s SOA. Results of such an assessment provide information concerning the progress and success of the realization of SOA. This way they have impact on the governance policies – a control cycle is created.

Most of the authors agree that SOA Governance is a fundamental requirement for a trouble-free adaptation as well as for successful operation. Regulations and control are the central elements which are to be effectively implemented. In parts, SOA makes the same demands on governance as common systems, but to some extent it exceeds the regulative support IT governance can provide. Regarding value contribution or IT-business alignment, existing IT Governance frameworks (e.g., COBIT) provide sufficient support. However, if confronted, e.g., with cross-organizational aspects, the scope of IT Governance frameworks is exceeded. SOA Governance needs to provide the abilities to guarantee sufficient SOA adaptability and integrity as well as to check services concerning capability, security aspects, and strategic business alignment.
3 CURRENT ISSUES AND PROPOSALS FOR SERVICE GOVERNANCE

3.1 Existing SOA Governance Approaches in Industry

An inherent characteristic of SOA is that services are not bound to existent entities, such as, e.g., accounting applications. They can be provided, bought or sold and executed in third party applications or environments. Additionally, to a greater extent than previous enterprise architecture concepts, SOA facilitates inter-organizational deployment of software artifacts. Thus, a governance approach must focus on the adoption and operation of SOA as enterprise architecture in a company. It must provide guidelines and mechanisms to ensure the integrity of an SOA and its adaptability to business and general administration processes. It also must provide tools to support the monitoring and control of services concerning security issues and the alignment to business processes. The main goal is to achieve adherence of the SOA system to various specifications and standards, such as the Sarbanes Oxley Act, ISO norms or internal regulations.

Numerous perspectives of SOA Governance exist driven by individual companies. Hence, they are rather product-oriented. Thus, approaches to SOA Governance do not always comprise all of the above mentioned elements. It depends on requirements and existing structures of the particular cases, which of the elements are considered useful. But they all concur that the adequate implementation of SOA Governance in a company requires an extra approach as extension to IT Governance to address SOA challenges (Woolf, 2007).

The SOA Governance approach of SAP AG, e.g. consists of a guidelines framework and an organizational institution, the Process Integration Content (PIC) Council. The framework has three parts: modeling and implementation guidelines, a special review process performed by the PIC council (guidelines enforcement), and the continuous execution of manual and automated service tests (SAP AG, 2007). The process is model-driven with the support of an integrated modeling suite. The actual engineering and development process starts with the identification and structuring of functional requirements. They are scrutinized in a map which covers deployment units, process components, and business objects. The resulting integration scenario model defines the necessary interactions and results in a service orchestration. The resulting content is subject to a review process by architects and an approval by management. The PIC Council guarantees quality of process integration content by reviewing interfaces for semantic correctness, ensuring standard conformity, encouraging reuse, establishing enterprise-wide consolidation and improving the integration guidelines. The SAP AG approach implements the idea of the SOA Center of Excellence, defining a council as central element of SOA Governance. The design of individual services is also governed by an enterprise service design guide which promotes a business-driven view based on processes and scenarios. Thereby services are not to be designed isolated from each other and are meant to be reused (SAP AG, 2005). The guidelines include concepts of service design for SAP internal development, business analysts, system integrators, and independent software vendors. The procedure is three-stage and consists of indicator-based service discovery, service design and documentation. The methodology also distinguishes the three levels of design of single services, service systems design, and service-enabling of enterprise applications. It is centered on the notion of so-called design contexts which represent patterns for improving applications with services.

Similarly, a comprehensive examination of the scope of SOA Governance is given by Oracle, Inc. (Afshar, 2007). Oracle’s approach attempts to give an overview of the entire Governance domain. Starting from a generic point of view, Oracle Inc. identifies eight decision fields. Within each field, policies for key issues are to be defined, in order to assure the according requirements on SOA and their reliable realization. The approach combines these decision fields with an SOA adoption model and a comprehensive set of best practices. The following table gives a short overview:
<table>
<thead>
<tr>
<th>Decision Field</th>
<th>Key Aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architecture</td>
<td>Standards, architectural assessment mechanisms, reference architectures,</td>
</tr>
<tr>
<td></td>
<td>application guidelines</td>
</tr>
<tr>
<td>Data</td>
<td>Data ownership, data service architecture, formats and standards, formalizing</td>
</tr>
<tr>
<td></td>
<td>of the description of data requirements in SLAs</td>
</tr>
<tr>
<td>Finance</td>
<td>Funding of business and technical services, of the hardware and software</td>
</tr>
<tr>
<td></td>
<td>infrastructure, backbones and assignment to accounts</td>
</tr>
<tr>
<td>Operations</td>
<td>Enforcement of policies and rewards/penalties, capacity planning, operational</td>
</tr>
<tr>
<td></td>
<td>model for cross-department deployment</td>
</tr>
<tr>
<td>People</td>
<td>Incentives for employees, organizational structure, roles and responsibilities,</td>
</tr>
<tr>
<td></td>
<td>SOA training</td>
</tr>
<tr>
<td>Portfolios</td>
<td>Project, service, and legacy portfolios for strategic planning of SOA and</td>
</tr>
<tr>
<td></td>
<td>support for project management</td>
</tr>
<tr>
<td>Project Execution</td>
<td>Project selection and adaptation, competence alignment, formalizing the life</td>
</tr>
<tr>
<td></td>
<td>cycle process control of business processes and policies</td>
</tr>
<tr>
<td>Technology Infrastructure</td>
<td>Strategic SOA platform, governance platform, migration of legacy systems,</td>
</tr>
<tr>
<td></td>
<td>design and implementation of infrastructure services</td>
</tr>
</tbody>
</table>

Table 1. Key Aspects of SOA Governance Decision Fields

Apart from these approaches, many other companies propose proprietary models. We give an overview of the variety of the field in the following:

Brauer and Kline (2005) at HP define different components supporting the implementation and management of SOA. They provide a holistic controlling framework, emphasizing the integration of people (organization), processes, and technology.

Bieberstein et al. (2005a) of IBM propose an SOA Governance Model. They identify six governance processes and three steps for launching the SOA Governance Model. The SOA strategy and SOA objectives should be defined in a way that both business and IT units have a clear understanding of them. According to them, policies, defined by governance positions, form the basis for any decision. Their model is completed by a set of best practices. Bieberstein et al. (2005b) also describe an approach to guide an SOA successfully, emphasizing transformation of organizational structures and behavioral practices. They propose the Human Services Bus (HSB) as a new organizational institution, streamlining cross-department processes, thus optimally exploiting the SOA approach.

webMethods' (2006) SOA Governance approach consists of two parts: Architecture Governance and Service Life Cycle Governance, the latter is divided into design-time, run-time, and change-time governance. Architecture Governance comprises issues such as corporate technology standards, the definition of an SOA topology and determination of an SOA platform strategy. Service Life Cycle Governance focuses on the regulation of design etc. of services through according policies and enforcement mechanisms.

The approach by Software AG (2005) identifies maturity and governance levels. Besides this 6-level-maturity model they define an SOA service life cycle, incorporating services, related artifacts, and roles. They provide a 5-step SOA adaptation plan as well as a set of best practices.

Bea Systems, Inc. (2006) clearly emphasizes the importance of the service life cycle as the most critical requirement of a successful holistic SOA Governance approach. Central policy definition and enforcement, regulating the design, building, provisioning, and operation of services, affect the whole SOA referring to quality assurance, monitoring, and SLA management. The primary goals are reduced development costs and faster time-to-service. With the acquisition of Bea Systems, Inc. by Oracle, Inc. it is reasonable to assume that their approaches will merge.

This list of governance model proposals shows the diversity of approaches to SOA Governance. However, most of them show congruencies which can be generalized. For governance within the Internet of Services, an approach is needed that lies between the requirements of an SOA and the more
general governance of IT. Thus, Service Governance, as we understand it, is a form of IT Governance and is mainly driven by the Corporate Governance of the host. It subsumes several points from the current best practices of SOA Governance, e.g. comprises it considers cross-company issues of multiple parties. Also, SOA as enterprise architecture is addressed. However, in addition Service Governance comprises the consideration of cross-company legal aspects exceeding those of current SOA Governance approaches, i.e. contract management over country borders, country-specific laws for data transmission and protection, and laws concerning the fulfillment of online contracts. Being a cross-company approach the framework needs to consider the interests of all stakeholders of the platform. Both the interests of the host and customers have to be included. Preconditions to be fulfilled by suppliers also have to be formulated, and vice versa. In contrast, SOA and IT Governance approaches normally focus on the operation within a single organization, considering a single stakeholder.

3.2 Building Blocks of a Common Service Governance Framework: Service Life Cycle

As a balanced starting point, the IT Governance frameworks of CObIT and ITIL v3 can be used. This is not only because they provide insights from relatively unbiased organizations rather than individual enterprises, but because they are at both ends of the governance spectrum: strategic governance and IT management. CObIT focuses on strategically important tasks (i.e. main processes) and ITIL focuses on management tasks (i.e. support processes), which are often subject to outsourcing and, thus, the ideal blueprint for managed third-party processes (IT Governance Institute, 2007, Office of Governance Commerce, 2007). Figure 1 depicts existing processes which have been taken over as-is or in an abridged form to focus on the specific needs of a Service Governance Framework. Most of the time however, the governance processes need to be extended to cater for the specific needs in an Internet of Services.
All relevant governance processes from these frameworks can be grouped in a life cycle consisting of five phases: design, deployment, delivery, monitoring, and change (Janiesch et al., 2008b). In each of these phases, several processes constitute the Service Governance Framework. The design phase contains all sorts of strategic aspects of the use or operating of such a platform and traded services. Identifying requirements, development of services, as well as the selection of third-party services are components of the deployment phase. The delivery phase comprises all aspects of service and infrastructure operations. It is closely coupled with the monitoring phase as they are executed concurrently. The monitor phase contains all aspects of service and infrastructure monitoring. The change phase includes all processes and tasks needed to adjust and change the infrastructure and services traded.

Functionality within the Internet of Services is centered on the central service broker component. It cannot be found in current general IT Governance frameworks or in SOA Governance frameworks. Thus, this process is new. There is a need for future refinements of the framework as this first version is intended to show the scope of the Service Governance. While this scope will also be refined, the main focus for detailed development needs focus on the process of broker operations. This is of particular importance if – as mentioned above – services are not only brokered like tradable goods but also moderated in a value-added manner similar to product-service bundles (Schroeder, 2008).

3.3 Building Blocks of a Common Service Governance Framework: Management Processes

Services in an SOA are tightly linked to business processes. These are controlled, monitored, and improved by management and its processes. Generally, a management process is the process of planning and controlling the performance or execution of any type of activity, e.g. projects as well as business processes or workflows. It is a tool for managers to control existing business processes actively – sometimes also referred to as factual leadership. It can relate to the top management of an organization as well as to project management and risk management.

Commonly, the management process can consist of the following phases, mostly described as a circle (Burghardt, 2000) as depicted in the following Figure.

![Management and Business Processes](image)

**Figure 2. Management and Business Processes.**

The objective phase describes a desirable realistic state, defining the aim to be achieved. The planning phase identifies possible ways to achieve the goal. The realization phase triggers, e.g. organization, human resources management etc. During the control phase the degree of target achievement is measured. During each phase, communication and exchange of information between the involved parties is crucial. During the planning and realization phase the subordinate business processes are designed and implemented. During the control phase they are monitored and assessed and the need for change management is evaluated. This can lead to starting over again with the objective phase.

Concerning IT management, there are a number of frameworks that support process management, as e.g. COBIT or ITIL. COBIT provides supporting mechanisms on leadership or governance level, while ITIL targets management of software production in general, i.e. on business process level (Kamleiter...
& Langer, 2006). In case of SOA, a governance framework is of particular importance, as in an SOA the link between IT and business processes is closer than in previous enterprise architecture approaches.

3.4 Building Blocks of a Common Service Governance Framework: Roles

Weill and Ross (2004) stress the human involvement in IT Governance in their definition of IT Governance as they suggest to specify “the decision rights and accountability framework to encourage desirable behavior in the use of IT”. Consequently, as a further step starting from the guidelines for organization forms and responsibilities, a more concrete shaping of the necessary roles is given in this section.

In the Internet of Services, several main stakeholder roles have been identified: service provider, service brokers or platform hosts, and service consumers (Barros & Dumas, 2006). The following Figure gives an overview of the stakeholder roles. While the service consumer and the service provider are actual persons taking the specific stakeholder role, the service broker is a virtual entity, e.g. a marketplace, a piece of software. Nonetheless it is operated by actual persons who act as a platform host.

![Figure 2. General Role in the Internet of Services.](image)

The service provider stakeholder supports agencies that hold governance and operational responsibility for a service, including organizational structures and other business aspects, as well as systems and other implementation artifacts. The service provider represents the role of a development party, producing and publishing services ready for execution. Largely, they are the service owners, responsible for the service implementation as well as maintenance.

The service consumer finds services via the service broker and requests and invokes them. He is the customer in the market transaction. For aggregated services, providers can act as consumers to create value-added services.

The main role of a service broker is to provide service location and description information contained in a service registry. So far the broker role is mainly associated with maintaining registries. As the central information database, its actuality is crucial to the success of the whole SOA system. However, intermediaries can play additional roles, e.g. mutually providing themselves brokering services, load balancing functionalities or negotiation support services (Erl, 2005). With an increasing number of services, registries become more and more important. They serve as a central location for tracking and managing services. The reusability of services depends on registries as these provide a way to share services across organizational borders. As a moderating entity the scope of the host is extended since it also has to attend to run-time and change-time issues such as services, which is updated while being in use. They must not be interrupted during execution.

Gu and Lago (2007) also give an overview of typical service life cycle models, including roles and responsibilities, developed mainly by software companies. Bieberstein et al. (2005a) defines various organizational SOA specific roles which are needed within one or more of the stakeholder roles.
4 CHALLENGES OF A GOVERNANCE APPROACH FOR A SERVICE MARKETPLACE PLATFORM

The scope of this paper has been the examination of Service Governance starting from two perspectives – SOA and IT Governance. For IT Governance, there are existing frameworks that provide structures, action scope, guidelines, and best practices. As this research concluded, the basic structure of IT Governance frameworks is applicable to SOA. However, they lack functionality or applicability concerning SOA-specific challenges, i.e. cross-company service deployment. Hence, unchanged, these frameworks are not fully suitable. SOA Governance requires at least an enhancement. In fact, approaches that focus on mere SOA Governance, lacking framework scope, already exist. An overview has been presented above. Most of them represent generic approaches and are hence applicable for adoption.

The organizational aspect, often identified as one major decision field of SOA Governance, is an important issue. We investigated in detail the service life cycle, roles and responsibilities, and management processes. Common SOA-specific organizational as well as stakeholder roles were presented. Our focus on a specific form of Governance for Services in the open Internet of Services is future facing. Papazoglou and Georgakopoulos (2003) elaborate on service marketplaces as an existing occurrence of SOA and Janiesch et al. (2008c) present an infrastructure and web-based business model for a generic service marketplace. They argue that services will have the largest share in the future business value networks. Thus, services have to be transformed into tradable goods; service marketplaces are considered to be an adequate vehicle to do so.

This entails new challenges such as monitoring and billing techniques and cross-company legal issues which require an improved form of service (marketplace) governance. It shows that the particularities of a moderated SOA cannot easily be addressed and managed by a common IT or SOA Governance approach.

We developed an outline of a governance framework which can be used for a service marketplace incorporating the above described role concepts. While investigating existing frameworks for IT Governance, we learned that an SOA introduces challenges for traditional governance frameworks. In fact, in the case of an SOA marketplace approach, it is to assume that the regulatory demands exceed existing structures in governance frameworks. This Service Marketplace Governance is a form of IT Governance and is mainly driven by the agenda of the marketplace host or service broker/moderator. It is also an extension of SOA Governance and considers the conformance and regulatory needs of a service marketplace host. In four main points, it can be distinguished from the common SOA Governance approach:

It comprises a form of SOA Governance including an according policy framework. (It is considered a super class of the common SOA governance approach.)

It pays special attention to cross-company legal aspects, e.g. data protection/security. Additionally, the term Service (Marketplace) Governance comprises the consideration of contract management over country borders, country-specific laws for data transmission and protection, and laws concerning the fulfillment of online contracts (such as the Fernabsatzgesetz (Distance Selling Act) in Germany).

It covers different service monitoring aspects. Technically, there are a variety of possibilities to realize service or SLA monitoring: decentralized, centralized, or hybrid monitoring. However, centralized monitoring is considered inappropriate due to the large number of service providers and executors and the very large number of services being offered on the platform. Hence, a decentralized monitoring approach might be the better solution. This, however, comes with additional requirements for a Service Governance Framework.

It includes the interests of multiple parties, i.e. stakeholders. Operating a service marketplace platform involves much more stakeholders than common SOA approaches. SOA platforms incorporate at most
two parties: the platform host (which is the service provider, broker, moderator, and developer) and
the service consumer. Being a cross-company approach this framework considers the interests of all
stakeholders of the marketplace platform: Consumers, platform hosts or service brokers, and service
providers. Both the interests of the platform host (broker/moderator) and service consumer are
included. Preconditions to be fulfilled by the service provider are also formulated, and vice versa: A
marketplace governance approach defines policies with respect to service consumers regarding the
interests of platform hosts and service providers. In contrast, SOA and IT Governance approaches
normally focus on the operation within a single organization, considering a single stakeholder.

One way to address these challenges is to make extended use of certification of platform conform
behavior of service providers, executors, etc. A generally accepted way for service providers to show
their competence and compliance with certain standards is the certification of the service provider and
its organization, e.g. based on the well established ISO 9001 audit and certification for quality
management purposes.

The above assumes a governance approach for one (centralized) marketplace platform which acts as a
broker (or moderator) for service providers. This governance approach does not yet cover governance
of multi-broker architectures which have different brokers communicating and trading obeying certain
rules. In fact, however, it addresses complex requirements for controlling and directing an SOA
service marketplace and all of its services and stakeholders, common existing IT Governance
frameworks fail to cover.

Acknowledgement

The project was funded by means of the German Federal Ministry of Economy and Technology under
the promotional reference “01MQ07012”. The authors take the responsibility for the contents.

References

from http://www.oracle.com/technologies/soa/docs/oracle-soa-governance-best-practices.pdf on 2009-
03-16.


Companies Reap the Full Benefits of SOA. BEA White Paper. Downloaded from http://www.itworld
 canada.com/Admin/Pages/Assets/DisplayAsset.aspx?id=e0a24263-a10a-4887-8d45-582261587176 on
2009-03-16.

(SOA) Compass: Business Value, Planning, and Enterprise Roadmap. IBM Press, Upper Saddle River,
NJ.

Enterprise Systems, Organizational Structures, and Individuals. IBM Systems Journal, 44 (4), pp. 691-
708.

Whitepaper. Downloaded from http://www.managementsoftware.hp.com/products/soa/swp/
soa_sw_p_governance.pdf on 2009-03-16.


Englewood Cliffs, NJ.

Gesundheitswesen. In SOA-Expertenwissen: Methoden, Konzepte und Praxis serviceorientierter
Architekturen (Starke, G. and Tilkov, S. Eds.), pp. 309-323, dpunkt, Heidelberg.


IT Governance Institute (2007). COBIT 4.1: Control Objectives for Information and Related Technology. IT Governance Institute, Rolling Meadows, IL.


Teaching Case: Leading the Change - ERP Implementation at Keda

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0251.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Teaching Case</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Enterprise resource planning (ERP) (packaged systems), Implementation, Critical Success Factors (CSFs), Teaching Case</td>
</tr>
</tbody>
</table>
TEACHING CASE

LEADING THE CHANGE - ERP IMPLEMENTATION AT KEDA

Abstract

While Enterprise Resource Planning (ERP) system demand in China has been steadily increasing, many of such implementation projects failed. Keda Industrial Co. Ltd (600986: Shanghai Stock Exchange), a manufacturer of large scale machinery in China, however, was one of the few that successfully deployed its ERP solution in 2005. In this case study, we document the ERP initiative of Keda from its conception to its deployment, study the factors that contributed to its success, and summarize Keda’s practices that substantiated the identified success factors.

Keywords: Enterprise Resource Planning
1 INTRODUCTION

According to a market study conducted by Gartner Research, ERP software demand in China increased by 29% in 2006. In fact, many enterprises in China have recognized the strategic importance of ERP in their rapid developments. Keda Industrial Co. Ltd. (600986: Shanghai Stock Exchange), a manufacturer of large scale machinery in China, was such a company. Keda produces and sells machinery for processing ceramics that are mainly used in construction. In August 2005, a mere five months after forming an implementation team, Keda successfully deployed SAP as their ERP solution. The project quickly paid off through lowered inventory and reduced operation and management costs. This success was all the more remarkable given the fact that an estimated 80% of ERP implementation efforts in China failed. Dr. Fan Zhu, the Vice General Manager of Keda, while utterly satisfied with the project outcome, could not help but wonder if the project success was the result of good practices that he could model future implementation projects after, or, just pure luck.

2 COMPANY OVERVIEW

Founded in 1992 by Lu Qin, with an initial capital of only 90 thousands Chinese Yuan (CNY), Keda began as a small manufacturer of ceramics machinery in Shunde of the Guangdong province. At that time, the ceramics machinery industry was dominated by European companies. By modeling its business after these market leaders, Keda enjoyed rapid growth in the local Chinese market in the mid-90’s and emerged as a major player in the industry, being recognized as one of the top 500 national machinery manufacturers in China, top 10 building materials machinery enterprises in China, among other acknowledgments and achievements. By early 2000’s, Keda had surpassed many competitors to become a world leader in building materials machinery, second only to SACMI of Italy (www.sacmi.com). Keda became a listed company on the Shanghai Stock Exchange in 2002, and in 2007, it reported a net profit of more than CNY 93 million, up from CNY 32 million in 2004.

Today, just 16 years since it was founded, Keda has more than 2000 employees and its product offering includes ceramics machinery, energy resource machinery, stone machinery, and building materials processing machinery. Keda also offers complete plant design and technical consulting services to ceramics and stone processing machinery. Due to the nature of its products, orders from Keda’s clientele typically involve high levels of customization and low volumes.

Keda’s business as a whole relies on several key business functions, such as research and development, purchase of raw materials and parts, inventory management, production that comprises of many assembly lines and workshops, logistics, and sales and marketing that connects the firm with its extensive customer base, which comprises mainly of construction materials providers from both the domestic and foreign markets. Each of these business units functioned autonomously, with little integration in between. This gave rise to a freewheeling corporate culture where non-standardized processes were adapted on the fly and issues were resolved in an ad hoc manner. With these, Keda continues its "pursuit of perfection" through "endless innovation" in the global market.

Innovation, in fact, has been Keda's emphasis. In terms of product innovation, Keda rolled out the first 3200 tons pressing machine in China in 1999. In 2005, three of the ten new machinery products were even considered world's first. Perhaps more remarkable was Keda's commitment to innovation for not just its products, but its business as a whole. In 2003, the company set up the national enterprise post-doctoral workstation, where post-doctoral scholars were invited to work. Their research areas included the business's supply chain management, human resource management, and so on. Keda also invested more than CNY 45 millions in setting up the ceramic engineering test center that was committed to research and development in state-of-the-art technologies. Through innovation, Keda established itself not just as a leader in market share and revenue, but also in product technology and management efficiency and effectiveness.
3 DRIVERS FOR ERP

Things, however, were not all good. The silo-based model was taking a toll on Keda’s business performance. Different business units often had duplicated tasks and data, resulting in redundancy, and in turn, heightened costs. Also, as there was little integration between departments, managers could not make timely and well-informed business decisions. This was especially taxing in the face of competitions from local and foreign companies, who challenged Keda in many fronts. To retain its leadership position and to continue growing its business, Keda needed to stay innovative with product development, business management and operation, and to be more informed about production, sales, and most importantly the customers when making decisions, or as Zhu put it,

“In managing the enterprise, our most important task is to provide the needed information for every decision maker every step of the way in the decision making process.”

An added external factor that also prompted for Keda’s ERP undertaking was pressure from Chinese government agencies. Recent years had seen China encouraging, and at times, pressuring for innovation in local enterprises through campaigns with explicit incentives to promote computerization in corporations. Keda’s ERP undertaking could be seen as a reaction to the government’s call.

A common challenge that enterprises like Keda face has to do with raw materials and parts. As mentioned above, Keda’s production is characterized by low volumes and high customization, which consequently lead to frequent and scattered demands for a large number, but not necessarily large quantity, of raw materials and parts. Some of these parts are common across different products, but most cannot be reused in different products. This makes purchasing decisions, or simply keeping track of the quantity purchased and used, difficult. This also has other repercussions, as Zhu explained,

“Before ERP deployment, due to the mess in material management, the cost of a product was unclear, and costing was based on experiences. It wasn’t clear how much profit or loss the sale of a product resulted in … and which part of the product or the production process contributed to that profit or loss. That made it difficult for us in pricing our products, too.”

Also observed was suboptimal utilization of resources and facilities. Reusable materials and parts were often scrapped; precious machine time of key facilities was often wasted idling. Zhu illustrated,

“For example, we had a key facility for production. We had someone time its usage in a week using a stopwatch. We were shocked when we learnt that it was only in production 24.6% of the time!”(Compared with over 90% utilization in a Japanese competitor)

Moreover, Keda’s rapid growth was not without consequences. For instance, competition from local and abroad prompted Keda to diversify its business and product lines. Keda’s commitment to R&D in turn resulted in a greatly expanded product offering. The single plant mode of production could no longer cope with the highly diversified business lines and production functions. This prompted Keda’s shift from single plant to multiple plants in 2004. By expanding its operation to multiple plants, it outgrew its existing Manufacturing Resource Planning (MRP II) system's capabilities. Keda's MRP II system at that time, CAPMS developed by Beijing Riamb Software IT Co. Ltd., did not support such multi-plant operations. Adding to the demise was the fact that Riamb ceased maintenance support for the system due to an internal restructuring of the software company. It seemed, to continue its success, Keda needed to rethink its IT, and it needed to do that quickly.

4 COMPUTERIZATION AT KEDA

Keda’s first take on computerization started in 2000. As a member of the IT department recalled,

“… there was no strategic goal. Others were doing it, so we decided to jump in as well. Since there was really no planning or even particular objectives, the IT projects were not
particularly effective in solving the company's problems, and the use of the systems at that time was less than satisfactory.”

Beijing Riamb’s suspension of support for the MRP system in 2003 did not help either. That was when Dr. Fan Zhu came on board as the head of the IT department. From an academic background, Zhu was familiar with the IT literature, and realized that success in an enterprise system project begins with clear objectives and expectations, and any computerization initiative would need to align constructively with the company’s strategic goals. In fact, one of the principal guidelines for enterprise system undertakings in the literature is to clarify strategic and organizational needs early (Davenport, 1998). Zhu explained the role of computerization in addressing such needs,

"Our goal is to be the world leader. Such goal cannot be attained by simply increasing labor hours. It requires a well established structure and system ... including computerization as a means. Computerization is an auxiliary tool that helps [Keda] to achieve its goals and develop its business."

What Zhu also realized was the fact that to harness the power of this tool, the computerization initiative had to be carefully planned, as a long-term disciplined plan was instrumental in any IT success (Feld & Stoddard, 2004). This multi-year plan would provide a detailed blueprint for systems adoption. With that in mind, he decisively ordered a halt to all ongoing IT projects. All resources and efforts were then put into working out a comprehensive and far-sighted five-year computerization plan. This plan would address the short term needs as well as the long term strategic goals, in order to propel Keda forward. In devising such a plan, Zhu had the help of Benjun Zhang, who has now become the head of the IT department.

"When working out the plan, we focused on how much investment was needed and in what kinds of systems, what our objectives are, what hindrances the company is facing, why we need computerization, and what problems we are targeting to solve." Zhang said.

One can easily imagine the complexities in working out such a plan. An assessment of status quo and a market analysis were first conducted. Existing problems were identified, and based on the business requirements collected from various levels of management, a set of objectives was set out. These objectives were then prioritized according to their urgency and how well they align with the company's strategic goals. Ranking high on Keda’s wish list was an integrated organization structure, breaking departmental boundaries, with streamlined data flows and business processes. Also considered crucial were increased management control, information availability, accuracy, and timeliness, by imposing standardized processes and procedures through systems adoption. Potential solutions for the requirements were considered as well. Throughout the plan, top management provided the high level direction, the long term vision, or as Zhang put it,

"... He (The Board Chairman) has a good vision ... wherever he is in the world, he wishes to be updated with the company's operation using a laptop computer, instead of relying on the reporting from the financial manager."

This planning effort lasted six months. The result was an encompassing plan that included the implementation projects for Enterprise Resource Planning (ERP), Product Data Management (PDM), Office Automation (OA), Manufacturing Execution System (MES), Customer Relationship Management (CRM), and Supply Chain Management (SCM) solutions, in phases. (See Exhibit 1.) This plan was directed towards a unifying platform on which all business applications run, uniting business units and promoting a centralized organization structure; a platform that was so often cited as a key ingredient in enterprise system success (Feld & Stoddard, 2004).

Also in the plan were detailed analysis of each project regarding its objectives, expected investments and benefits, feasibility in terms of factors such as staffing and technology, risks, and alternative solutions. According to the plan, the ERP implementation project was a priority, to be completed by the end of 2005, and had to go first. Time was critical.
5 CHOOSING A VENDOR

ERP strategy choice is crucial to project success but its importance is often underestimated (Holland & Light, 1999). One of the first choices Zhu had to make was between in-house development and third-party provider, and he felt that it was not a difficult choice. From the get-go, it was clear Keda would not be developing the system in-house. To begin with, time was not on Keda's side. Given the tight timeframe, partly due to Riamb ceasing maintenance support for the current MRP II system, building an in-house ERP solution that could easily take years was simply not an option. Moreover, ERP systems had been successfully adopted in many organizations in Keda's line of industry, rendering customizations needed and risk in packaged software minimal. Also, Keda derived its competitive advantages not from proprietary business processes, but from product innovation and price competitiveness. Hence there was little to lose adopting common existing solutions. Furthermore, there was a general lack of internal IT expertise at that time, but information technology was not perceived as a core competency of the company and accordingly, building a large team for such development was not favored. Instead, professional consultants and experiences vendors were to be leveraged, as Zhu explained,

"Enterprises need to be clear about their core businesses and competencies. For instance, we could have developed in-house our own ERP system, but that would not be very smart. Why reinvent the wheels when others have gone the same routes? Outside of our core business and competency, for anything we need, we prefer renting or buying existing solutions."

Keda thus went ahead and contacted twenty software vendors from both local and abroad. These included Kingdee, Lima, Tianxing, and several vendors for SAP, among others. These vendors were invited to visit Keda, and through these visits, they would briefly study the needs of the company, demonstrate how their software packages would satisfy those needs, share their past experiences, and discuss other implementation issues.

"The ERP project kicked off in August 2004. Vendor assessment was conducted. We invited both local and foreign software vendors to visit our company for detailed assessment ...
through these visits, we also aimed to let the middle and top management learn more about computerization and ERP in general.” Zhang recalled.

However, Zhu did not stop there. He decided to take on a more proactive role and, in turn, visited existing clients of these vendors. Zhang explained,

“We visited their existing clients, as we felt that only what we saw must be real. In these referrals, vendors would likely put forth their most successful cases, and through comparing the uses of their respective solutions in real production environments in the industry, we had a much better idea what suited us.”

In these visits, Keda gained insights about not only the vendors, but also possible complications in the implementation process; these insights proved valuable later in the implementation process, as they allowed Keda to avoid the same mistakes committed by others, as Zhu observed,

“After visiting several enterprises, we made an interesting observation: their IT departments were basically the only ones working on the ERP projects. They were churning out reports, workflows, etc, things we thought were outside the work scope of an IT department.”

Based on these encounters, Keda short-listed 9 vendors to respond to a Request for Proposal (RFP). This RFP specified the exact needs of the company and standardized the vendors’ responses thus allowing for comparisons of their products on the same ground, namely, how the particular modules of their packages would fill particular needs of the company. These 9 vendors presented their proposals in a three-day period. Zhu made sure senior management, including the director himself, was involved in this selection and negotiation process. After the presentations by the vendors, they would rate them according to how well the vendors’ solutions satisfy Keda’s needs and align with the management’s strategic goals. However, the reason for their involvement was much deeper, as Zhang explained,

“Through these presentations, we got a feel of what senior management was excited about. If he would not even attend these meetings, then his so-called support for the project remained superficial, and project implementation would be difficult. On the other hand, if they were interested, they would raise questions on the spot, after all, this is his enterprise and his 10-million yuan project.”

Finally in October 2004, SAP emerged as the winner. According to Zhu, SAP was chosen for various reasons. For one, SAP’s ERP solution was considered sophisticated and feature-rich. Specifically, unlike the company’s existing MRP solution CAPMS, SAP ERP was capable of supporting complex operational processes across multiple production plants. For another, being one of the industry leaders, continual and reliable support and maintenance could be expected. Moreover, with more than 35 years’ worth of experiences and more than 40,000 client implementations, SAP had a proven track record of successful ERP implementations. In particular, SAP was adopted by SACMI, a major competitor of Keda and the industry leader.

6 ASSEMBLING THE TEAM

“There were three main roles: the key users, the consultants, and the IT officers. The key users were at the core, consultants acted as coaches, and IT officers provided support, supporting the consultants and key users in their tasks. Success or failure hinged on the key users, since they dictated the future workflow. IT merely provided support; IT officers would not know how the production workflow should be,” said Zhang.

ERP is often regard as more about people than about systems, hence having the right people in the right roles is critical (Davenport, 1998). Zhu understood that, so with the ERP solution vendor on board, Zhu proceeded to painstakingly assemble the rest of the project team. Another external party, Digital China Management Systems Ltd., was recruited in a consulting role. Together with various stakeholders within Keda, such as the senior executives, key persons from all departments involved,
and all members of the IT department, the project team was officially formed in March 2005. In recruiting departmental representatives, Zhu made sure vital figures in the departments, like departmental managers or main operational staff who had thorough understanding of the operations and needs of the departments, were involved. He had an interesting philosophy, which illustrated his emphasis on user involvement as a key ERP critical success factor,

“In assembling the project team, we insisted that the top dogs of the various departments be involved. Also, we tried to pick those who were deemed indispensable by the departments. We relied on how vehemently departments opposed to the particular person's involvement in the ERP project to judge how indispensable he was.”

Zhu made sure every member had a clear idea of his role and responsibilities in the project, as he said,

"I think the definition and assignment of responsibilities were very important issues in the ERP implementation ... we had devoted a great effort on these issues throughout the process."

Exhibit 2: Keda’s ERP Project Team Structure (Source: Keda Industrial Co. Ltd.)

Exhibit 2 illustrates the team composition and assignment of responsibilities for Keda's ERP implementation project. In brief, top management of Keda would provide overall project direction and make critical decisions; overall project management would involve project managers from both Keda and Digital China; consulting support would be provided by personnel from Digital China; departmental representatives would oversee the business process redesign and system design aspects; the IT department would provide all necessary information technology support. In particular, each ERP system module was assigned an owner from the associated department, who was fully responsible for the workflows and operational details of that module. These key users ensured that system designs reflected business practices, and played a crucial role in training subsequent users in their respective departments. As illustrated by Zhang,

“Throughout the ERP implementation, we emphasized heavily on clear definitions of roles and responsibilities. They [Key users] were the core, consultants acted as coaches, and the IT department was in a supporting role. We could assist them technically but they were to determine how the system should function and what values it should create ... Obviously, the critical success factor lied with the key users.”
Besides clearly defining roles and assigning responsibilities, on a more tactical note, Zhu went to great length to ensure that the project team worked as an integral unit, as he believed that a determined team effort was not only beneficial but essential for project success. As Zhang recalled,

"... the arrangement was for the head of each module to work together at the IT department, with the marketing manager representing the marketing department, purchasing manager representing purchase department, etc ... at one point, their computers were moved to our office and we worked together as the ERP implementation was a crucial matter related to various business units."

This collocation scheme, which took place from March 8, 2005 to August 1, 2005, and from September 9, 2005 to January 4, 2006, is a further demonstration of the importance of user involvement and participation that is so commonly cited as a critical success factor in the literature (Holland & Light, 1999; Esteves & Pastor, 2000). The ERP effort was deemed important enough for all ERP team members to be physically working together. As a result, the departmental managers would only have time to work on departmental tasks after 5:00pm. Not surprisingly, this scheme was met with great resistance.

"There was a lot of resistance to this arrangement. For example, they said if they were to be working with the project team and be away from their offices for 5 months, and their departments continued to function normally, then they should probably all quit since that proved they were of no value to their departments," recalled Zhang.

Fortunately, management’s support was on Zhu’s side. After reporting to the director about the opposition, he worked with the director to devise an arrangement for reward and punishment in resolution of the issue. Since departmental managers had to work longer hours to keep their department afloat, they would be compensated with bonuses. On the other hand, if one were to be unsupportive of project work and miss project meetings, he would be assigned to serve at the canteen during lunch hours for a week!

"This arrangement was greatly effective as it somewhat frightened the staff, as no one wanted to be embarrassed. After all, it was the decision of senior management and if anyone really got punished this way, to serve at the cafeteria for a week, his career in Keda would be all but over," said Zhang.

7 IMPLEMENTING SAP

With these, the ERP implementation project in Keda went underway. Numerous changes were involved. The operations and workflow in many departments were modified and streamlined, and the organization structure also had to be adjusted, in order to accommodate the ERP system. In particular, as-is processes were to be studied and to-be processes were to be designed, and this was a responsibility given to the key users. According to Zhu, they would write the specific requirements for the system, specifying the inputs and outputs at various points for the particular modules they were responsible for, and the IT specialists would implement those requirements. Also falling on the shoulders of key users was data model specification, as explained by Zhang,

"Data was the main issue. The sheer amount of data was one issue, and the data model was another. We had to devise a data model to support all the operations and workflow, in order to minimize the disturbances after rollout. Since experiences in the operations were needed in this specification, we assigned this responsibility to the key users. Only they are most familiar with the operations and hence they should set the rules."

Much hard work was put into effectuating these changes, for instance, as Tian, the person-in-charge of the materials management (MM) module, recalled,
“Data preparation was the most difficult task. We spent a lot of time on stock taking and renumbering stock code so as to ensure data accuracy in the new system.”

With changes came disagreements, between different module owners, and between module owners and consultants. According to Tian, in resolving these differences, discussions among departmental representatives would first be held in seeking a compromise, and if no such compromise could be reached, the matters would escalate to the project management staff who would make the final call, balancing the way users wanted to work and the way the system was designed to work. Zhang added,

“The critical issue was balance; sometimes consultants would give certain advice but key users would advocate alternative proposals. Consensus could not be reached every time ... in those cases, project managers needed to make the decisions. In some cases, the IT department would have to submit a report analyzing the pros and cons of the proposals made by both parties in order to seek a decision from senior management ... As SAP ERP was a well-established system, in practice, in the absence of a consensus, we adopted the consultants' recommendations almost 100% of the time, since we were more inclined to adopt the model the system was originally designed for, and changes in the workflow and its management were to be expected.”

Implementation thus went on relatively smoothly, and speedily. Zhu again stressed,

“In managing changes, I deemed responsibility assignments the most important. There was no way the IT department alone could have achieved so much ... Take BOM [Bill of Materials] for example. We had over 30,000 products, if each was to have merely 3 levels, we would end up with 90,000 levels to specify; there was no way the IT department alone could do that.”

Also instrumental in managing these changes, as Zhu pointed out, was top management support, perhaps the most well-studied and central component of ERP critical success factor frameworks (Davenport, 1998; Holland & Light, 1999; Esteves & Pastor, 2000). Top management support is widely deemed essential for achieving project goals, and is manifested in managers’ involvement in the implementation process and commitment to allocating organization resources. Indeed, top management support was plenty throughout the implementation project. For instance, in 2005, the board chairman himself was present in five company meetings, four of which dealt with the progress and status of the computerization project. He was also present for some of the regular project meetings, blueprint design briefing, and the project launch ceremony. Zhu explained,

“... reactions of persons-in-charge of individual modules depended largely on the attitudes of the senior executives; the heads would not oppose to changes if top management had shown their determination.”

When asked about his success on securing top management support, Zhu offered his advice,

“You cannot let his [the director’s] support remain verbal, but to have it realize in actions. Firstly, he should be involved in the major events, milestones, and decisions making. Secondly, you should report all challenges and difficulties faced in the implementation process promptly, and not wait until the issues have already surfaced, so that you don’t appear passive.”

With the top management committed to the ERP project, the key users committed to the changes associated. In fact, the focus was then shifted to acclimating users to these changes. Besides involving them throughout the implementation process, staff training was also conducted, in parallel with system testing. Zhang explained on the training program,

“The first level of training was provided to selected key users by the consultants, the selected key users then in turn transfer their acquired knowledge and skills regarding system operations to other users. At the same time, selected key users were responsible for preparing operating manuals for all users.”
Moreover, according to Zhang, by conducting system training and system testing in parallel, not only was time saved, but users were also leveraged in identifying system issues. After all, the users of the system would be in the best positions to demonstrate how the system would function, or malfunction, in daily operations. Through testing, users also picked up skills in using the systems, as Zhang commented,

“Testing was training.”

8 GOING ONLINE

By August 2005, a mere five months after the project team was formed, Keda was ready for system deployment. Time for deployment was carefully chosen to mitigate risk and minimize business disruptions, as Zhang explained,

“All industry has its peak season and low season. For us, the months of August, September, and October were the low season and pressure on production would be lower. That was why we chose August as the rollout time.”

At this point, Zhu is faced with another strategic choice, this time a deployment one. While a phased implementation would be less risky, a “Big Bang” implementation offers full functionality right at cutover (Esteves & Pastor, 2000). To expedite the entire process, an aggressive Big Bang approach was adopted. Zhang explained,

“We did a rough assessment and did foresee the possible problems immediately after cutover, such as delay or even suspension of production. However, if we were to do a parallel rollout, it might take a month at least to stabilize, during which all data and work needed to be processed and done twice. Given the huge amount of data and the already heavy workload, parallel rollout was not viable and we went for a Big Bang approach.”

Work did not stop at cutover. Problems started to surface and they quickly translated into production delays. While busy fixing the technical problems and circumventing these delay issues, Zhu once again had the top management’s assurance:

“We notified the director himself before rollout about these possible delays in production, as the ERP rollout represented a fundamental change in how our staff operated. The director was extraordinarily supportive, and said that he could tolerate ERP-affected deliveries in the first year, during which ERP assumed top priority.”

User adoption was another issue. As most prominent ERP systems were originally developed with users in North America and Western Europe in mind, it was suggested that cultural complications might arise in the Chinese setting (Davison, 2002). In particular, opposition to technological changes might be especially pronounced in the Chinese context due to feared loss of status and discretion, new performance metric, and adaptation to something new (Zhang et al., 2003). According to Zhang, with the system online, workflow and operations dramatically changed, they faced a lot of challenges in human resource arrangement. Some workshops were unsupportive of the ERP initiative to say the least, since they did not view it as something the enterprise could not do without. In fact, resistance started to surface once again, as Zhang commented,

“When our ERP system first went online in 2005, some felt that work became more tedious. For example, materials for production now had to be checked in to and out from the warehouse before taking it to the production plant.”

He attributed users’ resistance to a power struggle brought about by the drastically changed business operations. Zhang explained,

“This resistance mainly stemmed from the redistribution of power following ERP rollout … Inventory management was decentralized and each workshop had its own inventory.
Following the ERP rollout, inventory management was centralized so that resources could be shared and better managed. On the third day after rollout, a workshop manager came to us and said that this centralization was wrong and he needed to set up his own inventory separately once again. Certainly, we said no since that would be in conflict with our ERP workflow. On the fourth day, he shut down the workshop, and the senior executives had no choice but to replace him.”

This replacement as a resolution was seen in another incident, as Zhang recalled, "We had new designs and new production orders every day, and these placed a lot of pressure on the purchase department and its manager. After one month, he could not cope anymore and decided to bypass the ERP system. Senior executives promptly replaced him.”

When asked about the rationale behind such responses, Zhang justified, “This is how things work in China; when a person is promoted to a new position, he has no conflict of interest with the new system, and is therefore willing to comply. Then everything else is easy, and can be solved by an increase in resources.”

Forced compliance worked well in this case. This apparent ease in change management stemmed from a country specific culture in the Chinese context, where organizational culture is imposed (Shanks et al., 2000). Authority, in this case the top management, was looked up to for direction. If top management demanded changes, they would be accepted. This demonstrated how country-specific culture could actually be exploited in attaining project success.

Forced compliance was of course not the only way Keda dealt with issues. For instance, after system rollout, some were not completely satisfied with system operations and interfaces, and requests for modifications to the system were reflected to the IT department. However, since the users were the original people who specified most of the system operations, Zhu was able to fend off these requests for the time being until system stabilized. He explained, “After system rollout, they [key users] found problems and demanded changes to the system. However, since they were the one who originally specified the workflow, the inputs, and the outputs, these shortcomings were their responsibilities. Yes, the system needed to be optimized, but that had to wait and they had to accept and adapt … Imagine if the IT department were to be responsible for specifying the requirements and system flows, then key users would demand endless changes, causing delays in the project. However, since they were responsible instead, they had no excuses.”

Within months, it was clear that the ERP implementation effort was a great success. Data accuracy, including that for production plan and inventory information, was estimated to have improved from 85% to 98%. Better data and information transparency and availability were also observed. “Information was much more transparent after system rollout, and tracing this information was much simpler through the system. This was in direct conflict of interest with some middle management. For instance, before ERP, it was very difficult to trace price paid for materials purchased, but after ERP, price analysis could be done conveniently. If price for a material was on a down trend, and your purchase price was getting higher, you have some explaining to do,” said Zhang.

These in turn allowed for improved market responsiveness, decreased level of stock, a significantly better product delivery time, and faster monthly financial reconciliation. The result was a dramatic decrease in operation and management cost, and improved efficiency, as Zhang illustrated, “In early 2005, we were struggling with our machine press production, and production delay was common. At that time, we were producing about 6 machine press per month, and that was under constant pressure from senior management. However, by March 2007, with
exactly the same facilities, we were producing over 30 per month. This dramatic improvement in production capacity is attributed to the information system in ERP.”

Due to improved dissemination of information, quality of decisions made was also improved, Zhang claimed,

“Before, he [the director] believed that orders from foreign markets would bring high profits, but after ERP rollout, through cost analysis, we realized that while a polishing machine sold for USD 50,000 higher, cost was CNY 400,000 higher! These orders did not bring high profits, and management shifted its focus from developing foreign markets to controlling costs.”

9 FUTURE OF ERP

Despite the success of the ERP implementation project, the initiative is part of a bigger computerization plan that would continue to satisfy the company’s information needs and improve the company’s productivity through better information management and dissemination for decision making. To fully capitalize on the centralized data and processes made possible by ERP, Keda remained committed to the computerization plan originally set out. Values were continuously created via disciplined new system additions and extensions. For instance, the Manufacturing Execution System, seamlessly integrated with ERP, tracked all manufacturing activities and provided real-time access to production progress and information.

“In the future, we would still be optimizing and extending on our ERP system, the ten modules and beyond. With the Manufacturing Execution System (MES) expected to roll out and be integrated with the existing ERP this year [2008], we foresee a much better monitored and managed production process. Besides improving resources and facilities utilization, we will also be able to tell exactly which stage of production each order is at, together with all other production details.”

ERP is an enabling technology. With ERP, Keda is strategically positioned to reap the benefits of computerization. ERP serves as the core on which systems, such as Customer Relationship Management, Business Intelligence, and Supply Chain Management, are built to enhance customer management, optimize decision making, and coordinate vendors respectively. To Keda, ERP is a beginning, rather than an end.

Reference

FUTURE REQUIREMENTS OF ERP SOFTWARE FROM THE VENDORS’ POINT OF VIEW

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0287.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Enterprise resource planning (ERP) (packaged systems), Service oriented architecture (SOA), Implementation, IT Project Management</td>
</tr>
</tbody>
</table>
FUTURE REQUIREMENTS OF ERP SOFTWARE FROM THE VENDORS’ POINT OF VIEW

Frick, Norbert, Institute for IS Research, Department of Computer Science, University of Koblenz-Landau, Universitaetsstr. 1, 56070 Koblenz, Germany, norbert.frick@uni-koblenz.de

Schubert, Petra, Centre for Applied Information and Communication Technologies – CAICT, Copenhagen Business School, Howitzvej 60, 2000 Frederiksberg, Denmark, psc.caict@cbs.dk

Abstract

CIOs and CEOs are confronted with the key question of choosing the right ERP software for their company. The evaluation process tends to be painstaking as the outcome usually affects the competitiveness and thus the future of the company itself. Packaged software has become so powerful in recent years that it fulfils the requirements of companies from different industries after a thorough customization process. Nevertheless, case studies and anecdotal evidence show that in many cases ERP implementation projects are demanding and results do not meet expectations. This leads to the question of whether and how ERP vendors are addressing the perceived problems of ERP users. Is there work in progress that will help facilitate the selection and implementation of ERP packaged software for users? Will there be tools available to adopt ERP modules to the specific business processes of a company and – taking into account that processes are likely to change over time – will there be mechanisms available to adapt the software to changing requirements (in the literature discussed as “agility”)? These are the questions that stimulated an in-depth study of the German-speaking ERP software market. In a qualitative study based on thirty interviews with ERP vendors we draw a picture of the value proposition of future ERP software.

Keywords: ERP systems, software development, requirements, IT value, packaged software.

1 INTRODUCTION AND LITERATURE REVIEW

Integrated business software (ERP systems) covers all relevant aspects of an enterprise (Davenport, 1998). “Business information systems can be either designed as custom applications or purchased as off-the-shelf standard solutions” (Scheer & Habermann 2000; Soh and Sia, 2005) (the latter is called “packaged ERP software” in this article.). According to the literature, today most companies turn to packaged software when looking for an enterprise solution (Schubert and Leimstoll 2004). Holland and Light recognized this trend almost 10 years ago when they said “companies are radically changing their information technology strategies by purchasing pre-packaged software instead of developing IT systems in-house” (Holland & Light, 1999).

Therefore many enterprises rely on external partners when it comes to business software. In recent years, researchers in the field of business software have investigated new architectural concepts and technologies for ERP systems including service-oriented architectures (SOA), Web services, XML and modularity. These are keywords which have been circulating for some years in the business software world. According to the experts, modern enterprise systems should be service-oriented, constructed in a modular form, and communicate via Web services. In addition the overwhelming variety of different interchange formats for business documents should be brought to an end by recognized XML standards (McGovern, Sims, Jain and Little, 2006; Yucesan, 2007). The question
which presents itself however, is much more far-ranging. Is a business solution sustainable if it does not fulfill the above mentioned requirements?

There are many systems in active use that are over six years old (Schmitt 2007; Winkelmann, Knackstedt and Vering, 2007) that follow few or none of the desired principles mentioned above. So for ERP vendors and ERP customers alike it is of reasonable interest to know about future challenges and opportunities. In small and medium-sized companies in particular, the introduction of a new system, or the migration of old to new platforms is accompanied by high costs in relation to turnover and, naturally, by associated risks (Scott & Vessey 2002; Winkelmann and Klose, 2008). The results of an SME study show that “the most important criterion used in selecting an information system is the best fit with current business procedures” (van Everdingen et al. 2000, p. 29). If a new system is needed, one should know exactly what it can do and how it will achieve this. A recent study confirms that “(...) companies often face the dilemma of whether to adapt to the software and radically change their business practices or modify the software to suit their specific needs.” (Dalal et al. 2004, p. 84)

Enterprise agility is understood as the ability of a company to adapt to new market requirements with the help of its ERP software and is a topic that has been discussed in academic literature in recent years (Gattiker et al. 2005; Sambamurthy et al. 2003). Concepts such as Service-oriented Architectures (SOA) (Liebhart 2007; Jankowska & Kurbel 2005) and Business Process Modelling (BPM) call loudest for research on flexible systems that can adapt themselves to user needs (Newcomer & Lomow 2005).

In order to be able to draw a picture of the state-of-the-art in ERP research we started our research process with a preliminary query of the EBSCO database using the search word “ERP”, “enterprise resource planning”, and “enterprise systems”. The search resulted into more than 40 articles of which 22 were interesting in the context of our project. Most of the papers stem from three special issues on ERP/enterprise systems (two in EJIS and one in CACM). Among the ones that we did not include were articles on performance measurement/cost-benefit analysis which are important topics for ERP but were not within the scope of our research. It is interesting to note that many authors underline the unvaried importance of enterprise systems for the competitiveness of companies and thus the importance of the topic for IS research. The following list contains an overview of the topics treated in the selected papers:

- ERP implementation: 10
- Process modelling: 3
- ERP system agility: 4
- ERP adoption: 2
- Miscellaneous (cultural aspects of ERP: 1, open source ERP systems: 1, ERP integration: 1)

The majority of articles deal with ERP implementations (Holland & Light 1999; Akkermans & van Helden 2002; Markus et al. 2000; Scott & Vessey 2002; Gosain et al. 2005; Biehl 2007). Two reports are on ERP implementation failures (Iacovou & Dexter 2005; Wei et al. 2005), two papers are specialized on upgrading ERP systems (Beatty & Williams 2006; Khoo & Robey 2007). There is a noticeably large number of case studies used in the articles to illustrate the findings. Process modelling (Scheer & Habermann 2000; Dalal et al. 2004; Delen et al. 2005) and ERP adoption (van Everdingen et al. 2000; Hwang 2005) focus on the need of the company to plan and adapt to the possibilities of the software system. The articles dealing with system agility show the need for flexible systems with regard to future requirements (Smith et al. 2003; van Oosterhout et al. 2006; Overby et al. 2006). It was interesting to see that none of the articles dealt with how or where companies operate their business software. The implicit assumption was that ERP software is operated by the user company itself. With the latest discussion of Application Service Providing (ASP), Software-as-a-Service (SaaS) as well as “on demand software” on our minds, we decided to also study the mode of operation (and thus the possibility for outsourcing) of future ERP systems.

With these considerations in mind the authors launched an investigation on the future requirements of packaged ERP systems. The underlying research question was the following:
What are future packaged ERP systems going to look like?

In order to answer this question, an empirical study of the ERP market in the German-speaking area of Europe was performed. We applied an explorative approach which used an empirical research method (interviews with ERP vendors) (Mayer 2004) in combination with an evaluation approach following Mühlfeld (1981) and Miles and Huberman (1994).

2 BACKGROUND OF THE EMPIRICAL STUDY

Since the study was geared at future requirements of packaged ERP software in the German-speaking market we chose an explorative research approach due to the unknown correlations that might affect our research objective. The study provides a structured overview of the different influential aspects such as technology or architecture and to reveal interdependencies among these components. We chose an empirical research method (Mayer 2004) because we wanted to collect knowledge from industry experts.

2.1 Research method

Mayer (2004, p. 30) suggests a general approach which starts with the development of a model. According to him five sources of information (theories, expert knowledge, common knowledge, literature and similar studies) are necessary to create a theoretical model which describes the research realm (the relevant part of reality). The next step includes the definition of all relevant concepts represented in the model followed by a dimensional analysis which deepens our understanding on a granular level. In the last step, the method of investigation needs to be selected (qualitative approach vs. quantitative approach). Since we intended to follow an explorative approach we selected a qualitative-oriented method. We created an interview guideline based on our theoretical model and evaluated the results.

2.2 Theoretical Model

The research was initialized by a preliminary query of the EBSCO database which gave us a general overview of relevant articles in the European Journal of Information Systems, Communications of the ACM and IEEE Software (mentioned above). In an initial workshop with an established ERP vendor and several other experts, key areas and relevant research questions were systematically identified. Areas of interest in the field of ERP systems such as business process modelling, SOA, Web Services and system architecture in general were discussed during the workshop.

The first version of the emerging theoretical model was based on practical experiences of the software vendor and the ERP experts who took part in the workshop. Architecture, technology and operations were identified as the initial three abstract model components. A consecutive literature analysis revealed that there are many publications which deal with architecture, technology and operations but there are additional studies which focus on market share and market development. We thus added the dimension “Market” to the model. The more detailed development of the model was performed by the authors based on existing ERP literature. In this phase we found a similar ERP study from 2007 with a focus on the user side with the title “Future development of packaged ERP-software” (Felley 2007) which provided us with valuable insights from the opposite perspective –the customer’s viewpoint.

The final result of this preparatory phase was a comprehensive multi-layer model which served as the basis for the development of the interview guideline (cf. Figure 1). The model includes all five sources mentioned by Mayer (2004, p. 30) (theories, expert knowledge, common knowledge, literature and similar studies).

The research questions for the analysis of requirements regarding ERP systems can be classified into aspects which contain conceptual and technical aspects as well as examined possible application
scenarios and market-relevant developments. The framework represents all the characteristics of an ERP system which the authors intended to study. In total, four aspects of an ERP system were defined:

- **Architecture**: This aspect contains the general concept, modelling as well as functional and integrative mechanisms.
- **Technology**: This aspect highlights necessary tools for internal operation and external integration (Volkoff et al. 2005).
- **Operation**: This aspect investigates the place of the software installation, maintenance and licence models.
- **Market**: This aspect includes the study of customers and competitors – the market environment which the commercial software provider faces.

![Figure 1: Theoretical model: the four main components of ERP systems](image)

### 2.3 Definition and Dimensional Analysis

In the next step we defined the relevant concepts of the theoretical model (Architecture, Technology, Operation and Market) followed by a dimensional analysis (Mayer 2004, p. 32) which describes the sub-categories of the main components in more detail (e.g. for Architecture: SOA, monolithic, client-server etc.). Due to limitations in space the sub-categories will not be described further in this paper.

### 2.4 Qualitative Research

In order to identify new patterns and correlations between the aspects of our theoretical model we followed an explorative research approach. Following Mayer (2004) we interviewed industry experts (*ERP software providers*) in the German ERP market. We conducted 32 interviews with company representatives in 8 different roles: 7 CEOs, 6 marketing directors, 6 product manager, 5 sales manager, 3 software development manager, 3 project manager, 1 business development manager and 1 key account manager.

Every interview was preserved in a written transcript. Afterwards all transcripts were analysed using recommended techniques by Mühlfeld (1981) and Miles and Huberman (1994).
2.5 The Interview Guideline

Questions for the four aspects of the model were summarized into an interview guideline. The outline of the questions can be found in table 1.1 and table 1.2. The interviews were conducted with representatives (experts) in each company. Every interview partner represents an expert “who provides in his defined area of expertise pure and repeatable knowledge.” (Mayer 2004, p. 40).

<table>
<thead>
<tr>
<th>Main component</th>
<th>Detailed aspects of the component</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1: Architecture</td>
<td>A1P1: Modelling</td>
</tr>
<tr>
<td></td>
<td>A1P2: Concept</td>
</tr>
<tr>
<td>• Do reference processes exist in graphic form in your system?*</td>
<td></td>
</tr>
<tr>
<td>• If yes, are they used to generate code?*</td>
<td></td>
</tr>
<tr>
<td>• If yes, are these reference processes used with the customer?</td>
<td></td>
</tr>
<tr>
<td>• Are existing industry standards used? (e.g. Score, GS1, ...)?*</td>
<td></td>
</tr>
<tr>
<td>• Can business processes be adapted by the user?</td>
<td></td>
</tr>
<tr>
<td>• Does the architecture follow a general concept?* (e.g. SOA) and how is your client-server architecture designed?</td>
<td></td>
</tr>
<tr>
<td>• Is the software platform independent?</td>
<td></td>
</tr>
<tr>
<td>• Which middleware is used?</td>
<td></td>
</tr>
<tr>
<td>• What is the applied technology in browser-based systems (e.g. Java, Ajax, ...)?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A2: Technology</th>
<th>A2P1: Internal usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Which operating systems and databases are supported by your system?</td>
<td></td>
</tr>
<tr>
<td>• Which programming language was used and why? (C/C++, Java, ...)</td>
<td></td>
</tr>
<tr>
<td>• Which application development system do you use?</td>
<td></td>
</tr>
<tr>
<td>• What is your view on open source as a tool in your system?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A3: Operations</th>
<th>A3P1: In-house</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Is your software customizable for several clients?</td>
<td></td>
</tr>
<tr>
<td>• Who distributes your software?</td>
<td></td>
</tr>
<tr>
<td>• Who maintains the software (support)?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A4: Market</th>
<th>A4P1: Customer</th>
</tr>
</thead>
<tbody>
<tr>
<td>• How many customers use your software?</td>
<td></td>
</tr>
<tr>
<td>• How many users can use your system?</td>
<td></td>
</tr>
<tr>
<td>• What distinguishes you from your competitors?</td>
<td></td>
</tr>
<tr>
<td>• How much influence do customers have on the finished product?</td>
<td></td>
</tr>
<tr>
<td>• How do you organize communication with the customer? (e.g. user groups, ...)?</td>
<td></td>
</tr>
</tbody>
</table>

* Questions in bold will be discussed in the findings, in the following chapter

Table 1.1: The interview guideline, part 1

In the summer of 2007 the questions of the interview guideline were trialled in three pre-tests with selected ERP vendors in Germany and Switzerland in terms of completeness and length of the interviews and subsequently optimized regarding content and time-effectiveness. The questions were then used for the following interviews with selected ERP vendors. 17 interviews were conducted until December 2007, further 12 experts were interviewed in January and March. Overall, 130 ERP system vendors in Germany were invited for an interview of which 32 companies agreed to conduct an interview. This corresponds to a response rate of 24.6 %. All interviews were stored as written transcripts for later evaluation.
Table 1.2: The interview guideline, part 2

<table>
<thead>
<tr>
<th>Research model aspects</th>
<th>Parts of the aspects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A1: Architecture</strong></td>
<td><strong>A1P3: Functionality</strong></td>
</tr>
<tr>
<td>Which modules does your software cover?</td>
<td>Is a connection to a third party system and respectively to an existing IT landscape possible, if yes, how (data/application level)?</td>
</tr>
<tr>
<td>Which modules can be obtained externally?</td>
<td>Are there modules which are heavily interdependent (on each other)?</td>
</tr>
<tr>
<td>Are specialized modules such as CRM or BI integrated?</td>
<td>Is an authorization concept (both internal as well as external) important and if yes, how? (e.g. roles/functions)</td>
</tr>
<tr>
<td>Do you use your own solutions for CRM/BI or are these bought in? (e.g. Cognos, Business Objects, ...)</td>
<td></td>
</tr>
<tr>
<td>Is there a release concept which allows customers to use individual adaptations in future releases?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>A2: Technology</strong></th>
<th><strong>A2P2: External usage</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Which document standards does your system support in regard to external communication with other systems, (e.g. EDIFACT, ebXML,...) and do you use converters in this regard?</td>
<td></td>
</tr>
<tr>
<td>Are Web services used and if yes, how?</td>
<td></td>
</tr>
<tr>
<td>Regarding Web applications: Do you rather consider existing resources (native Web) or proprietary solutions for improved throughput?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>A3: Operations</strong></th>
<th><strong>A3P2: Hosting</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Which operational models does your software support? (ASP, SaaS, stand-alone, ...)</strong></td>
<td>In which ways and means are the functional areas (modules) of your system made available to the customers?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>A4: Market</strong></th>
<th><strong>A4P2: Competitor</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>For which sectors is your product most suitable? (e.g. trade, manufacturing,…)</td>
<td>In which position in the software providers market do you regard yourself to be?</td>
</tr>
<tr>
<td>In which position in the software providers market do you regard yourself to be?</td>
<td>Do you regard open source software vendors as serious competition? If not, why?</td>
</tr>
<tr>
<td>Do you regard open source software vendors as serious competition? If not, why?</td>
<td>How or where in the market do you see yourself in ten years’ time?</td>
</tr>
<tr>
<td>How or where in the market do you see yourself in ten years’ time?</td>
<td>How do you see the future of the market?</td>
</tr>
</tbody>
</table>

* Questions in bold will be discussed in the findings in the following chapter

### 2.6 Evaluation

The evaluation was based on an empirical content analysis (Kromrey 2002, p. 311). We applied the method suggested by Mühlfeld (1981, p. 334) to analyze the transcripts. First, every answer to the question was marked within the written text. In a second step, the answers were categorized according to a general scheme. After that the single pieces of information were combined logically so that similarities but also controversies could be discovered. The resulting text represented a first analysis from the interview guideline. In the next step it was enriched by selected citations expressed by the interview partners. In addition similar answers were summarized and evaluated based on the Conceptually Ordered Display approach by Miles and Huberman (1994). Instead of focusing on one interview we condensed the answers in a cross-case analysis. The results of these steps serve as the basis for the following discussion of the findings.


3 THE FUTURE PROMISE OF PACKAGED ERP SOFTWARE

Altogether, several important trends emerged from the interviews and the subsequent responses to the interview guideline. Due to the large amount of data (192 pages of transcript material – 6 per interview) this paper can only present some selected findings which we thought would be most interesting for an academic audience, namely (1) process orientation, (2) service orientation, (3) operational model, (4) market challenges, and (5) flexibility/agility. The complete study results can be found in (Frick 2008).

The following table contains information from the study referring to the bold questions in tables 1.1 and 1.2. The table shows the correlation between the available capabilities of process modelling and the discussed architectural and operational model of the vendors (table 2). The underlying assumption for the table is that both, the design and implementation of processes as well as service orientation enable a software package to be more flexible regarding changes (agility). The rows display the number of vendors who supply the user with ways of designing and implementing their individual processes into the software. This process goes beyond a mere customization of a packaged software solution. “Graphical processes” enable a dialogue between the implementation partner (the “consultant”) and the future user (the “business expert”). With the help of graphical workflow engines, consultant and business expert can discuss the detailed steps of a business process (with its events and ramifications). This can help to form a common understanding of how the software should finally work. Examples for these kinds of ERP extensions are event-driven supply chains (Scheer & Habermann 2000) or workflow modelling tools.

The second row displays the number of vendors who can use these graphical models to actually generate ERP source code from them. Industry standards (third row) refer to the availability of pre-defined processes which are implemented into the ERP system, e.g. a typical pricing model for retail. The last row contains the vendors who offer all three options. The columns on the other hand indicate the correlation of process modelling (rows) with the underlying technical architecture and the possible mode of operation. The table reads as follows: 8 vendors whose systems follow a fully service-oriented approach are also able to provide graphical process modelling tools. 6 of these have the possibility to generate code from the graphical models. The same comparison is also drawn with the possibilities in the mode of operation. 7 vendors who are able to offer their software as a service are also able to supply graphical process modelling.

<table>
<thead>
<tr>
<th>Flexibility in Processes and Services?</th>
<th>SOA</th>
<th>SOA-capable</th>
<th>SaaS</th>
<th>ASP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphical processes</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Code generation</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Industry standards</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>All Aspects</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2: Service-oriented systems start to fulfil their potential by utilizing processes

In the following section we discuss the implications of the table together with the qualitative remarks of the company representatives.
3.1 Process Orientation: Graphical Design Support

The first requirement for continuous process support (Liebhart 2007, p. 91), the visualisation of processes, can be fulfilled by some of the vendors (cf. table 2.1). According to Delen et al. (2005, p. 107) “the synergistic combination of descriptive graphical models created using enterprise modelling methods (…) can deliver substantial results.” 50 % of the questioned software companies offer appropriate modelling possibilities, whether through connection of existing tools or by means of self-developed display methods. However, it will remain at the stage of visualization of business processes for the next few years. The next step, a coupling of process and application level in the framework of a process-driven code generation, is currently rarely implemented. Only 25 % of the vendors currently enable code generation from modelling. The offers range from workflow rules to whole class diagrams. The lack of support for automatic process transformation will not be enhanced. Complicated internal dependencies within an ERP system prevent building the needed program structure for supporting this feature. The software companies see no practical benefits in code generation, a procedure which, in their opinion, is simply impossible because of the internal dependence of an ERP system, as the following quote from a provider shows: “ERP systems are too complex to make business process modelling worthwhile.” (Frick 2008) Furthermore, there is little support concerning a standardized view of processes. Only 25 % of the vendors currently support process standards. Due to the need to provide fast and simple solutions for the customers it will be experience-based processes that are utilized in the future.

3.2 Service Orientation: The Concept of “Service” in ERP Systems

Services have established themselves as an important or even central building block in the system conception. 53 % of questioned vendors stated that they have integrated services into their system architecture, whether as integral service-oriented architecture (SOA) (28 %) or just as a SOA capability (25 %). Primarily, their benefit is gained from the possibility of offering functionalities externally. The reasons for this lie partly in differing definitions of SOA, which are internally seen as propagated by large producers such as SAP, IBM or Microsoft and therein make a standard architectural basis difficult. On the other hand, such a consistent new conception means a reprogramming or even new programming of many system components. The possibility of avoiding this expenditure and offering the required service by means of a service level agreement will be used more often in the coming years. Although it may sound strange, an interview partner expressed it as follows: “We can’t do SOA, we are SOA.”

3.3 Operational Model: Hosting Solutions for ERP Systems

Solutions on a rental basis have, since their hype at the end of the 1990s, found their way back into the customer’s awareness. SaaS has established itself as its own term for an operational model next to Application Service Providing (ASP) and will increasingly emerge as a possible solution from the software companies. Although SaaS is often equated with ASP, the differences between the earlier ASP (classical licence-based software is made available over the Internet to a customer) and SaaS (a software designed for Web use which serves several customers) are so clear that it can be defined as its own operational model. Although hosting solutions such as ASP (possible with 56.25 % of the products) or SaaS (possible with 31.25 % of the products) increasingly find their way into the offers of software companies, the in-house variant remains the main pillar of the product range. In addition, the technical possibility to host a system does not imply that vendors offer this service as operational functionality due to lack of customer trust. One vendor expressed the overall attitude towards hosting as operational model: “We provide in-house, ASP and SaaS but neither ASP nor SaaS are desired by customers.”
3.4 Market Challenges for the Future

For the future, 12 providers estimate an advanced consolidation as most likely. In the eyes of the questioned companies, the development of the market seems to clearly indicate a decrease in the number of providers. Exactly which group this will affect seems to be controversial. The niche suppliers are convinced that their excellence in their particular sector will protect them from great competition because cross-sector systems do not have the depth of functions which their customers need: “From our viewpoint the trend towards specialization in terms of sector-focus will continue further.” (vendor quote) 10 vendors forecast technological advancement as an influential factor for the market. Web-based architectures or fully modularized toolkits for building an ERP system are mentioned. Economical pressure by global players, in particular acquisitions by large providers like SAP or Oracle, is less eminent for the questioned vendors (7). So the statement concerning the ongoing consolidation does not necessarily imply a forceful buying strategy by global players that vendors are afraid of. Instead there is a general feeling that many vendors especially SMEs will perish during the process: “There won’t be several hundred vendors in five years time but only 20 to 30.” 5 vendors enhance this view by their comment towards a decreasing variety of systems in the market. Interestingly, there are only 5 system providers who focus on new operational models for differentiation despite the current obvious trend towards hosted solutions.

3.5 Flexibility of Processes and Services: Basis for Agile Enterprises

The already existing potential in Service Oriented Architectures (SOA) in conjunction with standardized process mapping (e.g. BPMN and BPEL) has, because of a lack of support from process modelling standards, not yet been exploited. A service orientation has already been implemented by 17 companies. In addition 13 of 17 vendors who use services in their system (internally and externally) do also have a graphical process representation. However, an orchestration of the services over and above the planning level is only possible to a very limited degree. Code generation is almost exclusively available for SOA based systems (7 of 8 vendors) so the infrastructure for combining processes with underlying services exists. Nevertheless the code which is currently produced is limited to specific code fragments (class diagrams, workflow rules, etc.).

The use of standards in process modelling is independent of the underlying architecture. Two vendors with a SOA and three with a SOA capability use industry standards. This means that the idea of a system which supports the ability to adapt to changed business processes will remain the exception in the coming years. The underlying architecture necessary to achieve such freedom and power of design would have to be component-oriented and also service-oriented up to current standards (Sprott 2000), so that the functionalities connected with the processes can, without any problem, be rearranged according to the restructuring carried out in the model. This will also help to meet increasing requirements regarding improved agility (Osterhout et al. 2006; Overby et al. 2006).

Based on these findings it is safe to assume that the graphical representation of processes will be more likely in combination with service-oriented vendors. The same conclusion is feasible for code generators that will be used within service-based scenarios. The architectural support for these tools on the other hand will be limited to a few vendors. The majority of vendors are not ready or able to rethink and rebuild their systems accordingly. When it comes to industry standards, service-oriented systems are more likely to include them. But without a common sense of well conceptualized and reusable process scenarios there will not be many systems in the future making use of standards.

4 CONCLUSIONS AND LIMITATIONS

Based on the results of our empirical study of the German speaking ERP market we discussed the aspects of process orientation, service orientation, operational model, market development, and
flexibility (agility) concerning future ERP software suites. We collected a large amount of data and decided to focus this paper on vendors whose systems are service-oriented and who offer hosting solutions for their ERP system. These vendors have, to some extent, already met the corresponding technological and organizational requirements for a flexible system that can cope with future technological and business-related challenges. However, in some areas there is still a clear need for development.

Process orientation is partly implemented in terms of a visual representation, it has not, however, come into its own yet as a standard organizational tool. Particularly in continuous process design and in the implementation of process standards there are still deficits requiring intensive further development. Modelling should not remain on the simple visual level but should, together with a service-oriented application system, allow for a reorganization of functions above the planning level. There are already corresponding approaches, although they are only used by a small proportion of the questioned companies. Additionally, there are standardized processes for cross-company process modelling for which implementation is lacking.

In the service-oriented areas, the necessary measures for a continuous application provision are to a large extent already implemented. Nevertheless, there are various aspects which still require further attention. SOA is still not clearly defined by all the companies questioned. However, the service idea is at least known of as a rough concept by all those questioned. Based on services, legacy systems can more easily be re-used, new modes of operations can be offered and acquisitions can be more easily integrated into the existing system. Nevertheless, a disciplined service administration is needed to ensure clarity, security etc., especially for large scale use.

Around one third of the companies see a progressive market consolidation happening in the future. This consolidation is likely to happen because of the growing pressure from global players such as SAP or Oracle who want to enter the SME market. Technological advancements as SOA and its possibility for simple application integration, or modern operational models as SaaS are contributing to the current market consolidation. Oracle alone acquired 41 companies within 48 months (Hill 2007). Some of the interviewees suggested a scenario in which some providers in the SME sector form so-called ERP clusters in order to arm themselves against the ever harder competition. For smaller ERP companies, it is reasonable to either secure a particular advantage in the market through niche excellence or to join forces and form a kind of “cluster” with other ERP companies in order to make a takeover unattractive: “It will be more important to link with others in the future.” (vendor quote) New technologies can even accelerate this process, as the expansion of Web services in the context of SOA has shown. One vendor summarized it: “As far the market goes, it’s: evolution not revolution!”

The paper presents selected aspects of a detailed analysis of the future requirements of ERP software in the German-speaking market based on an empirical study. We were confirmed in our belief that this topic still needs further research by the editors of the EJIS special issue on packaged software who stated: “In summary, we believe there is a need to theorize about packaged software and its place within the field of information system.” (Light & Sawyer 2007) Combining qualitative with quantitative research aspects, we shed light on the research question of what future packaged ERP systems are going to look like. Our research has several limitations. Firstly, the discussion is focused on the German-speaking area and it is possible the results may differ in other countries due to cultural implications (Soh et al. 2000; Davison 2002). Secondly – although the return rates were very favourable – this study is subject to the usual constraints regarding statistical representativeness. 32 interviews with vendors can only reflect a small portion of the complete universe of German-speaking ERP users and vendors.
References


Schmitt, K. (2007). ERP kommt in Deutschland langsam, aber gewaltig. silicon.de, Technologie und Business (20.03.07), URL: http://www.silicon.de/enid/wirtschaft_und_politik/26094, last access: 09.03.08.

Proceedings ECIS 2009
**EXPLORING ANTECEDENTS OF ORGANISATIONAL ADOPTION OF ERP AND THEIR EFFECT ON PERFORMANCE OF FIRMS**

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0314.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Adoption, Critical Success Factors (CSFs), Enterprise resource planning (ERP) (packaged systems), IS Performance</td>
</tr>
</tbody>
</table>
EXPLORING ANTECEDENTS OF ORGANISATIONAL ADOPTION OF ERP AND THEIR EFFECT ON PERFORMANCE OF FIRMS

Ram, Jiwat, School of Computer and Information Science, University of South Australia, 27 North Terrace, Adelaide 5000, South Australia, jiwat.ram@unisa.edu.au

Pattinson, Malcolm, School of Computer and Information Science, University of South Australia, 27 North Terrace, Adelaide 5000, South Australia, malcolm.pattinson@unisa.edu.au

Abstract

The research aims to address the existing theoretical gap in knowledge on antecedents to organisational adoption stage of the Enterprise Resource Planning (ERP) innovation process, and their relationship to performance of the firms. The predominant focus on investigating the implementation stage and its related issues has resulted in relatively limited understanding of the other stages of ERP innovation process. Additionally, little is known on how important critical success factors (CSFs) are to performance improvements of the firms. Grounded in innovation diffusion & IS Success theories, this study attempts to identify the antecedents critical to the success of the adoption stage of ERP innovation process and then examine their effect on organisational performance. Employing a mixed method research design for data collection and Structural Equation Modelling for analysis, the study empirically test a research model in order to develop an integrated, holistic framework for success of ERP innovation process. This research expects to contribute to the development of theory on innovation process by testing new antecedents to ERP adoption stage context as well as enhancing the understanding of CSFs’ role. The study will also provide recommendations for both the potential adopters and the vendors on successful uptake of ERP system.

Keywords: Adoption, Enterprise Resource Planning, Critical Success Factors, Firm Performance.
1 INTRODUCTION

Innovation is increasingly being hailed as the transformative engine that creates and shapes new economies of today’s networked world. The growth and survival of businesses is linked both to the creation of new products and services and to the adoption of novel ways of doing business whilst constantly improving the internal business processes, procedures, policies and business models (Damanpour & Schneider 2006). This ‘innovation pressure’ is thus translating into a tendency among firms to adopt information and communication technology / system (ICT/IS) innovations. One such technological innovation is the Enterprise Resource Planning (ERP) system.

With its hallmark enterprise-wide information sharing capabilities, ERP has gained popularity as the most rewarding information system of the 21st century (Davenport 2000). Paradoxically, the popularity of ERP does not seem solely to stem from its potential and promised benefits, but also partly comes from its notoriety for being complex to implement and various instances of its failed implementations (Ehie & Madsen 2005, Mabert et al. 2003). Reporting accurate information on ERP failures might be difficult if not impossible. However, it is generally believed that an estimated 60-70% of ERP projects, either fail to implement successfully or face various obstacles during implementation (Harvard 2007, Leung 2008). While such estimates seem very alarming, the lack of clear understanding on how ERP systems can be implemented successfully is still vague, thus adding to the burgeoning problem.

ERP projects fail due to a variety of reasons. Some of those reasons are: the complexity of the implementation process, misfit between ERP and organisational business processes, lack of availability of skilled human capital, cross-functional coordination issues, inappropriate planning and project management and incompetent consultants (Kim & Lee & Gosain 2005). All these reasons seem to embody one common theme: that is a sound understanding of the innovation process is essential in order to realise success in ERP projects (Laukkonen & Sarpola & Hallikainen 2007). This realisation has led to extant research, in terms of both the volume and diversity, on ERP related themes such as CSFs to ERP implementation, change management, cultural issues, trends and perspectives, use, maintenance and cross border implementations etc. (Moon 2007).

Despite the proliferation of studies on ERP, the research examining the organisational innovation process of ERP is still at a preliminary stage (Mabert et al. 2003) whereas the research on the adoption stage of the process is scant. The organisational innovation process is the combination of various stages or phases ‘along the temporal dimensions of their anticipated sequence’ (Rogers 2003, Zaltman & Duncan & Holbek 1973 p.52). Suffering from a lack of consensus on a globally accepted subdivision and nomenclature for the stages of an innovation process, the theorists have commonly identified these stages as initiation, matching, adoption, implementation, use and retirement of innovations (Kwon & Zmud 1987, Rogers 2003). The literature review indicates that the major thrust of the research on ERP has concentrated on the implementation phase generally and CSFs to ERP implementation in particular (Esteves & Bohorquez 2007, Moon 2007). This trend has resulted in a relatively limited understanding of the other stages of ERP innovation process such as adoption, use and retirement (Kamhawi 2008). Perhaps, the peculiar focus on investigating ERP implementation process is either the outcome of a sense of urgency among the researchers to find a ‘mantra for success’ to curb or to slow down the failures of ERP implementations, or it may be due to the reason that implementation, in ERP research context, is being treated as an inclusive process combining initiation, adoption, implementation and use. Regardless of the reason, it is crucial to avoid this “one size fit all” approach and any conceptual overlap between implementation and other stages of the ERP innovation process. This avoidance is vital, since understanding of each stage will enable the firms to proactively seek relevant solutions for achieving success at each specific stage of the ERP innovation process. Without diagnosing every stage of the ERP innovation process specifically, it seems unlikely that sound solutions to problems in ERP projects can be developed and its failures can be reduced or eliminated (Peslak & Subramanian & Clayton 2008). Equally important is the fact that the literature on ERP implementation is characterised by inconsistent, inconclusive findings and it lacks maturity (Ngai
& Law & Wat 2008). Therefore, there is a need for a fine-grained and well-structured approach to research in ERP context (Finney & Corbett 2007).

The adoption stage is the decision making stage in an innovation process (Rogers 2003, p.420). This is a very important stage as the ERP adoption decision entails huge monetary and resource commitments. An incorrect adoption decision could ‘well jeopardise the very existence of the organisation’. A well grasped understanding of CSFs relating to the adoption stage can help firms to achieve savings in costs, time and reduce the risk of failures by creating a more ideal environment and adequately preparing themselves (Verville et al. 2005). Furthermore, the identification of CSFs specific to every stage of the innovation process is of paramount essentiality, as each stage has different focus, priorities, stakeholders and performance metrics. With a very limited research on CSFs to ERP adoption stage, this study addresses these gaps in knowledge by investigating the factors critical to the success of adoption stage.

Though the significance of CSFs is recognised, little is known about whether the impact of CSFs is limited to the success of the relevant stage alone or does it extends beyond that, influencing the performance of firms in long run. Therefore, the study investigates the relationship between CSFs to adoption stage of ERP and their performance impacts on firms. An investigation of this relationship is vital in order to understand the real impact of CSFs (Nicolaou 2004).

Finally, literature presents a strong case for examining factors specific to the adoption of complex ICT/IS innovations, as the conventional factors presented by classical innovation diffusion theories (e.g. Rogers 2003) may not prove to be strong predictors to understand the innovation process of complex technologies (Hong & Zhu 2006). As such, the study fulfils this gap by testing factors specific to ERP adoption in particular and complex ICT/IS in general.

In the context of the literature gaps discussed above, this study investigates the following questions:

- What are the antecedents critical to successful adoption of ERP?
- What is the impact of successful adoption of ERP on the firm performance?
- What are the effects of CSFs for adoption of ERP on performance of firms?

This paper is organised as follows. First, in Section 2, a review of the literature pertaining to ERP adoption, implementation problems, CSFs and performance impacts is presented. In Section 3, the theoretical foundation of the study is discussed. Next, a theoretical framework identifying antecedents and their structural relationship is proposed in Section 4, followed by a discussion of methodology in Section 5. Finally, the implications and limitations of the research are discussed in Section 6.

2 LITERATURE REVIEW

2.1 ERP: significance and implementation problems

ERP, a term coined by the Gartner Group in the early 1990s, is a commercially available off-the-shelf software system, which embeds best business practices (Cotteleer & Bendoly 2006, Mabert et al. 2003). ERP offers extended functionalities to support critical business processes in many functional areas throughout the whole enterprise as compared to its specific-purpose predecessor systems such as Material Resource Planning (MRP) and Manufacturing Resource Planning (MRP-II) system of 1960 and 1970 respectively (Mabert et al. 2003). SAP, Oracle, Baan, PeopleSoft, JD Edwards, Technology One and MS Dynamics NAV are some of the well-known ERP systems.

ERP enables process and information integration across all functional domains of an organisation. It is defined as the customisable, modular structured software system that embeds best business practices and seeks to automate, standardise and integrate the key business & management processes and information using a common database and IT architecture; and provide real time information flow (Mabert & Soni & Venkataramana 2003).
The investment in ERP entails a variety of tangible and intangible benefits. Shang and Seddon (2000) and Basoglu, Daim and Kerimoglu (2007) list some of the benefits: (a) operational as in: reduced operating costs, accurate demand forecasts; (b) managerial as in: improved decision making and better resource management; (c) strategic as in: greater support for business alliances, building business innovations and cost leadership; (d) IT infrastructure as in: building business flexibility; reducing ICT costs; and (e) organisational benefits as in: supporting organisational change, facilitating better communication, business learning and empowerment. Historically, however, the implementation of ERP was carried out to replace fragmented, multiple, difficult and costly to use & maintain legacy systems – a transition that was neither easy nor quick.

Basoglu et al. (2007) cited figures from a Standish Group report on ERP implementation, note that ERP projects are, on an average, 178% over budget, take 2.5 times longer than intended to implement; and deliver only 30% of the promised benefits. Supporting this line of argument, several other researchers have reported a number of failed ERP projects with some of the projects involving multinational firms, including Boeing, Siemens, Panasonic, NAB Australia, AeroGroup, Dell, FoxMeyer, Whirlpool, and Dow Chemical (Ferguson, 2004, Karim & Somers & Bhattacherjee 2007). On the contrary, several ERP projects have been implemented successfully or have resulted in a number of benefits to adopting organisations too. Some examples are: IBM Storage system which achieved a reduction in shipment time for replacement parts from 22 days to 3 days, and Par Industries which reaped delivery performance improvement from 60% to 95% with a lead time to customers reducing from 6 to 2 days (Ehie et al. 2005).

The foregoing discussion highlights the opportunities and risks associated with adoption of ERP systems. Indeed, a failed ERP project could entail a potential disaster to the survival of the adopting organisation.

2.2 Critical Success Factors (CSFs)

The concept of critical success factors was originated in 1960s and it gradually gained practical acceptance in 1970s when it was used for the design of a management control system (Leidecker & Bruno, 1984). Since then, it has been variedly used across different disciplines such as organisation behaviour, business, management, marketing and information technology / systems. Also named as strategic factors, key success factors, key result areas and pulse points; the CSFs have been widely conceptualised as the ‘few key areas where things must go right’ for the firms to achieve competitive performance and profitability (Leidecker & Bruno 1984 p.23, Rockart 1979). Hofer and Schendel (cited in Leidecker & Bruno 1984) proposed a deterministic view of the concept and described CSFs as those areas, which can be influenced by management decisions to achieve competitiveness – a view that seemed limited in conceptual utility and sounded too simplistic. However, Leidecker and Bruno (1984, p.24) took a more comprehensive and deeply insightful approach in defining the concept of CSFs as “those characteristics, conditions or variables that when properly sustained, maintained, or managed can have a significant impact on the success of a firm competing in a particular industry.” This view and definition are shared and adopted as fundamental concepts in the present study.

The determination of CSFs is vital to organisations, to constantly and sufficiently focus on activities in already identified key areas or in identifying key areas for achieving the performance improvements. CSFs have been classified in a variety of ways such as strategic, tactical, internal / firm specific, external / industry specific, technological / innovation related and environmental (Finney & Corbett 2007, Leidecker & Bruno 1984). Given that the success of the ERP projects has been widely termed as ‘a moving target’ (Karim et.al. 2007), the identification of CSFs has attracted considerable literary interest. The CSFs are assumed to provide a positive, beneficial effect on the outcome, and thus knowing greater number of CSFs pertaining to a specific innovation process has been considered as critical to success of the process (Karim et.al. 2007). Despite the fact that conceptually CSFs are the areas that are linked to performance impacts on the firms, there seems to be a very significant gap in the literature in understanding the true role of CSFs on the firms studied. Generally, previous studies
investigating CSFs seemed to conceptualise them as an independent object relevant to a certain process, without investigating how critical the related CSF is to performance improvements. This situation warrants immediate attention and requires further examination of the role of CSFs in their interactive effect on performance improvements of the firms.

2.3 Adoption & Performance impacts of ERP

Referred to as a go no-go decision point, adoption is a key stage in the innovation process (Rogers 2003, p. 423) and has been extensively investigated in innovation diffusion literature. It has, however, received little attention in ERP literature as a bulk of studies have focused on investigating the implementation process of ERP (Esteves et al. 2007).

The literature points that assessment of benefits and risks of ERP has remained of strategic interest to the adopting organisations. Achieving operational efficiencies such as improvements in productivity, optimising inventory and data integration capabilities are some of the prime benefits being sought by ERP adopters (Kamhawi 2008). In investigating adoption motives, Raymond and Uwizeyemungu (2007) conclude in their study of Canadian Small & Medium enterprises (SMEs), that the firms with significant organisational capacities, commercial dependence on major customers and tendency of bringing innovative products are internally predisposed to ERP adoption, whereas those firms which are focused on networking and partnerships with other firms are externally pre-disposed towards adoption of ERP. Unfavourably disposed firms are profiled as having less conducive environmental, organisational and technological tendencies towards the adoption of ERP.

Extending the investigation on motives of adoption, Kamhawi (2008) finds that gaining strategic management and decision making capabilities are the main themes influencing ERP adoption decision, whereas startup costs and availability of resources are the main challenges to adoption of ERP. The findings indeed confirm the standpoint of Huang and Palvia’s (2001) study in which they highlight the differences in challenges faced by SMEs in adopting ERP, in developing and developed countries. Huang and Palvia (2001) find that the SMEs in developing countries face various cultural, economic and infrastructure challenges to adoption of ERP in contrast to their developed countries counterparts. However, Buonanno et al. (2005) contradict this position and conclude that SMEs do not regard financial constraints as the impediment to ERP adoption, but consider organisational and structural factors as the main influencing reasons thereof.

In an attempt to explore CSFs to ERP adoption, Deep, Guttridge, Dani and Burns (2008) suggest various factors across four key activities: plan, identify, evaluate and select. These factors are core team formation, review of requirement planning, ranking of needs / priorities, list of available resources, software and vendor short list criteria, arranging workshop and demonstration for end-users and customer site visits. Interestingly, Elbertsen, Benders and Nijssen (2006) discover that firms having less knowledgeable people / IT managers, who consider ERP as a complex and incompatible system, exhibit greater receptiveness to ERP marketing efforts and tend to adopt ERP more. The findings point to the opportunities available to ERP vendors for marketing customisable or best-of-breed systems. They also draw attention to the need for ERP vendors to review their marketing approach and step up efforts to target both well informed and less informed potential adopters.

Given the considerable investment of time, money and resources in ERP projects, the researchers have looked at the performance impacts of ERP in a post adoption scenario. Conducting a survey of Hong Kong based firms, Law and Ngai (2007) find that user satisfaction of ERP and business process improvement positively impact the organisational performance. They claim positive empirical relationship between the strategic intent behind the adoption of ERP and organisational performance. Velcu’s (2007) study reinforces the above findings, as the study reveals that firms driven by technologically-led motives versus business-led motives, perceive differently towards benefits of adopting ERP. Poston and Grabski’s (2001) longitudinal study of pre versus post implementation present mixed financial benefits to the ERP adopters. Hunton, Lippincott and Reck (2003) argue that a decline in financial performance of non-adopters relative to adopters should be anticipated and
studied, as adopting firms may not necessarily exhibit financial gains immediately for a number of reasons.

Among the studies measuring the non-financial impacts of ERP, O’Leary (2004), using data from companies which had opted for Oracle’s ERP, identifies several tangible and intangible benefits of ERP implementation. Considering ERP’s multi-directional impact, McAfee (2002) observes significant operational performance improvements in the pre versus post ERP implementation contexts. Supporting these findings, Cotteleer and Bendoly (2006) discover significant performance improvements in order fulfilment lead times in the near-term, as well as over an extended period following the deployment of an ERP system.

Interestingly, the time lag has been seen as quite important in assessing ERP performance outcomes. Although the ‘normal’ time lag has been found to be approximately three or more years, Matolcsy, Booth and Wieder (2005) report sustained operational efficiencies and overall improved liquidity for adopting firms after the second year of implementation. Similarly, Nicolaou’s (2004) study supports this view of early impacts on performance and found that ERP implementing firms realised performance impacts (ROI improvements) in their 2nd year after implementation.

Urging further empirical studies that consider implementation factors and their linkage to performance impacts, Nicolaou (2004) concludes that implementation factors such as vendor selection, implementation goals, scope and time on ERP implementation efforts are significant in affecting a firm’s realisation of performance impacts.

3 THEORETICAL BACKGROUND

3.1 Diffusion of Innovation

Over the past 50 years, innovation scholars have produced a number of theoretical models, terminologies, typologies and nomenclature, to explain the concept and the process of innovation (see for review: Kwon et al. 1987, Rogers 2003, Zaltman et al. 1973). To attribute this situation to the cosmopolitan characteristics of innovation or to the bandwagon effect of innovation is a matter of debate which is beyond the scope of this study, however, the resultant literature certainly seem to be converging towards a ‘core set of theoretical models’ (Kamal 2006). Supporting this line of argument, a review of various theoretical models indicates that the innovation process have been explained through four core processes: generation, diffusion, adoption / implementation and retirement / obsolescence. These core processes have further been divided into sub-processes such as initiation, matching, assimilation, use, and infusion to provide deeper insights into the various stages of the innovation process. Diffusion of innovation (Rogers 2003), IS Implementation Model (Kwon et al. 1987), Technology Acceptance model (Davis 1989), Innovation adoption and implementation Model (Gullivan 2001), Theory of Reasoned Action (Ajzen & Fishbein 1975) and Theory of Planned Behaviour (Ajzen 1985) are some of the core models of innovation (Cited in Kamal 2006). The linear, sequential and formulaic approach adopted by these models has attracted criticism (Bayer & Melone 1988) as well as validation in multiple studies across different disciplines (Damanpour et al. 2006, Molla & Licker 2005).

Rogers (2003) has described innovation process from individual as well as organisational lenses, explaining several attributes that influence the process. According to Rogers (2003, p.420), the organisation innovation process is a sequential combination of five stages, which is made up of a two-stage initiation sub-process, namely agenda-setting and matching; and a three-stage implementation sub-process, namely redefining/restructuring, clarifying and routinising. Rogers’ (2003) innovation models have been tested in cross-disciplinary contexts and many scholars have presented updates, extensions and revisions to this model. Rogers’ (2003) model has also been termed as being too formulaic and lacking in realism (Bayer and Melone 1988). Given the scope of this study, however,
the authors employ Rogers’ (2003) stage-based organisational innovation model as it lends well with the objective of the study.

3.2 IS Implementation Model

Consistent with Rogers’ approach, Kwon and Zmud (1987, p.233) proposed a multi-stage model of IS implementation, theorising that organisational innovation follows six stages, namely initiation, adoption, adaptation (development/installation), acceptance, use and incorporation. Extending Rogers’ (2003, p. 392) five-stage organisational model, the IS implementation model emphasises that four assessments: acceptance, usage, performance and satisfaction, should be incorporated into the innovation process model, as these assessments form the basis for implementation success. Since, the IS implementation model has been widely tested and has proved robust in explaining the organisational innovation process of information systems, this model serves as a framework in the present study.

3.3 IS Success Model

Delone and McLean (1992), in their famous IS success model, posited that system quality and information quality jointly and severally affect use and user satisfaction which, in turn, result in individual impact and subsequent organisational impacts. Further research into this model, however, had led to the addition of dimensions such as service quality, intention to use and Net benefits (replacing organisation impact construct) (Delone and McLean 2004).

4 RESEARCH HYPOTHESES AND MODEL

This study uses a synthesis of three models described above to test attributes critical to the adoption of ERP and their effect on firm performance. Consistent with broad category of attributes identified in these theories, the research uses environmental, technological / innovation related, organisational and IS Success related factors to create a hypothetical model, which is shown graphically in Figure 1.

1) Perceived Information quality
Information quality refers to measures of IS output, and has been measured in a variety of ways, e.g.: accuracy, output precision, output reliability, information timeliness, information relevance to decisions, completeness, format, understand-ability etc. (Delone & McLean 2004; Nelson & Todd 2005). Information quality has been found to relate positively to user satisfaction (Nelson & Todd 2005), so that ERP information quality would appear to be critical to adoption success of the system. Further, ERP systems are generally recognised as providing reliable, accurate and timely information. It is thus argued that ERP’s perceived information quality is a positive contributing factor to adoption decision, and the following hypothesis is developed:

**H1: ERP’s ability to provide information quality positively influences the successful adoption.**

2) Perceived System quality
DeLone and McLean (2004) found system quality to be one of the most important enablers of IS success, in their meta-analysis of the IS literature. Systems quality has been measured in a variety of ways, e.g. convenience of access, integration capabilities, reliability, ease of learning, resource utilisation, investment utilisation, flexibility of system, response time and usefulness of specific function etc. (Delone & McLean 2004; Nelson & Todd 2005). It has been seen as an important factor for ERP implementation and was positively associated with user satisfaction (Nelson & Todd 2005). ERP systems are believed to provide integration, flexibility, and optimum resource utilisation – and thus provide high system quality. It is therefore argued that ERP’s perceived system quality capabilities will be an important factor to adoption and its success thus develop the following hypothesis:

**H2: ERP’s ability to provide system quality positively influences the successful adoption.**
3) Organisational Readiness
Organisational readiness has been defined as “the ability of a firm to successfully adopt, use, and benefit from information technologies” (Fathian & Akhavan & Hoorali 2008). Prior studies have used a variety of measures to investigate readiness, e.g. awareness of benefits and risks of innovation; availability of human resources skills and capabilities; availability of technological, business and financial resources, commitment and support by top management, fit between innovation and organisational structure as well goals and values of organisation (Fathian et al. 2008; Molla & Licker 2005). Grounded in structural contingency theory of fit, Khazanchi (2005) believes that assessment of ‘readiness’ for an organisation to adopt a certain technology is an important criterion for successful implementation and performance impact. Researchers often posit that top management attitudes to change have significant influence on adoption outcomes (Wu & Mahajan & Balasubramanian 2003). Organisational readiness positively impacts adoption of technological innovations (Molla & Licker 2005). An ERP adoption introduces enterprise-wide change. Hence it is expected that the internal organisational preparedness will be critical to adoption success and therefore it is hypothesised:

**H3: Organisational readiness positively influences the successful adoption.**

4) Environmental Assessment
Chi, Jones, Lederer, Li, Newkirk, and Sethi (2005) note that: “an environmental assessment evaluates external information and identifies business needs, objectives, external opportunities, and threats”. Though environmental assessment has been measured using many indicators, it broadly encompasses: hostility, dynamics and heterogeneity of environment (Newkirk & Lederer 2006). These authors note that environmental hostility includes unpredictability of competitors’ market activities; legal, political or economic constraints; price / product quality competition; labour scarcity; and Product / service differentiation. By contrast, they argue that, environmental heterogeneity includes diversity in production and marketing methods; in customer buying behaviours; in the nature of competition; and in product line. Finally, growth opportunities, change in production/service technologies, rate of innovation, product / technology changes etc. indicate environmental dynamics. Competitive pressure, normative pressure and customer power have been found to positively impact adoption of technologies (Wu et al. 2003).

A dynamic, heterogeneous and hostile business environment may affect stability of demand, put strains on supply, generate a disloyal customer base; and result in fluctuating economic outcomes. Thus, a system that predicts, coordinates and forecasts market trends will enable the organisation to react swiftly and efficiently to market changes (Wu et al. 2003). ERP systems, with their flexible and integrative architecture, seamless data flow and real-time global connectivity features have the capacity to help forecast demands and supply variations, support sound decision-making, efficient utilisation of resources and achieve competitive advantage. It is therefore postulated:

**H4: An Environmental assessment characterised by environmental hostility, dynamism and heterogeneity positively influences the successful adoption.**

5) Perceived Strategic Value (PSV)
Information systems form the key strategic assets of an organisation’s asset portfolio. Thus their adoption is motivated by “business justification and strategic value” which the new systems are seen as bringing to the firm (Subramanian & Nosek 1993). Subramanian and Nosek (1993) pioneered the concept of perceived strategic value and presented three factors: operational support, managerial productivity and strategic decision aid, to measure perceived strategic value. Amit and Zott (2001) found four key factors of strategic value in e-business context: transaction efficiency, complementarities, lock-in and novelty. Grandon and Pearson (2004) concluded a positive relationship between PSV and adoption. ERP involves a huge monetary and resource commitment and promises benefits in operational, managerial and organisational domains. It is considered a tool for achieving operational efficiency and competitive advantage, thus it is argued that a positive assessment of ERP’s perceived strategic value would contribute to successful adoption of ERP. Thus the following hypothesis is developed:

**H5: Higher Perceived Strategic Value of ERP positively influences the successful adoption.**
6) Adoption, Implementation Success and Firm Performance

Success of ERP adoption is dependent on combination of various factors. A clearly defined vision and mission, the top management support and leadership, the detailed evaluation and selection process of ERP vendor and a detailed cost benefit analysis of the ERP project are the foundation stones for the success of ERP project at implementation stage and post-implementation performance of the organisation. This study adopts these items to measure the adoption of ERP and its success. Literature suggests that the implementation success have been measured in variety of ways, such as ERP project success, performance of ERP etc. Implementation in this study has been measured in terms of project success using four items: completion on time, within budget, as per expectations and level of user satisfaction with the ERP. It is argued that performance should be measured as the outcome of adoption and implementation of ERP, in order to understand the impact of ERP in post adoption and implementation context. Hence firm performance has been measured by two underlying latent variables: competitive advantage and overall performance of the firm. Dehning & Stratopoulos (2002) suggest “a company achieves competitive advantage by implementing a value creating strategy that is not being implemented by competing firms.” The adoption of strategic information systems such as ERP have been linked to achieving capabilities such as agility in offering innovative products and services, effective leveraging of resources, raising entry barriers for new competitors, achieving increased bargaining power with suppliers and customers and changing the dynamics of competition (Bhatt & Grover 2005). As such a ten-item measure as proposed by Teo and Pian (2004) has been adopted to tap into competitive advantage gained by the organisations in post ERP context. Overall performance of the firm is measured by seven item measure to check the impact of ERP on customer service, productivity and sales growth of the adopting organisation. Based on foregoing discussion, we argue,

H6: Adoption success positively influences implementation success.
H7: Adoption success positively influences firm performance.
H8: Implementation success mediates the relationship between adoption success and firm performance.
H9: Perceived information quality, perceived system quality, organisational readiness, environmental assessment and perceived strategic value have positive relationship with firm performance.

Figure 1. Preliminary Research Model

5 RESEARCH METHODOLOGY

This research project is designed to empirically test and validate the hypothetical structural model presented in Figure 1. The study employs a mixed method approach in data collection. The quantitative data will be gathered through a cross sectional survey and be used to identify the influence of various organisational, technological / ERP related, environmental and IS Success related variables on the adoption stage and their effect on the performance outcomes of the firms studied. It will be followed by semi-structured interviews with respondents in five firms to supplement and cross-validate the findings of the survey and probe further insights into the issues. Given the complex nature
of ERP adoption and implementation process, the choice of mix-method approach suits the objectives of this study as it will allow the testing of the hypothesised relationship among various factors as well as explore in greater depth the process by which the relationship among these factors occurred (Tashakkori & Teddlie 2003). The data collected through the cross-sectional survey is expected to provide greater external validity and generalisability of results, and afford an opportunity to analyse the collected data using a robust statistical tool. Considering that the conceptual model of this study is grounded in theory, takes a confirmatory approach to test the various hypothesised relationships a priori and includes relationship between multiple latent and measured variables, Structural Equation Modelling (SEM) is the most appropriate approach to be used for analysing data.

The population for the study is the Australian organisations that have implemented ERP systems. The sample frame has been identified using “Fairfax Business Research's MarketBase” companies database which includes decision-maker & company contact information, ERP software details and financial data of the company. Consistent with the prior research into ICT innovations and given the focus of this study, the target respondent group is senior managers for example: Chief Executive Officer, Chief Technology Officer, Chief Information Officer, Project Managers and business managers with dedicated involvement in the adoption and implementation of ERP in their respective organisations. This set of group is chosen as they fit the criteria of: (a) knowledge ability concerning the content of the enquiry and (b) ability to generalise patterns of behaviour after summarising either observed or expected organisation relations (Wu et al. 2003).

All latent variables / factors shown in the model (See Fig. 1 above) are being measured through multiple item scales, which are expected to provide a stronger construct validity. Wherever possible, measurement items which have been operationalised and tested in previous empirical studies have been re-used, which will increase the comparability and reliability of the results. All the items in the survey are being measured using a five-point Likert-type scale with response alternatives ranging from strongly disagree to strongly agree.

A two-phase pilot test of the survey instrument have been conducted, with fifteen academics and nine ERP practitioners / professionals respectively. The feedback has been incorporated into the survey instrument and where necessary modifications have been made to suit the context of the present study.

The web or Internet based surveys are seen as a convenient platform in achieving higher response rates due to their ease of use, low cost and greater interactivity (Dillman 2007). As such the survey questionnaire has been replicated into its web-based version and is available at www.surveymonkey.com. A link to the online questionnaire has been provided in the covering letter. Respondents are given the choice either to respond in hardcopy or online. To avoid duplication, an identifier number has been allocated to each respondent.

Data analysis will be carried out using a two-step approach: (a) assessing the measurement model; and (b) analysing the structural model for adequacy of data representation. The measurement model will be tested for validity and reliability properties. Construct validity will be assessed by convergent and discriminant validity, using confirmatory factor analysis (CFA).

6 CONCLUSION

This study intends to propose a new systematic direction to research on CSFs on innovation process of complex technologies in general and ERP in particular. By adopting a structured and an integrated approach, it is believed that the shortcomings of traditional CSF oriented studies could be overcome. The question, how important critical success factors are, requires an in-depth examination of the interactive effect of CSFs beyond the particular innovation process, and to further understand the role of CSFs in the spirit of the concept.

The findings are expected to have several important academic as well as practical implications which include: a) improving our understanding of the adoption stage of ERP through identification of key
attributes to the process. This would help firms to study their conditions against a set of identified CSFs, which is expected to improve their chances of eliminating or reducing the risk of failure or improving the chances of success of their ERP projects (Ngai et al. 2008), b) testing of a new combination of CSFs to ERP adoption, will shed light on the complex interaction and effect of these variables on the performance outcome, c) identifying new factors that could improve the predictive power of diffusion theories in explaining the uptake of complex technologies including ERP, d) providing a better understanding of whether CSFs role is limited in influencing the outcome of relevant stage in an innovation process or it goes beyond in influencing the performance of the firms as well, e) helping ERP vendors to understand the needs of potential ERP adopters, and thus formulate their marketing and product development activities to increase their market share and f) laying the foundation of a new theoretical framework of ‘successful adoption’ of ERP.

While unique in its design and perspective, the study has several limitations as well. Firstly, it only examines a limited number of antecedents pertaining to ERP adoption stage. As such there is a need for future research in testing more factors in getting an agreeable set of CSFs. Secondly, due to its geographical focus in Australian context, the utility of study’s findings in the cross-border and cross-cultural ERP projects context would be limited. Finally, the study provides a snap shot understanding of the adoption stage, and thus do not takes into account the dynamic nature of changes in effect of CSFs to improvements in performance and on the innovation process generally.

REFERENCES


Hong, W., and Zhu, K. 2006. Migrating to Internet-Based E-Commerce: Factors Affecting E-Commerce Adoption and Migration at the Firm Level. Information & Mgt. 43(2), 204-221.


The "fear Factor" in Critical Care Tele-Pediatrics

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0069.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Telemedicine, E-health, Case Study, Health information systems / medical record systems / care delivery /</td>
</tr>
</tbody>
</table>
THE “FEAR FACTOR” IN CRITICAL CARE TELE-PEDIATRICS

Gogan, Janis L., Bentley University, 175 Forest Street, Waltham MA, USA, 
jgogan@bentley.edu

Garfield, Monica J., Bentley University, 175 Forest Street, Waltham MA, USA, 
mgarfield@bentley.edu

Baxter, Ryan, Bentley University, 175 Forest Street, Waltham, MA, USA, 
rbaxter@bentley.edu

Abstract

How do emotions affect telemedicine adoption and usage for critical-care episodes? We report on early findings from a grounded theory approach in a multiple-case study of critical care tele-pediatrics (research in progress). Thus far our findings suggest that specialists believe that generalist clinicians at remote “spoke” hospitals are anxious and fearful when performing in a high stress crisis situation involving an acutely ill or gravely injured child, leading to tunneling of attention (overlooking important clinical details). The specialists at tertiary care “hub” hospitals feel they are able to keep an emotional distance from the situation and they also report feeling less anxious because they have had extensive training for these acute care situations. Both emotional distance and higher skill level seem to help the specialists to take in more clinically-relevant information and use this to guide the generalists. The generalists’ fear of negative evaluation by the specialists, or embarrassment in front of other clinicians, may have some impact on their decision to use tele-pediatrics, but does not seem to be a salient concern during critical care episodes. These findings suggest that emotions do play a role in telemedicine use and possibly also in other forms of technology-mediated inter-organizational collaboration.

Keywords: telemedicine, anxiety, stress, emotion

1 INTRODUCTION

The emotional side of technology-mediated inter-organizational collaboration has received little attention in the IS literature. Studies in other fields (reviewed below) find that stress leads to anxiety which in some situations will impede learning and in other situations improve learning. Stress may also increase or decrease the likelihood that individuals will seek knowledge from others. Under the stress of extreme time pressure people seek less information and performance degrades.

In critical-care pediatrics a child’s life hangs in the balance. A stressful situation such as this can elicit emotions such as fear or anxiety among clinicians who are not trained in pediatric critical care. Telemedicine can connect these clinicians to critical-care specialists located in tertiary-care teaching hospitals. In this paper we focus on the specialists who use telemedicine to provide assistance to physicians located elsewhere who are caring for acutely ill or critically injured infants and children (critical care tele-pediatrics). We report on early findings from case studies of two “hub” hospitals that provide pediatric critical-care consultation services via telemedicine. Conversations with pediatric intensivists (pediatricians who received additional post-residency certification in pediatric critical care) reveal that fear, anxiety and other emotions do apparently play important roles in both the acceptance and ongoing use of telemedicine in this domain.
This paper is organized as follows: We discuss prior research addressing the influence of stress on learning, information seeking, decision-making and collaboration. We describe our methodology, then discuss how the “fear factor” played out among several stakeholder groups – specialist experts at hub hospitals, generalist physicians and other clinicians at spoke hospitals, and hospital administrators. We discuss these findings and suggest opportunities for further research on the role of emotions in computer-mediated inter-organizational collaboration under time pressure.

2 EFFECTS OF STRESS, ANXIETY AND FEAR ON LEARNING

Human stress is a physiologic and emotional response to a stimulus (the stressor) (Gaillard, 2001). We experience stress and related emotions as follows: A triggering event or stimulus (such as a loud noise) precipitates an immediate pre-attentional orienting reflex (Rohrbaugh, 1984). Next the brain attends to key features of the precipitating event and matches them to past experiences, evoking a previously learned response. A state of arousal, initially experienced as uncomfortable, is subsequently experienced as a more refined emotional cognition. Unexpected stressors cause individuals to focus on their own emotional response, and less on the stressor itself. (Chajut and Algom, 2003; Zohar and Brandt, 2002). When a stimulus is perceived as threatening, the individual is more likely to use selective attention (Baddeley, 1972), attending to a subset of available information. “Under stress, attention appears to channel or tunnel, reducing focus on peripheral information and tasks and centralizing focus on main tasks … tunneling of attention can result in either enhanced performance or reduced performance, depending on the nature of the task and the situation. … When peripheral cues are irrelevant to task completion the ability to tune them out is likely to improve performance. On the other hand, when these peripheral cues are related to the task and their incorporation would otherwise facilitate success on the task, performance suffers when they are unattended” (Staal, 2004, p. 31).

Also, when anxiety impedes one’s ability to process relevant information, this itself can increase stress. Thus, anxiety is both a response to stress and a cause of it.

Through attentional processes a negative feeling is experienced, and through cognitive processes it is interpreted as an emotion such as anxiety or fear. Psychologists place situational anxiety and fear on a continuum; not separate emotions but different manifestations of the same emotional state. Anxiety affects one’s ability to learn new material or to apply previously-learned knowledge (Taris and Feij, 2004; Warr and Downing, 2000), but (consistent with the tunneling-of-attention findings above) the nature of that impact varies. One theory proposes that anxiety leads employees to seek help from people who they believe possess useful knowledge or skills. Lower performing employees reportedly engage in such interpersonal help seeking more than high-performing employees (Warr and Downing, 2000) and individuals are less likely to seek help if they fear learning bad news (Sheikh, 1998).

Time pressure is a stressor that can affect both emotions and performance. Novices benefit from being provided meta-information that describes the accuracy, timeliness and completeness of specific data sets (Fisher, et al, 2003), but in real life such meta-information is often not available (Ballou and Pazer, 1995; Cuppiello, et al., 2003). Under extreme time pressure decision makers may seek or process less information (Pearson and Clair, 1989) and perform worse (Entin and Srfaty; 1990; Van Galen and van Huygevoort, 2000; and Wickens, et al. 1991). Extensive training and rehearsal improves one’s ability to work effectively under high time pressure (Sniezek, et al., 2002) and to attend to the appropriate cues (Lerch and Harter, 2001). The theory of recognition primed decision making posits that when confronted with a new situation, an individual quickly assesses its key features, then recalls a similar previous situation and responds in ways similar to how he acted previously (Lipshitz, et al., 2001). Under time pressure, novices tend to focus on a single comparative situation and set of behaviors, even though a better outcome might be attained by altering the course of action to better meet the unique needs of the current situation (Flin, et al., 1997). Experts are better able to match salient features of a situation against appropriate prior experiences, while novices are more likely to focus on the wrong aspects (Kaempf, et al., 1996; Klein, 1993; Lipshitz, 2001).
While much research has studied the effects of prolonged exposure to chronic stress, we focus in this paper on technology mediated collaboration under acute (situational) stress.

3 TELE-PEDIATRICS

In the US and elsewhere there is a shortage of pediatric intensivists (doctors with advanced post-residency certification in critical-care pediatrics). Indeed, this field is young; following the establishment of pediatric intensive care units in several influential hospitals in the sixties and seventies, the first post-residency fellowship programs in pediatric critical care were offered in the eighties by hospitals in Toronto, Philadelphia, Washington DC, Baltimore, Boston, Dallas and Detroit (Epstein and Brill, 2005; Randolph, et al., 2004). As of 2009, 49 US hospitals offer pediatric critical care fellowship programs to just 116 candidates, according to the National Resident Matching Program (www.nrmp.org). Most fellowship programs admit just one or two fellows per year, and once their training is completed, most pediatric intensivists remain in urban centers; only 3% practice in rural areas. This is a problem since about 21% of U.S. children live in rural areas (Heath, 2008).

Fortunately it is a rare occurrence for physicians (especially in rural locations) to see critically-ill children; common conditions such as ear infections or minor injuries are far more likely. Unfortunately, when a child suffering an acute illness or critical injury is brought in, rural clinicians may not have all the necessary skills to care for them. At a community hospital a critically ill child might be seen by emergency room clinicians who are accustomed to treating adults, or by pediatricians who have little experience providing acute emergency care. The child may require immediate attention and stabilization before being transferred to a pediatric emergency department or intensive care unit at another hospital. So, some clinicians turn to specialists at teaching hospitals for help. Traditionally, that help is provided via brief telephone consultations, but today a few teaching hospitals offer telemedicine-based consultation services. Consistent with Thrall (2007) we define telemedicine as “the delivery of health care services over a distance.” In the current paper we focus on telemedicine-supported real-time consultations by experts working online to assist physicians located elsewhere, to care for acutely ill or critically injured infants and children (i.e., critical care tele-pediatrics).

Telemedicine is not new (early applications of it began in the 1960’s) and its scope is broad. Some telemedicine applications entail store-and-forward transmission of video, still images, documents, and graphical data while others utilize real time video conferencing. Numerous studies have led to a consensus among physicians that telemedicine services are safe and efficacious (Hersh, et al., 2001; Taylor, 2005) across a broad range of medical specialties. The top ten telemedicine domains (per the Telemedicine Information Exchange, which tracks more than 200 active programs worldwide) are: mental health, cardiology, dermatology, pediatrics, radiology, home health, orthopedics, neurology, oncology, and general surgery (Tulu, et al. 2007). While pediatrics is included in the above list, critical-care tele-pediatrics is not widespread. Fewer than six “hub” hospitals currently provide this service in North America, and the number of participating “spoke” hospitals is also quite low. One reason (as noted above) might be that the number of critical-care episodes is not very high and thus critical care tele-pediatrics is just not yet high enough on the priority list for health care administrators. Another reason might be that the supply of specially trained pediatric intensivists is not yet high enough to support this activity. Or, there may not be enough champions promoting this innovation. To date, studies in tele-pediatrics have largely focused on evaluating the technologies and establishing the appropriate networked applications for diagnosing and caring for infants and children (Marcin and Ellis, 2004; McConnochie, et al., 2006; Mehta, et al., 2001; Sable, et al., 1999).

Prior research hints that fear and anxiety may impact telemedicine use. For instance, a state hospital physician reportedly participated in telemedicine consultations in order to avoid dangerous exposure to patients with certain conditions (Paul and McDaniel, 2004, p. 200): “Would you like to get in the ambulance or the life-flight helicopter with somebody with a very contagious disease?” Yet, the impact of emotions on telemedicine use is largely unexplored.
Our research was conducted at two tertiary-care teaching hospitals that provide critical-care tele pediatric services. Our purpose is to explore the emotional side of tele-pediatrics. Little is known about the role that emotions play in influencing the behavior of the specialists and generalists on either side of the urgent telemedicine consultation. Thus, our research questions:

- Do emotions influence whether tele-pediatrics is used for critical care consultations?
- How do emotions influence what happens during these consultations?

4 METHODOLOGY

Our research is part of a larger study of telemedicine challenges in several clinical domains. In this paper we draw insights from interviews of 60 to 90 minutes each, conducted with three specialists in pediatric critical care (hereafter “intensivists”) at two teaching hospitals (UrbanHub and RuralHub). We interviewed one of the three pediatric intensivists on staff at RuralHub and two of the pediatric intensivists on staff at UrbanHub. We also conducted interviews with several administrators and other clinicians at each institution, to learn about other telemedicine services and to explore business, legal, technical and other challenges associated with them.

This study is ongoing and we are using an inductive, grounded theory approach, the constant-comparative analysis, and snowball sampling (Stake, 1995; Strauss and Corbin, 1998). Our open-ended interviews explore challenges and issues in pediatric critical care in general and in offering tele-pediatric consultation services. Interviews are recorded and interviewers also take field notes. In the first phase we interviewed specialists at two hub hospitals, and in spring 2009 interviews are being conducted with generalists at participating spoke hospitals. After interviews are professionally transcribed, we review them and add corrections and contextual notes to the transcripts. All members of the research team then code the transcripts, using four forms of coding: 1) Factual coding captures key events and facts (such as the cost of telemedicine equipment and number of pediatric intensivists on staff). Facts and viewpoints obtained in interviews are triangulated against information available in publicly available documents such as web pages, presentation materials, planning reports, news accounts and journal publications. 2) Comparative coding captures findings that are or are not consistent with previously-identified themes from earlier rounds. Our initial comparative coding entailed comparing interview segments with themes identified in earlier telemedicine studies conducted by two of the authors (such as the influence tactics of a telemedicine champion and concerns about licensure, credentialing, and reimbursement), and with findings identified in our literature review of other telemedicine studies. 3) Open coding captures new themes or sub-themes revealed in the interviews (such as in this study a theme about clinician fear which we had not anticipated). 4) Interpretation examines the relationships among themes and sub-themes to reveal the significance of events, actions or viewpoints (Stake 1995).

After each researcher separately codes the transcripts we meet to identify areas of overlap or difference in the coded themes and consolidate them into broader themes or break them out into narrower sub-themes. Each author then revisits the transcripts again with these themes and sub-themes in mind to substantiate or refute them and to identify additional informants and research sites.

In this paper we offer preliminary evidence of how stress, fear, anxiety, and telemedicine usage play out in the context of critical care pediatrics. In this research in progress the themes thus far reflect three experts’ views about critical care tele-pediatrics, but do not yet directly reflect the generalists’ views (although we expect to be able to discuss their perspective at the ECIS meeting in June 2009).

We next provide a brief background of the two hospitals and their critical care tele-pediatric services.

4.1 Case 1: UrbanHub

UrbanHub is a division of a well known urban teaching hospital. UrbanHub’s 14-bed pediatric critical care unit is considered a leader in this specialty, with intense competition for spots in its pediatric...
critical care fellowship program. The parent hospital has experimented with telemedicine in a variety of clinical areas, a few of which – such as its tele-stroke service – are now well established.

In the late nineties the UrbanHub pediatric critical care unit launched a grant-funded pilot test of tele-pediatrics with an affiliated suburban hospital. A video camera was set up in a treatment room at the spoke hospital’s emergency department. From the pediatric critical care unit at UrbanHub a consulting physician could closely examine the patient, see clinical information provided on various monitors, and observe the clinicians at work. More recently UrbanHub began offering the service to another suburban hospital and (thanks to private funding from a donor with ties to Puerto Rico) to a hospital in rural Puerto Rico. The department head would like to expand the service but reports that a financially viable business model for doing so has not yet been established.

4.2 Case 2: RuralHub

RuralHub, in a state with a very rural population, is the only tertiary care teaching hospital in its state. Its catchment area covers about one million people. Most residents of this state live in towns with populations less than 2500, and its mountainous terrain makes travel to RuralHub time-consuming and challenging, especially in winter.

In 1994 RuralHub had launched a tele-pathology pilot study, involving two remote hospitals. In the nineties pathology was the largest telemedicine user, followed by vascular surgery and nephrology. Some path-breaking telemedicine examples included remote endoscopy exams and a few instances in which telemedicine was used during live surgical procedures. Pediatrics was not a regular user of telemedicine for physician-to-physician consultations at that time. RuralHub began its critical care tele-pediatrics program in the early years of this decade. A RuralHub pediatrician had trained at the prominent UrbanHub pediatric critical care program in 1998, just as UrbanHub was starting to use telemedicine in this domain. When this physician returned to RuralHub after completing his fellowship, he began to offer tele-pediatric consultation services (funded by a federal grant) and to build a department of pediatric critical care, which now has three intensivists.

Before launching a telemedicine service the intensivists visited ten rural emergency departments to help them learn how to comply with new guidelines for pediatric critical care and advanced life support. The field of pediatric critical care was evolving rapidly, however, so there was still a concern that rural clinicians’ skills might erode in comparison with best practices in just a few years. The RuralHub intensivists reported that since pediatric emergencies are rare, rural clinicians who underwent training might nevertheless be poorly prepared to provide critical care when needed. Motivated by these concerns, more federal funding was sought to set up a pilot test of a critical care pediatric telemedicine service involving ten rural emergency departments. As was the case at UrbanHub, intensivists provided 24/7 coverage. This pilot test demonstrated the feasibility of using telemedicine for critical care consultations to rural clinicians, and the RuralHub pediatric team continues to offer this service, even though they are not yet being reimbursed for their consultations.

5 EMOTIONS AND CRITICAL CARE TELE-PEDIATRICS

The three specialists at the hub hospitals were unanimous in the belief that clinicians at spoke hospitals are eager to transfer critically-ill children to tertiary-care facilities, because they are anxious about caring for these patients. A RuralHub intensivist discusses this:

“*These are low frequency, high morbidity events. ...The level of comfort can become extremely low for someone who doesn’t see children very often and is presented with a tiny baby with serious health problems; it’s very difficult for them. ...The goal for that provider is to make this problem go away as fast as you possibly can. So they don’t hesitate to call us... None of the pediatricians, zero, want to take care of these children; they want them gone. Emergency department physicians feel the same way, because they don’t have the technology to care for these children. ... When physicians, nurses, respiratory therapists on the other end are clearly scared to death and we walk them through how to*
do absolutely everything and the transport team comes and the kid leaves, they are so happy with the service that we have given them”

Another hub doctor presumed that a physician at a spoke hospital was anxious out of concern about being observed by the experts in the new telemedicine consultation service:

“It took us a while for the person there [at the spoke hospital] to rely on us and feel confident...(It) is a very sensitive relationship. At first, physicians are afraid we are seeing what they are doing wrong; they are concerned about that. Like, ‘Oh my God, these guys are going to see us, we’re not doing well.’ So we have to overcome this barrier. Once they feel that we are friends and that we’re working together and we’re partners (then it is fine). The new person over there is now learning gradually. She prefers to call us over the phone more than talk with telemedicine. We tell her, ‘Turn [the video camera] on.’ She says, ‘Oh no it’s just a quick question on the phone.’ [Laughs] It takes some time.”

Before telemedicine, anxious clinicians at spoke hospitals in the rural state would place a telephone call to request that a critically ill or injured child be transferred to RuralHub, and the intensivist would attempt to determine whether the patient was stable enough for the transfer to take place. However, the intensivist found it difficult to offer advice based on clinicians’ verbal descriptions of what was going on. Being able to see the child, the various monitor displays, and the clinicians at work via video is a huge plus, in his view. He described one particularly vivid case:

“I turn on the telemedicine and can see them doing chest compressions on this kid. I say, ‘You need to slow down a little bit and push harder, make the chest move more.’ Then I see they have a breathing tube in and they’re squeezing the bag and I say ‘Slow down, squeeze harder and longer.’ I could see the cardigram and monitors and can make recommendations. ‘You put the breathing tube way down too far in the lung.’ That happens all the time because they are anxious and they push it in too far. I say ‘Pull the endotrachial tube back and you will ventilate better.’ I could see the patient had a type of lung injury called non-cardiac pulmonary edema. I could tell the respiratory therapist, who had not had experience using ventilators with children in the past 15 or 20 years, exactly what dials to push on the ventilators and what settings to do exactly what we wanted. I told them the resuscitation drugs and the doses to use, because they may remember what drugs to use but they can’t remember how to use them or how much to use. The child was placed on mechanical ventilation (and transferred to RuralHub). I sent the transport team out and directly supervised them; they talk to me over telemedicine, I told them exactly what to do.”

A RuralHub intensivist felt he was able to “see the big picture” better than clinicians at the spokes:

“...When they’re intubating, everyone’s field of vision focuses down, very, very narrowly on the procedure they’re doing. I can sit up there, move the camera around, pan in and out and I can say, ‘You know something? The level of oxygen in his blood has dropped to a very dangerous level and you need to stop doing what you’re doing right now and fix that and then try your intubation again.’ They don’t hear the alarms and they don’t see the stuff because they’re busy doing other things. And there is just the question of familiarity. I do this for a living; it doesn’t scare me anymore. And we can provide that sort of experience and expertise.”

An UrbanHub intensivist also commented on his ability to retain some emotional distance:

“We have our remote control from here that we can use to control three cameras that they have over there. We can see all the monitoring equipment, see the patient, see the x-rays, see the EKG’s and all the other stuff that we need. … We supervise them, and move the camera wherever we want … The system surprised us, how well it works. We really can see a lot. … We could see the pupils of the patient’s eyes. Clinicians on the ground … are so stressed caring for the children that they don’t pay attention to things that we can see since we are less emotionally involved. We can see what is happening on the whole … We can point out that a valve is not functioning too well. For example, a cardiac arrhythmia case: the ventilator connected to the child was not working so well; oxygen was not coming. Even before they had time to get that information on the monitor, even before alarms went off we could see that the chest was not moving well and we could tell them what was wrong. Another
time, they were trying to intubate but the child was nauseous and about to vomit. There again they are so anxious trying to maintain the airways to make the kid breathe. If you say, ‘Hey wait a minute, turn the kid to the side’ to avoid aspirating. It’s amazing sometimes by not being over there (we can be more effective). …We really can see a lot. The clinicians on the ground, in the field are so stressed caring for the children that they don’t pay attention to things that we can see, since we are less emotionally involved. We can see what is happening on the whole at that point.”

One hub intensivist states that what the spoke clinician needs is “having another set of eyes consulting them, helping them to make the decisions.” Another expert offers a similar comment about the importance of his role in reassuring the spoke clinicians that they are making the right decisions. In the interview, we asked whether there were circumstances in which he might say to a clinician at the spoke, “You are doing the right thing.” He replied: “There is a lot of that. ‘Yes, I agree absolutely with what you’re doing.’ These are the – yes, yes, yes – ‘I suggest doing this and you’ve already done that – we’ll send a transport team out.’”

Sometimes the clinicians at the spokes just need assurance that a specialist is present. In one episode a specialist suggested that the two doctors work together: “I said ‘You know, why don’t we observe this kid together, give him a couple of nebs (nebulizers) one after the other and let’s observe him in twenty minutes again. We’ll see how the kid responds.”

Many spoke hospital clinicians want to transfer patients immediately; however this can be risky. A Boston physician contrasts what might have happened had the transfer details been done over the phone versus the actual outcome in which they were able to use full video:

“…[without video] it would be a description over the phone. ‘We have this 17 years old that’s doing fine, having some mild headaches, having a little fever, we don’t know what it’s going to involve and want to transfer the kid. We have called the ambulance.’ I said, ‘wait a minute, let me look at the kid. We need a critical care ambulance, we need a critical care doctor, we need her to be ready for the transfer. …If she had been transferred without us seeing her it could be a disaster en route. The old fashioned way that most of the hospitals do over the phone, they never described her as sick as she was; when we looked at her and said, ‘This is a sick patient, we cannot put her on the ambulance and just send her, especially now during the traffic hours. We have to stabilize her over there as well.’ And as we are talking, within 10 minutes the patient collapsed over there. So at least they were prepared.”

Thus, the specialists that we interviewed at two hub hospitals contended that a) spoke clinicians are more anxious than experts because they don’t have the same level of skill; b) some spoke clinicians are, at least initially, anxious about being observed by hub experts; c) hub pediatric intensivists feel that when they use telemedicine they achieve a degree of emotional distance that allows them to take in more relevant information and thus more effectively guide and reassure the generalists.

6 DISCUSSION

From interviews with pediatric intensivists at two “hub” hospitals we uncovered evidence about the influence of emotions on telemedicine adoption and use. Anxiety – whether about using a new technology or out of fear of being exposed as incompetent – apparently led some clinicians to choose a more familiar technology, the telephone, when requesting help from hub specialists. This illustrates the familiar pattern of anxiety impeding new system adoption. Yet, some spoke clinicians did use telemedicine when confronted with critical care cases for which they had little prior experience and inadequate skills, and the specialists who provided help believe that their guidance helped the spoke clinicians provide better care under extreme stress.

We further learned that the intensivists believed that telemedicine extended their expert capacity to take in a great variety of clinical information and to focus on important details that escaped the attention of anxious clinicians at the spokes (because presumably stress caused tunneling of the spoke clinicians’ attention). Telemedicine also allowed the experts to achieve some emotional distance; they reported feeling less emotionally involved and better able to see the “big picture,” remain calm and
effectively guide the spoke clinicians. These early findings suggest that while stress can cause tunneling of attention for those clinicians who have less specialized and/or less practiced skills, the combination of emotional distance and high-level expertise allows medical experts to take note of important clinical details. This is an important finding that warrants further investigation in other time-pressed clinical situations (such as trauma care, ambulance care, and complex surgical care). Most observers contend that telemedicine provides the “next best” alternative to having an expert physically present in the room to guide the less-expert clinician. However, our finding about emotional distance suggests a more intriguing possibility. The experts that we interviewed indicated that telemedicine gave them a degree of emotional distance that they found beneficial, suggesting that in some circumstances telemedicine might be seen as superior to face-to-face consultation.

It is important to note that all physicians are trained to perform under stress in a variety of clinical situations. However (as mentioned in the introduction to this paper) pediatric critical-care situations are rare occurrences. Thus, even during their internships and residencies, pediatricians or emergency physicians may get few opportunities to learn how to provide effective care in these situations. Post-residency, clinicians at remote spoke hospitals will get especially few opportunities to participate in critical-care events involving young children, and also most of these remote hospitals do not have pediatric intensivists (with their higher level of post-residency training) on staff and available to coach other clinicians in these specialized techniques. With fewer than 120 physicians entering pediatric critical care fellowships each year, this situation is likely to persist for quite some time. While the numbers of board-certified intensivists on staff at tertiary-care centers and perhaps at suburban hospitals will gradually increase, there will continue to be a physicians with this specialized skill set at rural hospitals and clinics. And, with pressure to control costs and with the low frequency of pediatric critical-care episodes, most hospital administrators will choose to allocate scarce resources elsewhere.

Thus, when a critically ill or injured child does arrive at a rural clinic or hospital, the attending clinicians may not have all the knowledge and skills necessary to provide the best possible care and this shortcoming may increase their experience of stress, anxiety, and/or fear. Our interviews revealed that telemedicine can be helpful in these situations. The intensivists that we interviewed explained that specific system capabilities -- such as real-time video transmission and the ability to remotely control video cameras -- were important technical features that helped to reveal important clinical information to them and improved their ability to guide the clinicians at the spokes.

Because pediatric critical-care episodes are rare occurrences, the skills, knowledge, and confidence of even well trained clinicians will tend to erode due to lack of practice. To combat this issue, experts at the hub hospitals offered refresher training (sometimes delivered via the telemedicine tools) to help generalists keep their skills current. One specialist at UrbanHub felt that these training sessions actually reduced the number of requests for telemedicine system for consultations, because (he believed) the spoke clinicians now feel better prepared to handle pediatric emergencies and thus don’t feel as anxious when these emergencies arise.

In our study clinical experts freely shared their knowledge via telemedicine consultations, with little concern that this activity would reduce their power or professional standing. Physicians in a teaching hospital are, of course, expected to teach, and a telemedicine consultation is a form of education. Nevertheless, it was interesting to observe this apparent beneficence, since in contrast, some prior knowledge management studies (e.g., Cabrera and Cabrera, 2002) report that business experts sometimes decline to contribute to knowledge repositories, presumably out of anxiety or fear (of losing status or being deemed expendable). This finding also deserves further exploration to uncover specific aspects of hospitals’ work climates and incentive schemes that facilitate “unselfish” knowledge sharing. Perhaps organizations in other industries can draw some important knowledge management lessons from health care.
LIMITATIONS AND CONCLUSIONS

As a research in progress, to date our findings are based primarily on interviews with three key informants at two hub hospitals (interviews with other administrative and technical employees provided valuable background information but did not speak to this particular issue). In spring 2009 we are scheduled to interview clinicians at the spoke hospitals that these two hub hospitals serve. Until then, we will not be in a position to speculate as to whether generalists really were motivated to request telemedicine consultations out of fear/anxiety versus out of a dispassionate assessment of the skills that are missing from their repertoire, or some other reason. One vignette hinted that some spoke clinicians preferred to use a less transparent communication technology (the telephone) compared with video-conferencing, which might put this clinician’s mistakes in full view, causing embarrassment or further stress. A hub specialist believes that once trust is established, such concerns will fade into the background. This view is consistent with Dunn and Schweitzer’s (2005) behavioral experiments in which they find that emotions do not influence trusting behaviors when the trustee is someone familiar. Other studies have explored issues surrounding emotions, affect, and trust (e.g., Williams 2001, Weber, Malhotra, and Nurninghan 2005).

Our study findings thus far only hint at circumstances that may impede the use of telemedicine for pediatric critical care, and further study is needed to attain a clearer picture of the facts that impede or encourage adoption by the spoke clinicians. Also, our informants thus far have not brought up any instances when they felt that telemedicine impeded their ability to provide care or had adverse effects, but that does not mean that we can conclude that there have been no problems. Nor at this time are we able to conform whether the infrequent occurrence of critical care episodes is strongly or weakly related to participants’ experience of anxiety or fear, although the comments that we heard from our three informants at the two teaching hospitals were strikingly similar. At the June ECIS meeting we should be able to speak to these issues. Further in-depth interviews are needed to tease out a clearer picture of both positive and negative aspects of telemedicine use under time pressure.

Early findings from our research in progress point to fascinating issues concerning how emotions affect technology-mediated collaboration across organizational boundaries in risky, time-pressed situations. Several vignettes reveal specialists’ beliefs that fear motivates generalists to seek their help. However, we cannot as yet say whether and how specific features of the technology artifact affected this collaboration. In learning that hub specialists believe that generalists are motivated to participate in telemedicine services because of fear or anxiety about their ability to provide the best care to the critically-ill young patients in their charge, we wondered what actually goes on in the generalist’s mind. We do not know if these participants would feel more or less comfortable if the specialist were in the room with them. We do know that most rural hospitals will not have this choice (if they want their clinicians to have access to the guidance of specialists during critical-care episodes, they will need to use telemedicine).

At one of the hub hospitals that we visited, an administrator stressed the importance of having their hospital personnel regularly test the system and its network connectivity in order to ensure that spoke clinicians will be able to use the system when it is most needed. This individual observed that while clinicians are accustomed to employing a great variety of improvisational skills when providing medical care, many clinicians are taken aback when confronted with a piece of video conferencing equipment that is not operating correctly. It may just be a matter of noticing that a plug has come loose or a setting changed, but the clinicians will not know what to do to get the system working correctly. Clearly, while some clinicians are comfortable with a variety of information and communication technologies, others are less so. If they perceive that telemedicine is their only avenue available to get the expert guidance they need, then perhaps they will embrace it even if they normally shy away from using leading-edge technologies. In that case it can be helpful to have nearby technical support personnel on call, and it also may be helpful to give the hub specialist control over the cameras, able to zoom in and out on his/her own. Our reasoning here is that the hub specialist is under less stress
during a critical-care consultation and may be in a better position to perform these tasks than if the harried clinicians at the spoke site are expected to do this.

Our interviews revealed that the experts value their ability to use remote video to take in many more clinical details than are available in telephone consultations, as well as their ability to control the video from their end. This, combined with the findings about emotional distance suggests that system trust (trusting beliefs about ICT applications) interacts with emotions in ICT-enabled collaboration. In future studies we suggest exploring how emotions such as anxiety and fear interact with organizational trust, interpersonal trust, and system trust. For example, further research is needed to investigate specific technical features used by clinicians on both sides of telemedicine consultations (for chronic versus urgent care), and to closely compare how experts coach novices when using ICT versus in face-to-face situations.

This study explored specialists’ views on the effect of fear/anxiety on telemedicine usage in a single high-stress domain (critical-care pediatrics), characterized by infrequent episodes and extreme time pressure. Our interviewees emphasized that critical-care episodes occur infrequently in pediatrics, and that ongoing training is thus necessary to ensure that spoke clinicians will initiate telemedicine consultations in those situations when the hub experts can be most helpful. Further research is needed in order to determine whether these findings about emotional aspects of ICT-enabled collaboration would apply in other high-stress domains in which infrequent episodes and time pressure are at play, and to tease out the separate and interacting influences of frequency, time pressure and risk.

These are early days for critical-care telemedicine, especially in pediatrics; we are eager to learn more about this promising use of information and communications technologies to bring needed expertise to underserved populations in this and other clinical domains.

8 ACKNOWLEDGEMENTS

We gratefully acknowledge the assistance of the unnamed individuals who we interviewed at RuralHub and UrbanHub, who took time out of their busy clinical schedule to share their thoughts. We thank also the McDowell Research Center of the University of North Carolina at Greensboro, who provided funding that partially supported this study, and Bentley University which provided additional financial support.

9 REFERENCES


Proceedings ECIS 2009


CONSUMERS’ ACCEPTANCE AND USE OF PERSONAL HEALTH RECORD SYSTEMS: A THEORETICAL MODEL

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0245.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>Adoption, E-health, Health information systems / medical record systems / care delivery /, Human computer interaction (HCI)</td>
</tr>
</tbody>
</table>
CONSUMERS’ ACCEPTANCE AND USE OF PERSONAL HEALTH RECORD SYSTEMS: A THEORETICAL MODEL

Assadi, Vahid, McMaster University, DeGroote School of Business, 1280 Main Street West, Hamilton, Ontario, Canada, L8S 4M4, assadiv@mcmaster.ca

Hassanein, Khaled, McMaster University, DeGroote School of Business, 1280 Main Street West, Hamilton, Ontario, Canada, L8S 4M4, hassank@mcmaster.ca

Abstract

Recently, there has been a growing trend towards consumer-based healthcare in which consumers are increasingly becoming partners in their own care. One way of accomplishing this is to provide consumers with access to their health records through the use of Personal Health Record (PHR) systems. In spite of their potential benefits, recent research has shown that PHRs are not yet popular or well known to consumers. The overall objective of this research is to investigate the influences of various personal, behavioral, and environmental factors on the adoption and use of PHR systems by Canadian consumers. Drawing on both the information systems and behavioral healthcare literatures such a model is developed and presented. The proposed model will be validated using a longitudinal design over a period of 16 months involving patients from two local clinics. The study participants will be introduced to an existing PHR system at those clinics. The system will subsequently be made available for their potential use. Users will be surveyed at various points in time regarding their perceptions about the system utilizing both close-ended and open-ended questions. Collected data will be analyzed using structure equation modeling and qualitative data analysis techniques.

Keywords: e-Health, Personal Health Record system, Technology Adoption, Social Cognitive Theory, Longitudinal Study
1. INTRODUCTION

Two important trends can be observed in the Canadian healthcare system (Urowitz et al. 2008): the advent of e-Health giving rise to a more important role for information technologies in healthcare (Eysenbach 2001; Tan 2005); and a shift towards consumer-based healthcare (Eysenbach and Diepgen 2001; Runy 2000) where patients are considered as partners in their own care process (Urowitz et al. 2008). For example, today’s educated and computer literate baby boomers who make up almost one in every three Canadians (Folker 2007) are facing health-related conditions as they age and are increasingly seeking health-related information from various sources including the Internet (Bliemel and Hassanein 2007). Providing access to personal health information through innovative technologies could potentially reduce the cost and complexity of healthcare delivery through efficient use of resources in the healthcare system (Raghupathi and Tan 2002). One such innovative technology is the use of Personal Health Record (PHR) Systems.

A PHR can take various forms including a stand-alone application or an Internet-based system (Endsley et al. 2006). PHRs are created, owned, updated, and controlled by an individual and/or others authorized by her/him. They contain a summary of a consumer’s lifelong health information such as allergies, home monitoring data (e.g. blood pressure), medications, laboratory test results, conditions suffered, treatments given, vaccinations, etc (Thomas 2006). Numerous benefits have been suggested for consumers utilizing PHRs. For example, they can access a wide range of reliable and credible health information leveraging this access to increase their understanding of their health condition and to be active participants in their own care (Cimino et al. 2002; Moehr and Grant 2000; Ueckert et al. 2002). PHRs put consumers in control of their own health information by allowing them to update their records either manually or by automated polling of information from visited care facilities (e.g. hospitals, physician offices) (Tang et al. 2006). By leveraging the control and access provided by PHRs, consumers could become empowered to better manage their health (Tang et al. 2006). For example, they could in collaboration with their physicians, detect disease in the early stages by observing trends in their health status (e.g. changes in blood pressure). They can also consult with their physicians on any usual conditions observed in their health records. Moreover, the system can alert people when their health records show such unusual conditions or exceptions (Tang et al. 2006) (e.g. a conflict between newly prescribed medications and previously or currently used ones). It is important to note that PHR consumers are not necessarily dealing with immediate medical concerns and can be ill or healthy.

PHR systems are also suggested to be beneficial for patients with chronic diseases (Heubusch 2007). Chronic diseases are often characterized by long latency requiring patients to be continuously aware of their condition in an ongoing collaboration with their caregivers (Folker 2007; Heubusch 2007). PHRs can facilitate patient-physician communications in an efficient manner through changing such communications from episodic encounters to continuous interaction (Tang et al. 2006). Furthermore, self-management activities and active patient participation in the care process are major parts of a successful chronic disease management program (Lankton and St. Luis 2005). A PHR system can facilitate such a high level of patient engagement (Tang et al. 2006).

In spite of all the aforementioned potential benefits for PHRs, recent research has shown that they are not yet popular with or well known to consumers (Sittig 2002; Cronin 2006). Very few studies have covered the reasons responsible for the lack of PHRs’ popularity. Existing studies have mostly concentrated on overviews to clarify the characteristics and functionalities of PHRs (Abrahamsen 2007; Atkinson et al. 2007; Brown 2007; Cronin 2006; Kim and Johnson 2002; Lafky et al. 2006; Sittig 2002; Tang et al. 2006; Thomas 2006; Yee and Trockman 2006). The few studies performed on the adoption of PHRs were for the most part not empirical in nature (Denton 2001; Iakovidis 1998; Jones 1999; Lafky and Horan 2008; Winkelman et al. 2005). These studies have put forth numerous factors that bring about the lack of PHR popularity. Of particular interest, Tang et al. (2006) suggest...
that behavioral factors may impact PHRs’ adoption, yet the role of such factors was not empirically tested. Hence there is a need for additional research in this area.

By developing and validating a model that explains the behavioral/social factors influencing consumers’ adoption and use of Internet-based PHR systems, this study pursues the following objectives: (i) to identify the behavioral/social factors influencing PHR system adoption and use among the consumers; (ii) to investigate the impact of individual characteristics (perceived health status, age, sex, Internet experience and education) on relationships in the model.

This paper is organized as follows: the proposed research model is presented in section 2 along with theoretical background and hypotheses. Details of the proposed research methodology are presented in section 3. Finally, the potential contributions of this research are discussed in section 4. Throughout this paper the words consumer, individual and patient are used interchangeably unless otherwise specified.

2. THEORETICAL BACKGROUND AND RESEARCH MODEL

Social Cognitive Theory (SCT) (Bandura 1986) will form the basic framework for conceptualizing a research model for this study. SCT explains human behavior as an interaction between behavioral, environmental and personal factors. SCT has been widely employed in the IS literature with demonstrated validity (e.g. Chan and Lu 2004). Specifically, this theory is a widely accepted model for explaining individual behavior in the IS area (e.g. Compeau and Higgins 1995). SCT states that a person takes an action that has personal cognition in a social environment (Bandura 1986). This notion fits well into the context of this study since individuals need to make a decision to adopt a PHR system for managing a healthy life while interacting with the medical environment through the system. The proposed model in this study (Figure 1) builds on the above SCT categories while incorporating related constructs from both the IS and healthcare literatures. The underlying theories for the proposed model are examined below.

Since PHR systems are information systems, IS-related constructs are incorporated in the proposed model. The Unified Theory of Acceptance and Use of Technology (UTAUT) aims to explain user intentions to use a new computer technology and subsequent usage behavior (Venkatesh et al. 2003). This theory was developed through a review and consolidation of the constructs of eight theories that were previously employed in the literature to explain usage behavior regarding a new computer technology. The eight underlying theories include Technology Acceptance Model (TAM) (Davis 1989), Theory of Reasoned Action (TRA) (Fishbein and Ajzen 1975), Theory of Planned Behavior (TPB) (Ajzen 1985), Motivational Model (Davis et al. 1992), combined TPB and TAM (Taylor and Todd 1995), model of personal computer utilization (Thompson et al. 1991), innovation diffusion theory (Rogers 1995) and social cognitive theory extended to personal computer usage (Compeau and Higgins 1995). UTAUT holds that three key constructs, namely performance expectancy, effort expectancy, and social influence are direct determinants of usage intention (Venkatesh et. al. 2003). Moreover, gender, age, experience, and voluntariness of use are shown to moderate the effect of the three key constructs on usage intention (Venkatesh et. al. 2003).

Due to the healthcare context of PHRs, there is also a need to consider relevant theories from the healthcare literature. Lee et al. (2007) argue that the design of PHR systems should be informed by the Health Belief Model (Janz and Becker 1984). The Health Belief Model seeks to identify antecedents for an individual’s health behaviors which include activities undertaken by the individual for the purpose of preventing or detecting disease (Rosenstock 1966). It asserts that the likelihood of a person taking a preventive healthcare action is the outcome of his/her perceived health threat as well as the benefits/costs associated with engaging in that action. These variables are in turn influenced by demographical, social and psychological cues (Janz and Becker 1984; Jayanti and Burns 1998; Rosenstock 1966). Using PHR systems is similar to such behaviors since such systems are intended to help consumers maintain a healthy life. Hence, it is reasonable to consider using PHRs as a preventive healthcare behavior. Consequently, constructs such as subjective health knowledge and health
consciousness are added to the model from the healthcare literature and hypothesized to have a direct impact on behavioral intention. Additionally, inter-relationships between the aforementioned constructs from both the IS and healthcare literatures are incorporated in the model as appropriate. The relationships between the constructs in the model shown in Figure 1 are explained below in detail.

The main objective of this study is to identify the behavioral/social factors influencing PHR system adoption (intention to use) and usage (actual use) among consumers. Based on Fishbein and Ajzen’s (1975) definition, intention to use, in this paper, is a measure of the strength of an individual’s intention to use a PHR system as a preventive health care behavior. On the other hand, actual use is a measure of the frequency and extent of using a PHR system by users (Venkatesh et al. 2003). Prior research has shown a strong correlation between behavioral intention and actual system use (Venkatesh and Davis 2000; Venkatesh et al. 2003). Thus, we hypothesize that:

**H1:** Stronger consumer intentions to use a PHR system will positively influence her/his future usage of such a system.

**Facilitating conditions** is defined as the degree to which an individual believes that organizational and technical support is provided for using a system, and has been shown to be a direct determinant of system usage (Venkatesh et al. 2003). Within a healthcare setting, caregiver support has been suggested to have a positive impact on consumers using PHRs (Johnson and Singal 2006; Tang et al. 2006). Caregiver support is incorporated as part of the facilitating conditions in the proposed model as described in the methodology section. Therefore, the following is hypothesized:

**H2:** Facilitating conditions will have a positive influence on an individual’s actual usage of PHR systems.

**Social influence** is defined as “the degree to which an individual perceives that important others believe he/she should use the new system”, and it has been shown to be a major determinant of behavioral intention to use new technologies (Compeau and Higgins 1995; Venkatesh et al. 2003). Moreover, Janz and Becker (1984) have shown the positive impact of social influence on intentions to

![Figure 1. Proposed research model](image-url)
perform preventive healthcare behavior. Since it has been argued here that using PHR systems is a form of preventive healthcare behavior, the following is hypothesized.

**H3:** Social influence will have a positive influence on an individual’s intentions to use PHR systems.

**Effort expectancy** is defined as the degree of effort an individual believes is required for using an information system, and has been shown to negatively impact an individual’s intention to use new technologies (Thompson et al. 1991; Venkatesh et al. 2003; Wilson and Lankton 2004). In a preventive healthcare context, Janz and Becker (1984) acknowledged the impact of difficulty of taking a preventive healthcare action as a barrier to performing that specific action. Such difficulty is similar to the concept of effort expectancy. Therefore, the following is hypothesized:

**H4:** A higher effort expectancy associated with using PHR systems will negatively influence an individual’s intention to use such systems.

**Performance expectancy** is defined as the degree to which an individual believes that using an information system will help him/her attain benefits in his/her job (similar to health benefits attained through using a PHR system). Performance expectancy has been shown to be a determinant of behavioral intention to use technology (Davis 1989; Venkatesh et al. 2003). On the other hand, Jayanti and Burns (1998) have acknowledged the positive impact of the degree of ease associated with performing a preventive healthcare behavior on individual’s likeliness of performing such behaviors. Moreover, the belief of important others supporting a particular behavior has also been shown to impact an individual’s assessment of outcomes associated with that behavior (Compeau and Higgins 1995). Finally, effort expectancy has been shown to be a direct determinant of performance expectancy (Venkatesh 2000). Similarly, Davis (1993) has shown the relationship between perceived ease of use and perceived usefulness of an information system which are similar to effort expectancy and performance expectancy in the proposed model (Venkatesh et al. 2003). Thus, we hypothesize that:

**H5:** A higher performance expectancy associated with using PHR systems will positively influence an individual’s intention to use such systems.

**H6:** Social influence will have a positive impact on an individual’s performance expectancy in using PHR systems.

**H7:** A higher effort expectancy associated with using PHR systems will negatively influence an individual’s performance expectancy associated with using such systems.

Computer **self-efficacy** refers to an individual’s belief of having the capability to use computers (Compeau and Higgins 1995). This definition can be extended to the belief of having the capability to use an Internet application such as a PHR system. It has been shown that belief of important others supporting the use of technology, positively impacts an individual’s self-efficacy regarding such usage (Compeau and Higgins 1995). Moreover, self-efficacy has been shown to also have an effect on perceived ease of use which is incorporated here as effort expectancy. Individuals with higher levels of self-efficacy will perceive the system as being easier to use (Venkatesh 2000). Finally, self-efficacy was shown to have an impact on performance related outcome expectations regarding the use of a computer system (Compeau and Higgins 1995). This latter construct is incorporated in the proposed model as performance expectancy, thus, we hypothesize the following:

**H8:** Social influence will have a positive impact on an individual’s self-efficacy regarding the use of a PHR system.

**H9:** A higher level of an individual’s self-efficacy regarding the use of PHR systems will negatively influence his/her effort expectancy in using such systems.

**H10:** A higher level of an individual’s self-efficacy regarding the use of PHR systems will positively influence her/his performance expectancy in using such systems.
Subjective health knowledge is defined as an individual’s storehouse of healthcare information (Jayanti and Burns 1998), and has been shown to have a positive effect on individuals’ likelihood of performing preventive healthcare behavior (Boechner et al. 1990). Subjective measures for health knowledge are well established in the literature, and they cover general health information rather than information about specific symptoms and cures. Moreover, Jayanti and Burns (1998) argue that subjective health knowledge has a positive impact on an individual’s assessment of positive outcomes related to performing a preventive healthcare behavior. Therefore, we hypothesize that:

H11: A higher subjective health knowledge level possessed by an individual will positively influence his/her intention to use PHR systems.

H12: A higher subjective health knowledge level possessed by an individual will positively influence her/his performance expectancy associated with using PHR systems.

Health consciousness is defined as “the degree to which health concerns are integrated into a person’s daily activities” (Jayanti and Burns 1998). Prior research has shown that individuals who are health conscious are much more likely to engage in a preventive health care activity (Jayanti and Burns 1998; Kraft and Goodell, 1993; Rosenstock, 1966). Since in this research using PHR systems is considered to be a health care behavior, it is expected that individuals who are more health conscious will exhibit higher intentions to use PHR systems. Moreover, Janz and Becker (1984) showed a positive influence of social influence on health consciousness. Therefore, we hypothesize that:

H13: Individuals with a higher level of health consciousness will exhibit a greater level of intention to use PHR systems.

H14: Social influence will have a positive influence on an individual’s level of health consciousness.

3. RESEARCH METHODOLOGY

Research setting: The research model presented above will be empirically validated through a longitudinal field study involving patients using an actual Internet-based PHR system. Performing a longitudinal study will allow the investigation of factors as they evolve during the active process of adoption decision-making (Venkatesh et al 2003). The study will be conducted at two local clinics, namely Stonechurch Family Health Centre1 and the McMaster Family Medicine2, located in Hamilton, Ontario, Canada with almost 35,000 registered patients. There are currently 27 physicians, 60 residents and 30 nurses, nurse practitioners, and allied health professionals practicing at these clinics. MyOSCAR3 is an Internet-based PHR system associated with the above clinics which incorporates typical PHR functionalities including communication with healthcare team, requesting copies of records, prescription renewals, appointment requests, access to reliable health information, etc. There are currently 100 users registered with this PHR system which will be used to conduct this research. Prior to testing the hypothesized relationships, and in order to refine/enrich the proposed model, the concerns and experiences of a subset of the current users of the PHR system will be investigated through conducting two focus group sessions with 5-8 participants. Each session will be lead by a facilitator and a recorder will be available in order to capture all the responses (Basch 1987). Collected data will be analyzed in order to identify particular patterns, themes or concerns which are mentioned repeatedly by the respondents (Bender and Ewbank 1994). NVivo will be used as the software tool for analyzing the data (Richards 1999). Mingers (2001) suggests that research results will be richer and more reliable if different research methods from different paradigms are mixed together.

Pilot study: Prior to conducting the actual study, a pilot study involving 30 new users of the PHR system under study will be performed. The pilot will also have a longitudinal design which will be

---

1 http://stonechurchclinic.ca/
2 http://fammedmcmaster.ca/
3 http://www.myoscar.org
conducted over a period of two weeks. At the beginning, instructions on using various features of the system will be provided to the participants using a short online video. Then, participants will be asked to fill out a survey containing the model measures as well as questions about demographics and perceived health status. Two weeks later, log files of the PHR system will be examined to measure actual usage through objective measures based on frequency and extent of system use. Participants will be encouraged to use the system by asking them to complete specific tasks simulating realistic scenarios encountered by typical users during the pilot study to ensure the collection of enough data to achieve the objectives of the pilot study. Results from the pilot study will be used to: conduct a qualitative analysis based on the open-ended questions and to refine the model; measurement scales for the model constructs will be assessed and refined; and any potential technical issues or problems with procedures will be identified and resolved.

Main study: Prior to starting this research study, a research ethics protocol will be prepared and submitted to the McMaster Research Ethics Board which is the body responsible for reviewing and approving all research studies involving human subjects at McMaster University. This protocol will include a consent form that all participants will be asked to sign prior to joining the research study. The consent form will provide potential participants with information regarding the objectives, nature as well as any potential risks of the study. It will also explain the measures put in place to ensure confidentiality of subjects' collected information and the option to drop out of the study anytime at their will.

Data collection for the actual study will be performed at two points in time over a 16-month period. At the beginning of the study participants will be provided with a short online video tutorial on how to use various features of the PHR system. Then, a survey will be conducted to gather measurement scales for the model factors except for actual use and facilitating conditions which should be measured after users have actually experienced using the system. Individual characteristics (demographics, details of previous computer and Internet use as well as perceived health status) will also be collected at this point. Data collected at this point will be used to validate a model for antecedents of behavioral intention to use PHR systems. Sixteen months later at the end of the study, actual system usage will be measured by analyzing the log files of the system. In addition, subjects will be asked to respond to a survey instrument measuring facilitating conditions. Data collected at this point will be matched to previously collected data and used to validate the full model (including both behavioral intention and actual use). Moreover, perceived health status will be measured again in order to investigate the existence of any meaningful change in this construct before and after using the PHR system. Based on prior research experience in the abovementioned clinics, a 16-month period was chosen to allow for sufficient patient-caregiver encounters thus providing participants with motivation to use the PHR system. The time period between measurement of model constructs and actual use will also allow the establishment of temporal sequencing thus avoiding common method bias (Compeau et al. 1999).

Measurement instruments: In order to ensure content validity, measurement scales for constructs in the proposed model will be selected from extant literature, and will be slightly adapted to reflect the context of this study. Measures that require considerable change will be re-developed following the guidelines suggested by Moore and Benbasat (1991). Intention to use will be measured using the 3-item scale by Venkatesh et al. (2003). Other constructs will be measured as follows: performance expectancy using the 7-item scale by Compeau and Higgins (1995); effort expectancy using the 4-item scale by Moore and Benbasat (1991); self-efficacy using the 10-item scale from Compeau and Higgins (1995); social influence using the 4-item scale from Thompson et al. (1991); health consciousness using the 6-item scale from Jayanti and Burns (1998); subjective health knowledge using the 5-item scale from Jayanti and Burns (1998). Finally, a formative construct will be devised for facilitating conditions to capture both concepts of technical support (Thompson et al. 1991) and caregiver support (Paswan and Young 2002) relying on both IS and healthcare literatures. Measurement scales for this construct will be developed following the three stages of item creation, scale development and instrument testing suggested by Moore and Benbasat (1991) while considering guidelines on specifying formative constructs by Petter et al. (2007). Perceived health status will be measured using
the 2-item scale from Kaplan and Baron-Epel (2003). A subjective measure for health status is employed as an individual’s perception of his/her health status is likely to influence his/her adoption and use of a health service like PHR system (Ware et al. 1981).

Qualitative analysis: Participants will also be asked to respond to open-ended questions relating to their concerns and experience with using the PHR at the two points of data collection in this study. Subjects’ responses to open-ended questions will be analyzed using NVivo in order to enhance the robustness of results as well as to strengthen the findings through triangulation (Benbasat et al. 1987). Triangulation involves validating the results by combining a range of methods (Tashakkori and Teddlie 1998).

Model validation: Structural Equation Modeling (SEM) will be used to validate the proposed model. SEM allows the analysis and investigation of unobservable variables that are indirectly measured from observable variables (Chin 1998). In particular Partial Least Squares (PLS) will be used as it is applicable to both exploratory and confirmatory research which is appropriate for this study and as it supports having both reflective and formative constructs in the model (Chin et al. 2003; Gefen et al. 2000). Further, PLS gives optimum prediction accuracy because of its prediction orientation (Fornell and Cha 1994). The measurement model in PLS will be assessed in terms of item loadings, internal consistency and discriminant validity (Gefen et al. 2000) using SmartPLS4 software.

Impact of individual characteristics: In order to investigate the impact of individual characteristics on PHR system usage, various statistical techniques will be employed. First, Chow’s (1960) test will be used to investigate whether the validated model varies for individuals with various perceived health statuses. To this end, subjects will be divided into two groups of well and un-well people based on their responses to the perceived health status scale. One PLS model will then be developed for each group and corresponding path coefficients will be compared across the two models using Chow’s test to identify any significant variations (Chin 2000). The same approach will be used to compare PLS models for males and females. Second, the impact of age, education level and frequency of Internet use will be examined by creating a series of control models. To this end, for each of the aforementioned variables, one construct will be added to the model and additional paths will be created from the new construct (e.g. age) to all the existing constructs. Then, variance explained for the constructs in the original model and the controlled models will be compared. Finally, we will examine other possible relationships which are not hypothesized, through a saturated model analysis (Chin et al. 2003). In addition, any possible interaction effects between independent variables will be examined using PLS as suggested by Gefen et al. (2000).

Sample: The sample size for validating the model in PLS is determined by the maximum of 10 times the number of paths leading to a construct and 10 times the number of items for the most complex construct (Chin et al. 2003). Self-efficacy is measured using the most number of items (10) in the proposed model resulting in a minimum required sample size of 100. Performing the Chow test will require twice as many subjects (200) since the models should be tested for the two groups separately (Ghilagaber 2004). To allow for possible spoiled surveys 300 participants will be recruited. Previous research experience in the clinics indicates a 50% initial response rate. Considering the longitudinal nature of the study and the fact that questionnaires will be sent out at two points in time, it is reasonable to expect that part of the subjects might drop out of the study. Assuming a 50% dropout rate, there is a need for initially targeting 1200 adult registered members of the clinics in order to ensure having 300 participants throughout the study. Invitations to participate in the actual study will be sent by regular mail to a randomly selected sample of 1200 adult patients and followed up with reminder phone calls. Approached members of the clinic will be incented to participate through opportunities to win prizes. In order to assess non-response bias, respondents and non-respondents of the both surveys will be compared, based on demographic information (Compeau et al. 1999).

4 http://www.smartpls.de
4. POTENTIAL CONTRIBUTIONS

Personal health record systems have the potential for helping consumers to take over more of their own care, thus reducing the burden on the healthcare system. But it is essential to improve the adoption rates of such systems if they are to have an impact on patient care. Behavioral issues have been identified as major inhibitors to adoption of PHR systems by consumers. Thus, there is a special need to focus on consumer behavioral factors influencing intention to use health information systems in general and personal health record systems in particular. While there is a plethora of research discussing adoption models for information systems in general, only a few studies have focused on the adoption of information systems in healthcare settings. The proposed study attempts to address this gap by seeking to develop and validate a theoretical model explaining the factors influencing an individual’s intention to use personal health record systems as a preventive health care behavior. Although this research is being carried out in a Canadian context, it is highly relevant for other developed countries that have similar demographic and healthcare system characteristics.

From an academic perspective, results of this research will contribute to the IS and e-Health literatures by developing an adoption model specific to PHR systems. It is hoped that this research will attract the attention of researchers to further develop and test constructs and models applicable to consumer intention to use personal health records and other health information systems as a preventive healthcare behavior.

Practitioners will also gain a better understanding of consumer preferences through this work, resulting in practical guidelines for PHR systems’ development, promotion and use. Results from this research can help direct attention to the most influencing adoption factors while proposing solutions that mitigate consumer resistance. Such solutions will enhance the PHR benefits for consumers and the healthcare system. Technology providers will benefit by informing the design of their proposed systems based on these results. This, in turn, will lead to higher rates of adoption and success of the Internet-based personal health records. Health care providers will also benefit from the results of this research by being able to deliver a higher quality level of care at a lower cost and complexity by involving patients in their own care through PHR systems. Given the growing importance of consumer-centered healthcare and e-health, adoption studies of this nature are both timely and relevant.

References


ATTITUDES TO INFORMATION TECHNOLOGY IN HEALTH CARE PROFESSIONS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0639.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Health information systems / medical record systems / care delivery /, Ethnography, Sensemaking, Technology adoption</td>
</tr>
</tbody>
</table>
ATTITUDES TO INFORMATION TECHNOLOGY IN HEALTH CARE PROFESSIONS

Ann Svensson, University West, Department of Economics and IT, S-461 86 Trollhättan, SWEDEN, ann.svensson@hv.se
Ulrika Lundh Snis, University West, Department of Economics and IT, S-461 86 Trollhättan, SWEDEN, ulrika.snis@hv.se
Pia Svanberg University West, Department of Economics and IT, S-461 86 Trollhättan, SWEDEN, pia.svanberg@hv.se
Lars Svensson University West, Department of Economics and IT, S-461 86 Trollhättan, SWEDEN, lars.svensson@hv.se

Abstract: The purpose of this paper is to explore attitudes towards IT among various categories of health care staff; health care professions. We will identify problem areas that may be the reasons for why different attitudes among different professions at a healthcare organisation exist, and subsequently we will analyse how this may have impact on how to make sense of IT use. The research question is: What factors may explain differences regarding attitudes to IT among different professions in a health care organisation? The paper reports from a particular study of the “NU” healthcare organisation in west Sweden. The results reveal two main problem areas: i) the infrastructural and; ii) the socio-organisational. These are discussed as analytical implications for bridging the gaps between different professions in health care organisations.

1 INTRODUCTION

As modern information technologies (IT) now offer new possibilities for improving almost every aspect of health care, their implementation is a highly relevant and fast accelerating process. The initiatives themselves vary greatly from scattered single projects of various ICT solutions to large national programmes. The domain has been studied intensively during the last decades. The studies have been conducted in several ways and from different perspectives. Many researchers have used empirical approaches to analyse how IT can support standardisation and flexibility (Hanseth, 1996), work practice and coordination (Lundberg, 2000), IT and gender (Wilson & Howcroft, 2000) and adoption in context (Chiasson & Davidson, 2004; Cho et al, 2008). All these studies address the issue of IT and its design and use in a medical work context. For instance, Heath & Luff (2000) present several examples where seemingly innovative and reliable systems have failed when introduced into specific and critical organisational contexts, such as health care work. In much of these examples the system in question does not necessary cause severe problems, but it is simply “under-used” or rather. They call for further examples that actually take the serious account of detailed user studies attempting to exploit the fundamentals of and attitudes towards new technology, in order to find implications for re-design or adopt them. Wilson & Howcroft (2000) highlights another aspect when they point out that IT use is more or less closely related to socio-cultural perspectives such as gender, roles, and the individual competence level. The results of their studies show that health care workers associate
information systems and information technology to effectiveness and control, which is perceived as being in conflict to the traditional values of healthcare: care and humanity. In Chiasson & Davidson (2004) they discuss the use of IS theory in understanding IT adoption in health care. They call for a deeper consideration of the actual context of health care settings in order to explore new challenges for both the health care industry (practitioners) as well as to develop new IS theories (researchers).

We argue that research aiming to understand people’s assumptions, attitudes, and knowledge is of special interest when discussing change in organisations, especially when change is caused by IT. Many varying factors contribute to the understandings of attitudes towards IT in health care. Many of the factors concern the flexibility of the systems and whether they were 'fit for purpose', along with the confidence and experience of the actual IT users. We believe that attitudes of practitioners are a significant factor in the acceptance and efficiency of use of IT in practice. There is a clear correlation between people’s mental frames and their technology use. Previous research indicates that during conditions where the technological frames of key persons and divergent stakeholder groups, such as managers, technical experts and users, differ a lot the result can be even more complex and cause problems and conflicts related to IT design and use (Orlikowski & Gash, 1994; Constantinides & Barrett, 2006).

In this paper we take this previous research seriously. Treated as purely technical in nature, IT issues in health care should gravitate towards a more soft and people-oriented approach. Point of care delivery is vital for the success of any application in the clinical healthcare environment. We aim to explore attitudes towards IT among various categories of health care staff; health care professions. We will identify problem areas that may be the reasons for why different attitudes among different professions at a healthcare organisation exist, and subsequently we will analyse how this may have impact on how to understand and use IT. The research questions is: What factors may explain differences regarding attitudes to IT among different professions in a health care organisation? The paper reports from a particular study of the “NU” healthcare organisation in west Sweden. By attitudes we mean peoples assumptions and expectations that relate to their understanding and use of IT. We assume that these are tightly related to the knowledge and experiences developed during IT implementation and adaptation processes. This is based on assumptions used for instance in theories developed by Henfridsson (1999). Orlikowski & Gash (1994) use the concept ‘technological frames’ when they analyse different attitudes towards technology. They say that individuals create mental models and expectations about their future use of IT in work organisations. Here we will identify the gap that may exist in between the different models of the technological aim and how this gap may be bridged.

2 RESEARCH METHOD

We have used ethnographic methods as a means to study contexts where the organisation of human interaction and computer support is apparent. Ethnography is a collection of methods for collecting and analysing data from empirical settings. According to Hammersley & Atkinson (1995) the importance of attitudes is obvious, which means that we try to understand and make sense of actors and activities in the organisation under study.

Our research was conducted as an ethnographic study of a health care organisation in Sweden, the NU Health Care. The NU health care is divided into five hospitals in different cities, two of the hospitals are complete with emergency departments and the other is classified as local area hospitals. The study took place in two of the five sites of the highly distributed organisation. In order to get a
coherent understanding of the work conditions and the people's attitudes to IT we strove for a broad representation from various key professionals from both the emergency department and the IT department. All together we interviewed 20 persons that represented the operational as well as management categories at both departments. Their professions were ranging from doctors, nurses, secretaries, management, IT co-ordinators, supporters to ambulance drivers. As mentioned earlier, the two sites we studied were located in two towns. Here we can suppose that there exist different cultures that have developed over the years. However, this perspective addressing a more geographical and cultural aspect we have not been into so deeply.

We applied semi-structured interviews. Some questions were adjusted in order to capture the specific conditions of the different interviewee categories. The questions were grouped under headings such as current work activities, attitudes to IT, the fusion and integration of the organisations into “NU”. The interviews lasted for 45 to 75 minutes each and were all recorded. Except the interviewer we were at least one more to take notes and comments during the interview session. In our study the informants were interviewed in their real work place and wore their actual working clothes. A consequence of this may be that the attitudes recorded in the interviews are rather expressed in a role-playing game, from a perspective of what they should say rather than what they in fact like to say.

By using qualitative analysis to the empirical data, many researchers see the potential for informing appropriate lessons learned for improving the design and use of IT. (e.g. Hughes, Randall, and Shapiro, 1992). In our analysis we try to be constructive in the situations where we found a “gap” between different technological frames among the different professions. In this analysis we did not necessarily see typical design issues but rather analytical findings that may have impact on the design as well as use and adaptation of IT in the organisation. Before the analysis took place we read the material several times. We interpreted the data through continuous discussions. At the same time it was possible to reflect upon the situation recorded and this is of special importance in long, qualitative interviews, according to McCracken (1988).

3 ATTITUDES TO INFORMATION TECHNOLOGY

There are several perspectives that relate to people's attitudes to technology, on the one hand the perspective that relates to the individual and on the other hand the one that relates to social processes within the organisation. To be able to interact with technology people have to make sense of it, and in this sense-making process, they develop particular assumptions, expectations, and knowledge of the technology, which then effect their actions toward it. While these interpretations become taken-for-granted and are rarely brought to the surface and reflected on, they nevertheless remain significant in influencing how actors in organisations think about and act toward technology, (Orlikowski & Gash, 1994). Further, Orlikowski & Gash (1994) suggests a conceptual framework when studying people’s underlying assumptions, expectations, attitudes, knowledge and experiences of technology. These interpretations of technology, technological frames, are central to how people interact with technology. The framework consists of three categories that can be applied on empirical data to characterise the result.

- **Nature of Technology** – refers to people’s images of the technology and their understanding of its capabilities and functionality.
- **Technology Strategy** – refers to people's views of why their organization acquired and implemented the technology. It includes their understanding of the motivation or vision behind the adoption decision and its likely value to the organization.
• Technology in use – refers to people's understanding of how the technology will be used on
a day-to-day basis and the likely or actual conditions and consequences associated with such
use.

These conceptual models or technological frames are essential when the individual interpret and
understand his and others actions toward information and communication technology. It is however
important to emphasise that these interpretations have a distinct social dimension. They appear,
develop and change in interaction with the information technology as well as with other people.

A sociologist who has great influence on the understanding of how people relate to information
technology is Karl Weick. Weick (1995) consider that sensemaking can start with someone noticing
something different in the organisation. There are several barriers to make these assumptions public
and accepted, although one ought to consider that the problem is obvious. It may be something very
complex and therefore difficult to understand. Orlikowski & Gash (1994) refer to the same
phenomenon when they argue that pre-understanding play an important role under these
circumstances.

Henfridsson (1999) has been studying sensemaking in organisations related to adaptation of IT. He
proposes that Karl Weicks (1995) sensemaking perspective is especially useful in order to understand
IT-adaptation in a micro-perspective. Henfridsson has been studying the process, which an
organisation transits when IT first is introduced and until the technology is a part of people's taken-for-
granted conceptions. He also defines three elements of this transitional process, between the phase
associated with ambiguity about an IT-artefact and the following phase of common-sensible use of the
same IT-artefact. The elements are referred to as “enactments” and can be described as identity-
construction, self-fulfilling prophecies, and organisational defences. In our study Orlikowski and Gash
may help us in understanding the different attitudes from the perspective of an individual professional
 technological frame. The categories that they are suggesting can make it possible to obtain a clear
picture of an actual condition in the process of IT-adaptation.

4 RESEARCH SITE: THE “NU” HEALTH CARE ORGANISATION

The study was conducted in the NU health care in Sweden. The NU health care is divided in five
hospitals in different cities, two of the hospitals are complete with emergency departments and the
other is classified as local area hospitals. In the below sector the organisational structure, that faced the
reality when the study was conducted, is described. At this time there was although a very unstable
situation characterised by organisational changes. This comes from political turbulence as well as from
constant threats about economical reductions for the health care sector in general, but most important
was however the discussion that dealt with the question of closing one of the emergency departments
in the NU health care.

4.1 The IT-department

There are three people sharing the role of IT-manager, one of those are also responsible for the IT-
strategy within the NU healthcare and is also a doctor. As an IT-strategist he also participates in the
regional IT-managing group, which are supervising the IT-director and have the primary responsibility
for IT in the region. There is also an IT advisory team that function as a link between the health care
management and the IT managers, in this team both practitioners and IT-people is taking part.
4.2 The emergency department

The emergency department is divided in two departments; each of them located on different hospitals in different cities. The two departments have one common manager. There are several section managers on each department responsible for scheduling, personnel businesses among other things. The emergency department has a medical department, a surgeon department and an emergency department for children, ears and gynaecology and during the nights and weekends. In the clinic nurses, assistant nurses, doctors and medical secretaries works. Every department has their own IT-coordinator that function as the primary support to the personnel.

4.3 The information in the healthcare system

The information flow in the healthcare sector is complex. It can be divided into three categories. There is the information that is protected by the law of secrecy, the so-called medical records where patient information is collected in archives. The information includes individual information, diagnoses and treatments for every contact that the person has had with the department. The information system also includes an administrative information volume, for example messages from the personnel department, protocols, and notes from meetings among other things.

When it comes to information that aim to support medical treatments, documents are created by one individual that has the expert knowledge within the specific area. This information is then used in a variety of different documents and is part of many different document-flows. Either the document is used as it is or it is read and rewritten. This duplicating of information involves many risks. When the document is distributed throughout the organisation, it is difficult to determine what version is the most recent. It is impossible to know how many who has interpreted and rewritten the information, and therefore it is hard to know if the original message is changed. As expressed by one of the IT-coordinators:

“The information handling in the healthcare is extremely complex. There is a need of one solid group of people who really understand how systems is constructed in order to describe the handling of information in the health care properly and to deliver the specification to the people who are going to develop this. I do not think we have ever succeeded with it...we must develop solutions that generates something useful for the people that are making the beds and are helping the patients on the health department. If we cannot do this soon, then they won't understand why we buy a computer and put it in the expedition. ...We have a long way to go, the health care need to mobilize a huge knowledge to be able to do the descriptions that is needed.

5 SENSE-MAKING OF IT IN THE “NU” HEALTHCARE

The ethnographic study has provided an understanding of the sensemaking of IT in the organisation. There are large differences between different groups of professions, though within one profession the sense making has shown to be relatively homogeneous. From our empirical material we have tried to crystallise the different approaches to IT, put them together in groups and identify the contrasts between them. To explain the different approaches to IT we have derived two comprehensive categories that are argued to constitute the essential problem areas: infrastructure and socio-organisational problems.
5.1 Infrastructural problem areas

In the infrastructural processes the present systemic conditions can be seen as a main problem. We found three subcategories, which we decided to name activity-adapted functionality, communication and integration between systems together with co-ordination of initiatives.

5.1.1 Activity adapted functionality

In the NU health care they have only started the work with descriptions of work processes and routines. This work has been started to avoid creating IT solutions with weak coupling to the organisational activities.

The reason is that some people in the organisation have realised the importance of considering the IT infrastructure to support organisational activities. The IT technicians often see as their speciality to secure the technical specifications and other requirements on the IT equipment. People in the NU-organisation are concerned with that IT have proper functions to help them in their daily work. But they are not always aware of the need for functions to cooperate. Different professions here think in different ways. The knowledge about the activities within different areas in the organisation, co-ordination between departments and the technical knowledge do not always walk hand in hand.

People who order new systems need knowledge about organisational needs. When ordering new systems the discussion easily drifts to first of all specify the equipment and its details. The staff in the healthcare is enthusiastic, but at the same time disappointed or dissatisfied. The management often is perceived as alienated from the need of IT support in different areas in the organisation. For example when an Intranet is created there is a risk that management sees and compare this to an information site on the Internet and the references become incorrect.

There is also a problem consisting of communication, where management is not aware of the complexity of managing information. If IT should satisfy the requirements of different work activities in the organisation, a different approach to is needed. The ambition could be to go from standardisation to methods which are characterised by de-central requirements and co-ordination.

5.1.2 Communication and integration between user levels and system levels

The vision is to have access to the patient’s records wherever the patient is located. The vision is also that the systems are open and able to communicate with each other. From a user level perspective the routines are almost the opposite today:

“We are too bad in describing our requirements in the NU health care. Often it is too urgent, the clinics miss the long term thinking about IT support. We are often too late...”

People within the IT profession at “NU” needs to participate in an earlier phase in the requirement process. Both the IT department and the clinical departments need to get together and come to an understanding about functions in a desired system. It could then be easier to formulate the requirements before looking for suitable systems. There is a great amount of work going on at the IT department on standardisation of the technical platform and the products in use. They also try to standardise the user interfaces on the screens.

But from a system level perspective, there is another gap as well. How should the systems exchange information? Today there is a lack of knowledge about principles of standardisation. They are quite unaware of that the standardisation often has an impact on making a choice of a system and that it could be crucial in creating suitable systems- and communication solutions. The integration
between systems is almost non-existent. For example the system for patient case books is completely separated from the system for patient administration, which handles notices to appointments, next visits, diagnoses, registrations and so on. At one clinic they use three different systems to complete the patient care documentation. The systems have to be activated at the same time because some of the information in one of the systems must be manually registered in another system. The work of harmonisation of systems in the NU health are is going on, but is often met with suspicion.

“The Västra Götaland region wanted that system [Melior], we, used to work with the old system was not asked, as usual.”

When choosing a system it is important to analyse what the system should support and if there is need of a system which allow integration between systems in a smart way.

Most of the patient’s records are paper-based today. Though there is an objective that all records should be computer-based in 2002. Medical records are handled separately by each clinic, and each clinics have separate records for each patient they have treated. If the patient is coming to a new clinic it is impossible for them to see if and what medicine is previously prescribed. The consequence is that much time in spent looking for old records in archives at the different clinics or records are freighted by taxi between different hospitals.

“It should be possible to fill in the forms of X-ray referrals, laboratories referrals etc at the computer. It is bad that it is not possible today. At Huddinge, where I have worked before, the possibility existed already in 1979. It is badly managed by the management. Therefore there are no uniformed systems. It will be difficult for us because we need all three systems. It will probably change very much.”

5.1.3 Co-ordination of ideas, initiatives and projects

The results show that there is an optimistic attitude to the future possibilities with IT, both at the operative and the management level. Some different initiatives have already begun, which could be seen as embryos to a positive work towards changes. There is a tendency to invest money and resources on many different small IT-related projects.

In the organisation there are many fruitful ideas about systems for more efficient administration. Workings groups exist within clinics at both NÄL and Uddevalla to co-ordinate the work when new jointly systems are introduced. As an example a project with development of an Intranet could be mentioned. The Intranet should be created and implemented to give the users access to information in an easy and efficient way. The ambition is also to emanate from an original document, and use this original document in different document flows. The aim of the Intranet project is furthermore to develop the infrastructure of the information processing. It is clearly expressed that the Intranet project should not affect and changes local work practice. When each small project is carried out with its own budget the effect is that the advantages of co-ordination is missed. The way that the systems will communicate is also missed when small projects are carried out without co-ordination with each other. Often a more overall project is missed. It is the integration and communication between systems, which are very important when the systems should be used in an efficient way. Other consequences of small projects can also be that small information islands are created, from which information is distributed. The very same information can be created in different places and the risk for an uncontrolled information overload without co-ordination is obvious. No common information sources are created.
5.2 Socio-organisational problem areas

Problems due to the practical use of IT appear continually. We refer to these problems as a socio-organisational problem area. These are extremely intertwined with and in a sense derived from the infrastructural ones, but here we outline particular categories such as lack of communication, the dialogue about IT, attitudes to their daily IT use, about uncertainty and differences.

5.2.1 Lack in communication – the dialogue about IT

The results show an obvious lack in the communication between different groups and professions within the organisation. In the empirical study there are many examples on this phenomenon. Regarding the NU health care IT vision the personnel has lack of knowledge. All of the personnel told that they did not know anything about the IT vision. The conclusion is that the IT vision is not anchored in the organisation and that there is an obvious lack in communication between the different groups.

The dialogue concerning user and systems requirements has shown to be problematic. The health care personnel consider themselves as bad in describing their requirements. The reason is not fully clear. But the mental pictures of what to focus in the choice of system are different in different groups. For example has the health care personnel more knowledge about which functions the system has to have. The IT personnel are instead focussed on which communication standard is required. Today the health care personnel come up with ideas. These ideas are also rather incomplete due to the technical requirements. One reason to the problematic situation is that the technical competence is coming in too late in the process of describing the requirements. The following example also shows the difficulties in communication between the IT personnel and the health care personnel. The language used within the IT community is rather technical and that will cause problems.

“It is very difficult for me to really explain in the project administration group, what we are able to create in this project, and to have them to understand this theoretical discussion and give them a chance to really say yes or no. They really understand the activities in the health care, but… the project has interpreted the clinic requirements to a computerised solution and the project administration group is going to take a decision to the amount of resources to create this solution. This implies a lot of things; there are always risks with development projects. It is heavy principles they have to decide on, because it is there principles we have to build upon later on in different IT systems. But when the project administration group says yes, I cannot clearly feel that they have understood what I have told them.

There is a prevalent uncertainty among the IT personnel and the management level at the clinics in decision processes, where IT personnel have knowledge which is hard to communicate to them who are going to make the final decisions.

“Unfortunately we don’t have this, we should need this, we not always agree with the catwalk… so we have talked loosely about this. But it is still so new…”

“[It is] much dependent on whom you are talking with up there [at the IT department]. If I talk to someone it is away with the fingers, do anything at all. If I talk to someone else it is yes, you can do this… I don’t really know which leg to stand on.”
5.2.2 Attitudes in the daily IT use

The functionality in the systems is not always adapted to the clinic’s routines, which influence the attitudes to IT in a negative way. The knowledge about the systems functionality is therefore important. Regarding signing documentation in the patient’s electronic case books the consequence sometimes could be that one have to sign what others have documented. It is perceived as risky because it is possible to have the responsibility for something one has not being involved in. The consequence will be that they not document as much as before.

An essential problem in the daily use of the system is the lack of education in using the documentation system. For example at a clinic in Gothenburg were some of the personnel dedicated for some weeks only to support the use of the system, which was seen as a prerequisite to handle the documentation in a satisfying way.

In spite of these problems with the system the personnel is very positive to the system. In a few years they think it will work very well. The system only has to be more adapted to the clinics’ routines and the personnel need education in using the system.

“If we had good systems and if they worked we could dedicate more time to the nursing work.”

“I see very good possibilities with IT but nothing is happen or it goes to slowly and there are many problems…”

The clinics’ personnel want a better co-operation, and they wish that they could call the IT department and have support instead of having a service order number.

“I stumble against many problems every day, which is due to the users lack of knowledge.”

Amongst the health care personnel there is a negative attitude towards IT, which tends to be caused of fear to changes and to the overall technology. The IT personnel are also aware of this, which the following citation shows:

“… I have also encouraged everyone to use the computers and the Internet to learn and exercise. So the use of hotmail is absolutely OK. Otherwise they will be kept away from the computer if the only are allowed to do certain things. Then they get more exercise to use the keyboard an so on.”

The health care personnel see the functions closed to the work activities, which could be computerised. The attitude to IT is much troubling and there are many problems with IT, but if it works it will possibly be good.

“I think we can use this much more than we are doing today. But we need education and information about this. I feel that if one is not trained or interested from the very beginning it is difficult to dare. One does not go into the system then.”

The prevailing situation is that there is a kind of negative resignation towards IT. The development is too slowly as it takes too much time etc. The health care personnel also feel that they are bad in handling computers:

“The majority of us are women and have low interest of computers. We see them mostly as type writers. Our computer technicians have not come to their advantage. Probably we could solve many of the problems but we have no time to learn. I see good possibilities with IT but nothing is happened or the process is going too slowly and there are much problems…”

“As a section leader I see the computerised system as a huge problem. We are going to have a
new salary system at the new year, Palett, so we hope it could be better. The old system was miserable, what should take one hour takes a whole day. Now we are doing the salary reports manually so that the salaries will be paid in time before Christmas. It has been hard to log in to the systems.”

5.2.3 Uncertainty and differences

Individuals could perceive differences between events going on and events which told to happen (Weick, 1995). It can be exemplified of when a clinic should choose a system. A clinic had a wish of a specific system adapted to their activities. The IT co-ordinator perceived that the choice of system been something else and not the system they wish to have. Even if the IT co-ordinator in this situation tried to convince himself that it really was the original wish of system they still were searching for, the result was not what was thought. He interpreted the situation as a gap between how it was and how it should be. He tried to make sense of IT based on the intentions of the system, his knowledge and experience about the health care activities and the presumption he had.

The personnel then discovered that the system was not that integrated as expected and they experienced interruptions in their routines. It can arise situations hard to interpret which comprehend unstructured problems, there both the goals and the means are problematic and unclear (Henfridsson, 1999). Weick (1995) means that there is no need for more information when overwhelmed of ambiguity. Instead there is need for appraisals, priorities and clarity of preferences.

One more expressive example relates to the development of an Intranet, which was conducted as a project with a reference group. The reference group is composed of user representatives from five different clinics. Though it is a relatively small group of users. The situation is risky in that it can be more confused and that more interpretations can be added if more information is spread to the individuals in an organisation (Henfridsson, 1999). An IT co-ordinator expressed it as:

“It is tremendously hard for the average user to understand a theoretical discussion of how an Intranet will work and how we had planned the information handling, when the user cannot see the practical solution in front of him/her. And I would say that it gives no more than creating more question marks, therefore we have chosen to work with a reference group of users instead.”

It is problematic if the IT co-ordinators would like to change the configuration of the computers at the clinics. Even if they have the education to do this, and if they have a storage from old computers, they are not allowed to mount the storage without the approval from the IT department. At the same time is the IT department only responsible for the equipment until the wall socket at the clinic. The consequence will be that the personnel feel watched over. There is also a scanning program at each computer, which every time when the computer is stated, reads the configuration of the computer and detect if any new programs have been installed. The personnel is also monitored in their use the Internet. The IT department have the possibility to see which pages the personnel are looking at and for how long time they are connected to the Internet.

6 DISCUSSION

An analysis of the attitudes towards IT with respect to the three dimensions of technological frames (Orlikowski & Gash, 1994) starts with the fairly obvious reflection that the meaning of technology is perceived differently by the two professions: IT- and health care personnel. The technical education and competence of the IT-personnel results in a more imaginative and change oriented attitude. At the
same time, this profession shows a tendency to underestimate the potential problems related to changes in work practice brought on by the use of new systems. In that sense, IT staff can be argued to have primitive view of what Orlikowski and Gash refers to as technology use.

Work at the emergency health care unit is highly complex and involves processes where a large number of people, often under great stress and time pressure, shall co-ordinate their respective tasks into a well-oiled machinery. Even the smallest disturbance or flaw might result in serious consequences for the patients. In light of this, it is perhaps not surprising that the staff mobilizes an organizational defence (Henfridsson, 1999) as a reaction to plans of new systems and applications in the clinic.

The strategy of technology refers to the way an individual or a group of people understands the purpose or the agenda behind the implementation and use of information technology. Wilson and Howcroft (2000) argues that healthcare operatives tends to perceive IT in general as having the primary purpose of monitoring, controlling and rationalizing work. This is in clear conflict with the traditional ideals of nursing, caring and healing, and can consequently result in a rejection of new technology. Subsequently they explain this conflict between technological and professional ideal as being rooted in traditional gender-roles, with nursing and caring as female attributes whereas control and rationality are traditionally male. This gender aspect probably gives an important contribution to the analysis of the case study, but other aspects can offer equally powerful contributions.

As pointed out by Goffman (1969) people tend to adopt different roles when presenting themselves in public life. It cannot be ruled out that the critical attitudes towards new IT systems found in this study has other origins but the ones conveniently afforded by the professional role of a stereotypical nurse or doctor. Perhaps it is easier and more socially accepted to use such a line of argumentation instead of expressing for instance a fear of change and a reluctance to abandon familiar work practices. Furthermore it is worth noticing that this technological strategy ("the vision") is not expressed among IT-personnel. Instead they tend to emphasize efficiency, professionalism and "keeping up with scientific progress". Technology is viewed as having a large potential with respect to increasing the quality of healthcare professionalism. The fact that healthcare operatives still view IT as primarily an instrument of control can be interpreted as a lack of confidence in IT staff as legitimate authorities in the area of healthcare.

Dissonance with respect to how the work practice is perceived and understood could have a negative impact on performance. This study shows that health care operatives are a fairly homogeneous community engaged in a process of common sensemaking with respect to their technological frames. This is so, inspite the fact that the community contains a number of different professions and specialities that needs to communicate and co-ordinate their activities. The study also shows how the IT-staff has arrived at different, yet shared, understanding of technology. This group expresses a somewhat uncritical and unproblematic attitude towards IT, which does not make sense for the end-users in operative healthcare.

The operatives report problems and calls for action regarding the complex processes of information management. IT professionals design 'solutions' aiming at addressing these issues, that in turn is perceived as not appreciating the nature of existing work practices, and consequently run a high risk of being rejected. This fundamental conflict creates a gap that calls for organisational changes in order to be bridged.

In this context it is important to point out that healthcare operatives should not be perceived as being totally hostile towards all technological change. Healthcare in general and emergency healthcare
in particular is constantly invoked in adaptation of new technology. New machines, instrument, and methods of treatment is reforming and revolutionizing work practices all the time. These innovations are however originating from sources that healthcare operative perceives as legitimate professional authorities, hence national or international medical specialists and or new government legislation.

This could contribute to that some of the suspiciousness connected to implementation of new IT systems is replaced with positive expectations and faith, something that dramatically would open for individual and collective sensemaking. In this sense, it is all about giving the IT systems "a fair chance" and consequently end up with faster adaptation, where members of the organization are tolerant and hospitable (Ciborra, 1999), overseeing initial problems related to IT. Such processes should be seen as benign, when they result in improvements with regard to the quality of healthcare, but there is also an evident risk of prolonging the lifetime of less successful ventures. Both scepticism towards IT and positive expectations regarding other care-technologies run the risk of becoming self fulfilling prophecies (Henfridsson, 1999).

What makes the situation somewhat hopeful for the future is the fact that healthcare sector have long traditions of coping with situations where people of different professions (assistants, nurses and doctors) and different area of expertise (medicine, surgery, pathology etc.) manages to integrate their efforts into an efficient operation. The challenge at hand is consequently to involve also IT-specialists and administrators.

Addressing the socio-organisational aspects we believe that a strong candidate as a possible future link between health care professions and IT profession is the IT-co-ordinators. At present their situation is problematic, especially since their profession is fairly new to the organisation, and has not yet positioned itself in relation to other professional groups. The sensemaking of their environment and themselves has not converged to a situation where all stakeholders agree on what part they will be playing. Apart from IT-co-ordinators, also the operative managers might play an important role. Having the benefit of being legitimate authorities in both camps, this category of people could hopefully catalyze a fruitful dialogue that will result in a shared culture with positive attitudes towards sound IT interventions.

The analysis highlights the infrastructural aspects as important aspects of understanding the way people in NU-healthcare relate to IT. The results can be understood as contradicting paradox, where calls for flexible systems that are well rooted in work practice is in conflict with the demands for co-ordination, efficiency and standardisation. It is highly problematic to run large-scale bottom up projects, and at the same time impossible for top down initiatives to gain sufficient momentum, because of organisational defence mechanism (Henfridsson, 1999).

Some changes in the healthcare sector, such as implementation of new apparatus or methods of treatment, are local and situated to one or a few specialist clinics. Other systems (Melior, Adapt etc.) relate to information management (patient records, process reports etc.) and concern the entire organisation. This type of exhaustive systems are difficult to adapt to micro-level requirements, and are to a large extent designed in a way that makes integration with other systems very difficult, (Lundberg, 2000; Hanseth, 1996). Also Lundberg (2000) passes the ball back to designers and system-vendors for the healthcare sector, when arguing that the paradigm of designing generic systems must be replaced a paradigm of designing work oriented infrastructures. Existing systems should be cultivated into supporting a given work practice. This does not imply a conservation of existing routines, but rather that innovation and change should evolve from a close co-operation between designers, IT-professionals and healthcare operatives, making sure that common sensemaking is
aligned with an optimal use, meaning and strategy for information technology. This may call for a profession-oriented infrastructure rather?

7 CONCLUSION

The aim of this study was to explore problem areas concerning different attitudes to IT in health care. The conclusions are derived mainly from the ethnographic study but some theoretical discussions have also been developed. Two main problem areas were found: the infrastructural and the socio-organisational. We have argued for a dialectic approach, which means that we consider both the individual as well as the social processes that build up the attitudes to IT. We found reasons for the differences that include lack of co-ordination between professional groups, human communication problems, and hindrances in systems integration and low level of knowledge and understanding in the daily IT aim and use. Here we have tried to be constructive and find mechanisms that may trigger the social processes between people with different attitudes. The adaptation process will be extremely important as a stage for these processes to take place. We believe that open up the dialogue and reason about the technological aim may foster attitudes that make a better understanding and use of IT in the “NU” health care organisation as well as in similar organisations.

8 REFERENCES

Chiasson, MW and Davidson, E. (2004), Pushing the contextual envelope: developing and diffusing IS theory for health information systems research, In Information and Organization 14(3), 155–188.
Hanseth, O, 1996, Information Technology as Infrastructure, Department of Informatics, Göteborg University
Henfridsson, O, 1999, IT-adaptation as sense-making: Inventing new meaning for technology in organizations, Department of Informatics, Umeå University, Sweden.
McCracken, G, 1988: The Long Interview, California, Sage Publications
Orlikowski, W, Gash, D, 1994, Technological Frames: Making Sense of Information Technology in Organizations, ACM
INNOVATIVE CAPABILITY DEVELOPMENT PROCESS: A SINGAPORE IT HEALTHCARE CASE STUDY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0091.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Case Study, Health information systems / medical record systems / care delivery /, Innovation Driver / Enabler / factors / process, IT innovation</td>
</tr>
</tbody>
</table>
INNOVATIVE CAPABILITY DEVELOPMENT PROCESS:
A SINGAPORE IT HEALTHCARE CASE STUDY

Say Yen, Teoh, RMIT University, Building 108, Level 17, 239 Bourke Street Melbourne, 3000 Victoria, Australia, sayyen.teoh@rmit.edu.au

Shun, Cai, The Logistics Institute-Asia Pacific: National University of Singapore, E3A, Level 3, 7 Engineering Drive 1, Singapore 117574, tlics@nus.edu.sg

Abstract

With the global spending of more than US$4.1 trillion in 2007, healthcare has been one of the largest industries worldwide. To better plan and develop the healthcare system, an effective way is to leverage on the advancement of information technology (IT). However, despite of many attempts in the computerization and automation of healthcare services, most of the healthcare information systems failed to deliver in the era of rising expectations. With respect to this aspect, this study focuses on the success stories of a private-owned hospital in Singapore. We illustrate our lessons learned from this hospital, the Alexandra Hospital, which has successfully innovated a healthcare system setting and making a benchmark standard for the Singapore healthcare industry. By conceptualizing on the generation of innovative capability process, this study complements the scarce innovation literature in the service industry. Theoretical and practical contributions were shared and documented in this paper.

Key Words: Information Technology in Healthcare, Innovation in Services, Innovative Capabilities and Case Study
1 INTRODUCTION

Healthcare is the single-largest industry worldwide with a total spending of more than US$4.1 trillion (World Health Organization, 2007b). In which, the U.S. healthcare expenditures alone make up 14% of its GDP, and this figure is projected to rise to over 16% ($2.3 trillion) by 2010 (Plunkett Research, 2003). Recognizing the massive size and continuing growth of this industry along with its potential of transforming societal health-and-wellness quality, the Singapore’s Ministry of Health sets the mission to develop the world’s most cost-effective healthcare system to help keep Singaporeans, the country’s key resources, in an ideal state of health (World Health Organization, 2007a).

One of the most prominent ways to improve the healthcare system is to leverage on the advancement of information technology (IT). Healthcare experts, policymakers, payers and consumers find that computerization and automation of health records and processes could critically transform the scenarios of healthcare industry (Dwivedi et al., 2007). Though much in-depth studies had been done to improve the healthcare information system, yet it fails to rise on the occasion.

In view of rapid changes and uncertainty in the healthcare industry, it is crucial, practical and interesting to investigate how organizations can innovate and stay competitive in such a volatile environment. From expert’s opinion, organizations could shape the future of their industry through innovative changes (Lawson & Samson, 2001). The problem is that, an acceptable, comprehensive and systematic framework towards successful innovation is difficult to come by. This means that, an ideal and systematic framework of innovation has yet existed (Lawson & Samson, 2001). The scarcity of studies on innovative change in the servicing industries (Amble & Palombarini, 1998, Aranda & Molina-Fernandez, 2002, Martin & Horne, 1993) has motivated to focus on this subject. In this study, we intend to bridge the research gap by conceptualizing the generation of innovation services and innovative capabilities to illustrate lessons learned from a private hospital that has successfully designed and injected an innovative philosophy into its healthcare system-a service firm.

The other parts of this paper are structured as follows. First, a review is presented on the background of innovation services with innovative capabilities before an outline of our research methodology which is coupled with a description of the Alexandra Hospital’s case study. Second, the findings and discussion on the generation of innovative capability processes are listed in detail. Finally, the theoretical and managerial contributions are put in place.

2 LITERATURE REVIEW

2.1 Innovation in services

Despite the growing awareness that innovation is not confined to technical processes and products alone, contemporary research on innovative activities is still largely focused on technical innovations in the manufacturing sectors (Gallouj & Weinstein, 1997, Koch & Strotmann, 2008, Tether, 2003). Until recently, researchers began to recognize that there are differences in the nature of innovation in services in comparison to manufacturing. Probably, the two most prominent characteristics of services compared to manufacturing are: intangibility and interactive aspects of services. The reasons may be that firstly services tend not to have an independent physical existence and also service innovations can be invisible. Secondly, as services are interactive, they are often being co-produced by the providers and users acting together” (Tether, 2003). For these reasons, it is interesting to explore how innovation can be specifically carried out in the services sector of a hospital setting in this study.

Three schools of thoughts have emerged so far in the services innovation. The first school of thought advocates supplier-dominated innovation and suggests that services innovation is largely dependent on adopting externally developed technologies in facilitating the provision of new services and/or...
enhancing the productivity of their services (e.g. den Hertog, 2000, Pavitt, 1984). Services are often considered to be “laggards” in terms of innovation and dominated by the technology provided by suppliers (Pavitt, 1984). More recently, some researchers began to suggest that service innovation might be beyond pure technological innovation. Beginning with the “interactive model” suggested by Kline & Rosenberg (1986), the second school of thought on service innovation emphasizes the differences between services and products, and suggests that services are far from dull providers of standard activities, but are instead dynamic and fluid, constantly changing to meet customer requirements, and achieving them through creative combinations of hard (e.g., equipment, computer software, etc.) and soft (human skills, knowledge, etc.) technologies (e.g. Camacho & Rodriguez, 2008, Gallouj & Weinstein, 1997, Kline & Rosenberg, 1986, Sundbo & Gallouj, 2000). This perspective of innovation in services strongly emphasizes interaction with users and is believed to be particularly relevant for a knowledge-intensive business services (KIBS), such as consultancy. The third school of thought sees competition and innovation as related processes, which are themselves related to the scope for strategic positioning (e.g. Tether, 2003). It further suggests positioning and innovation through creative combinations of hard and soft technologies. While this perspective accepts that the adoption of technologies can be hugely significant for service development, it also considers that adoption is rarely as passive or as dependent as the supplier-dominated view (Tether, 2003). With the basic knowledge on these schools of thought fundamental to services innovation, it provides us a clearer notion to further explore them from a hospital perspective. Most interestingly, in this case, we examine the services innovation in the process of adopting a new workflow and information system in a hospital setting. To further extend the background understanding of this study, we look into IT innovation literature.

2.2 Information technology innovation

Information technology (IT) innovation has been extensively studied by researchers in the Information Systems (IS). Early studies on IT innovation from IS perspective have been focused on the impacts of perceived technology characteristics on adoption of technology. Relying on the theories of innovation diffusion, key perceived technology characteristics have been identified, such as perceived complexity, relative advantage, compatibility, etc (e.g. Rogers 1995, Moore & Benbasat, 1991). More recently, scholars suggest that the fit between technology and the task might significantly influence the success of technology adoption (e.g. Goodhue & Thompson, 1995, Zigurs & Buckland, 1998). On the matter of the main stream philosophy, IT innovation has also been recognized as a process of movement, i.e., transforming inventions into usable forms and use in practice (King et al. 1994). Holding fast to the same belief, we examine the service innovation of this healthcare case study from a process perspective.

By adopting a Health Information Systems (HIS), we could most probably share some similar concepts and situations in normal IS adoption. Nevertheless, the adoption could be quite different from the generic IS approach, given its specific purposes in serving the healthcare practice. For example, Heeks’s (2005) “design-reality gap” conceptual model has highlighted HIS implementation challenges. Hence, it would be interesting to find out the generation of innovative capabilities introduced by AH hospital. Especially, in the case of AH, the extent of its innovation does not only evolve around the technology, but also include the change in its hospital workflow and structure. This has added the creditability to this case study.

2.3 Innovative capabilities

The concept of innovative capabilities refers to a firm’s ability to develop new products, services and/or markets through aligning strategic innovative behaviors and processes to achieve the usual and novel solution (Wang & Ahmed, 2007). In general, there are two types of innovative capabilities: incremental and radical innovative capabilities (Dewar & Dutton, 1986). Incremental innovative
capability is defined as the capability to generate innovation in a progressing manner, which refines and reinforces existing products and services (Subramaniam & Youndt, 2005). Radical innovative capability is the capability to significantly generate and transform existing products and services (Subramaniam & Youndt, 2005). Fundamentally, the distinction between incremental and radical innovative capabilities lies in the ability to draw organization knowledge (Subramaniam & Youndt, 2005). Such capabilities are particularly associated with the ability to enhance a firm’s performance (D’Este, 2002, Deeds et al., 1999), evolution and survival in the light of external competition and change (Deeds et al., 1999, Lazonick & Prencipe, 2005).

To innovate, it would require the recognition of the new, unfamiliar capabilities of IT for rethinking business processes instead of its familiar ones (Whitman, 1996). Unfortunately, most organizations often do not provide sufficient supportive resources for individuals who are prone to IT innovations (Wu, 2003). Thus, it would be interesting to study an organization which provides supportive resources to innovate. In this study, we intend to increase the understanding of AH’s IT innovation by unveiling its processes and capabilities. Supposedly, the more innovative a firm is, the more it possesses the capability to search for sustained competitive advantage (Wang & Ahmed, 2007).

To innovate, an organization has to have the ability to develop new products, services and/or markets with continual and incremental learning and experimenting from inside and outside of the organization (Aranda & Molina-Fernandez, 2002). Teece, et al. (1997) identified learning, reconfiguration and coordination/integration as the key processes in dynamic capabilities. In this case, learning refers to the ability to acquire, assimilate, transform and exploit existing knowledge to integrate and generate new knowledge (Aranda & Molina-Fernandez, 2002, Teece et al., 1997). It is acknowledged that knowledge and transfer of learning is of utmost importance for the implementation of system innovation (Metcalf, 1995)

Integration and coordination have been viewed as important elements that facilitate the deployment of resource reconfiguration. Such two elements could specify how different resources are integrated to create new ways of performing new sets of activities, for example, by transforming the existing resources into new potential resources that could correspond to volatile environments (Teece et al., 1997).

In this study, we aim to explore the process of innovative capabilities development and its deployment in resetting a new healthcare industry model (including structure and IT systems). In view of our research interests, our study builds on the innovative services and capability conceptualization and commonalities identification derived from the case study.

3 METHODOLOGY

We chose a case study approach to investigate our research interest, as it provides the opportunity for the researchers to explore contemporary events in the case company (Winter, 2003) with the empirical inquiry research on “how” and “why” questions (Lazonick & Prencipe, 2005). The main interest of this study is to discover how a service firm- Alexandra hospital (AH) innovates in a volatile environment to set a new healthcare industry model for Singapore.

In addition, the scarcity of research in this phenomenon (Amble & Palombarini, 1998, Aranda & Molina-Fernandez, 2002, Martin & Horne, 1993) has further substantiated the choice of research method in this study. In particular, an interpretive case study is an appropriate means of empirical inquiry (Lazonick & Prencipe, 2005) when the phenomena to be studied are complex and not easily separated from its original context (Yin, 1994). Thus, drawing from interpretive perspective, a process model of innovative capabilities development will be presented. This study will use the rich insights available in the case by examining the major events (narrative as instance) of the entire innovative response process in the volatile hospital environment.
One of the most dynamic Singaporean hospitals- AH was invited to participate in this study. This case study is particularly appropriate for the purpose of our study because it is cited by the Singapore Health Minister as the piloting prototyping center, which leads and sets the future standards of the Singapore Healthcare industry. In particular, the case study focuses on how innovative they are in deploying resources and capabilities in response to the government challenge to formulate a new explanatory model for the Singapore Healthcare industry.

This case study was conducted over a period of six months with the total of 19 face-to-face interviews with nine distinct interviewees, who were nominated by the Director of Operations according to their expertise and relevance to this case study. In addition, due to the size of the hospital and its limited resources, no more than ten staff were allocated to handle matters relating to all its IS projects. At the period of change, employees were bound with multiple tasks in their daily routine. Tasked with heavy workload, our informants were still willing to sacrifice their valuable time over the lunch or even after office hours with us over the interviews, demonstrating their passion for this project. Most interviews were tape-recorded and transcribed, with photos and additional notes taken to collect the best possible events of the case. Each interview session lasted between 1.5 to 3 hours, with informants ranging from nurses, IT specialists, doctors and top-management personnel having an average working experience of three years. We adopted a practical way of understanding textual data that was suggested by Klein & Myers (1999), i.e., via personal visits, emails and phone contacts to foster relationships among researchers and key informants. By doing so, it enables us to understand their rich depictions (Yin, 1994) underlining the meanings of their expressions (Hirschheim et al., 1991) so as to discover the core-case information that is necessary for the comprehending, analyzing and evaluating of the case study (Klein & Myers, 1999). To ensure the quality of data collected, we triangulated the data collected with other resources including empirical observations, follow-up email clarifications, along with about 230 MB softcopy documentations and archive records.

4  CASE STUDY

4.1  Case background

Alexandra Hospital (AH) was first established in 1938 as a British Military Hospital in Singapore. In 2000, AH had undergone a restructure and became a member of the National Healthcare Group (NHG), which is one of the two healthcare clusters in Singapore today. Since then, it has undertaken many initiatives to upgrade its major facilities with the desire to become the Mayo Clinic of Asia complimenting with the objective to provide patients with the hassle-free experience.

As one of the smallest hospitals in Singapore and a lack of government supports, AH was limited to use its scarce resources wisely to achieve its aim towards the upgrading of healthcare service standards in Singapore. That is to offer a better, faster, cheaper and safer-care service to patients on top of developing new business processes, and exploring innovative technologies. Ultimately, they decided to set up a new healthcare standard in Singapore by moving towards a Hassle-Free, patient-centric hospital to be located in the Northern Singapore in 2009.

4.2  Planning and preparation for an innovative breakthrough

To achieve the Hassle-Free Hospital vision, AH referred to the two best practices of US-Japanese Medicine/Healthcare systems (the Mayo Clinics (United States of America) and the Kameda Medical Centre hospitals (Japan)) to realize this vision. In addition, AH staff were sent to visit the Toyota plant in Japan and its workshop in Singapore to study how to transfer the learning into the hospital via the approach of thinking out-of-the-box mindset. Coupling with the impressive Japanese ‘Kaizen’ (continuous improvement) philosophy, a management team member:

“It worked for Toyota cars, so why not for patients at hospitals?”
Motivated with the inspiration of ‘Kaizen’ initiative of the Japanese, AH management team has to nurture the interest of the relevant stakeholders for their intellectual and emotional buy-in before kicking off a series of ‘Kaizen’ activities. Knowing that the change would bring tremendous impact to the hospital, creatively, the management team capitalises on its choice of having a passionate and committed IT personnel to spearhead the change. Through her social network and personal relations in the hospital on top of a supportive medical doctor husband, they managed to bridge the ideas and interests over informal social discussions and gatherings. News spread with the growth of interest and eventually the casual talks and gatherings were supported and attended by more specialised staff forming into an action-oriented team. Such a carefully crafted and planned management initiative has eventually and successfully nurtured the interest and cultivated the need for change in the hospital with minimum hassle.

Following the change, there are improvements whether significance or not were encouraged by the Director of Operation with the aim to save resources, time and add value to patients. According to a senior top-management:

“We believe in taking small steps in rapid succession approach, where we test new technology and deploy them if suitable”

Complementing the knowledge learned through the ‘Kaizen Flow’ (refers to Figure 1) initiated by the AH operational team with a small number of volunteered staff, including medical doctors and nurses, to participate in the change, the management incorporated the patients ‘wish list’ into the system. The initiative kicked off with scrutinizing the traditional work processes to identify bottlenecks, reviewing the existing patient-care processes and also re-examine patient services via the exploitation of technologies.

To further support the innovative change, all management staff including the operation team were assigned with a mobile phone and relocated to an office with no designated seats to encourage dynamic interactions, responsive and attentive management practices.

With the belief that innovation does not lie in the hands of the top-management, weekly system review session is organized to welcome suggestions from all levels. According to the Director of Projects in Operations:

“The contribution is from everyone because the process flow and the initiated changes have to happen on the ground then we (the management) help to “systematize” the components or the requirements.”

The hospital is filled with synergy to innovate. It is not uncommon to detect members of the management team gathering first-hand-on-site problems from the ground staff to understand the practicality and the feasibility of the change. Since then, the process of identifying and solving problems remains part and parcel of everyday work, more importantly, the project specialist 1 claimed:

“We feel that we should not get a system that changes our workflow. So we started to review our processes and request for an integrated IT system that can support us.”
On-going trainings were provided to staff to increase their awareness of the undergoing projects in relation to future changes to the hospital performances. In addition, the Andon board \(^1^\) is used to notify the new changes and act as a means to justify the improvements and benefits of the newly introduced workflows and healthcare system to staff.

### 4.3 Articulating and integrating resources

A series of continual learning and experimenting activities were carried out by the AH management to identify the necessary improvements before sourcing for further resources to accomplishing the task. The Chief Operating Officer alleged:

“We will continue to source for relevant technology in the market for our new facility, which will be ready in 2009”

To offer a more efficient and effective safer-care service to patients, an internal operational team was set up to examine the workflow of emergency-ward patients at the triage. After the review, three potential areas of improvement were identified— the physical layout of the department, the business workflow and the need for an integrated healthcare information system. The proposed improvements were then readily agreed by the management as prompt attention to emergency-department patients is vital and immediate change is targeted to upgrade its conditions and services.

After several intensive meetings on costs-and-benefits analysis, the management decided to invite external collaborators for strategic partnerships in improving the department physical layout, business workflow and the healthcare information systems. The assistant director of operations explained:

“We are a small hospital. So for some of the things we implemented, cost is important, as we don’t have the financial arm to bring in the big software players…”

To initiate the innovation, AH proposed the idea of setting up its integrated healthcare information system as a business case for a Singapore polytechnic college to take up as an assignment topic.

“To be able to solve an actual business problem is a precious opportunity for those polytechnic students…” Explained the Assistant director of operations

Through the ‘kaizen’ management approach and the workflow process review under the guidance of the head of emergency department, the physical layout of the emergency ward was innovatively improved over a short period of time by the internal operational team in ensuring that the point-to-point stations are positioned for the convenience of patients. After due considerations, AH made a bold and historical move to revamp the triage system by allocating a senior doctor (experience doctor), instead of deploying a junior doctor (trainee medical officer) to serve as the first point of contact with patients in the initial consultation process. Within the 90-day commission period, AH surprisingly received an endorsement for IS prototype from the Singapore polytechnic college regarding the newly improved workflow.

To put the IS prototype provided by the Singapore polytechnic college into practice, AH extended their partnerships with some other industry partners to pilot and test use various advanced technologies in its process of transforming into a hassle-free hospital. Along the exploration journey, Frontline Solutions (FS) Corporation implemented the prototype and played the role of system integrator to merge all the medical records and information, currently stored in various databases into a central-integrated-database system, on top of customizing information for efficient delivery across multiple devices. This collaboration has indeed created a win-win scenario for both parties. While AH benefits

---

\(^1^\) Japanese name for a visual production-control device (usually a lighted overhead display) that continuously shows changing status of the production line and sounds alerts if a problem is imminent (Business dictionary.com, 2008)
from the integrated healthcare system, FS Corporation enjoyed the privilege of advertising its success story with reference to AH experience. Based on the synergy and creativity manifested by AH staff, an innovative healthcare system is successfully designed from scratch.

4.4 Paving to the innovative hassle-free hospital service model

Taking a bold move to improve the workflows and innovatively develop a better system and workflow process from scratch is an achievement and a great leap forward for AH. The entire innovation service transformation was a great success: 1) AH has tremendously improved patients’ satisfaction, but also successfully expedited patient flow by 400% (from 22 to 70 patients attended per hour), 2) Reduced patient waiting time by 50%, and 3) The newly designed system can genuinely contributed to the recording and retention of patients’ medical history via the paperless administration in supporting the go-green campaign. By so doing, AH boosted its brand-name in the hospital industry by ranking consistently number one in the Ministry of Health Patient Satisfaction Surveys from 2004 to 2007.

5 DISCUSSION

Basing on the AH organization-wide innovation effort, a definition of a high-performing innovation organization suggested by Lawson & Samson (2001), we developed a process model of innovative capabilities development in a hospital-service settings, as showed in Figure 2. As suggested by our model, the development of innovative capabilities is a complex-helix process in articulating and integrating organizational learning and knowledge with tactical partnership and coordination to technologies, processes, services via organizational intelligence, strategic integration and operational articulation. In other words, the capability to innovate is developed through the capability to integrate and to articulate internal and external knowledge, resources and expertise to develop a new standard of products/services through aligning organizational intelligence to create a novel solution.

The conceptualization of strategic integration refers to critical managerial decision process in identifying, modifying and improving the internal and external resources. Operational articulation accounts for the operational activities of a firm in reconfiguring hospital standard, building stability, efficiency and profitability including the reduction costs, risks, and the achievement of economic of scale. Organizational intelligence is known as the capability to deal, process, interpret, manipulate and access information to resolve complicated problems/issues in a purposeful, goal-directed manner so that an organization can increase its potential to adapt and operate strategically (Glynn, 1996, Lawson & Samson, 2001). In other words, organizational intelligence refers to the capability for an organization to harness its preserved knowledge from different problems/issues, then deepen the knowledge through learning and further legitimize its perceived value according to its goal. AH focuses its initiative by learning from foreign hospitals in US and Japan and among its employees through the ‘kaizen flow’ which allows it to build the capability to identify new avenues for investigation and to innovate (Burgelman & Maidique, 1988).

Given that this model was inductively derived from the AH case study data, we present in a way how the existing literature corroborates with the model and how the model enriches our present understanding of innovative capabilities process model.
Figure 2: Process model of innovative capabilities development derived from the AH case study.

Figure 2 suggests that the process begins with an innovative aim to build a hassle-free hospital by providing faster, cheaper and safer-care services to patients. Such an idea was initiated by the top management since AH became a member of the NHG. To address the initiative, the hospital articulated knowledge from various sources by integrating information and expertise from both internal and external sources (Aranda & Molina-Fernandez, 2002) and transforming them into hospital-embedded knowledge or ‘organizational intelligence’. This is in line with previous research which suggests that a firm can develop superior capabilities through learning mechanisms, including continual experimentation, and the analysis of small mistakes (Eisenhardt & Martin, 2000). The ability of a hospital to integrate and articulate knowledge gained externally through the benchmarking of others’ best practices and technological opportunities coupled with the incorporation of its internal capabilities is critical to the success of an innovative change (Cohen & Levinthal, 1990; Anderson & Ejermo, 2005).

5.1 The evolution of capability to articulate in the process model of innovative capabilities development

The ability to articulate knowledge plays a critical role in innovation process (Andersson & Ejermo, 2005). Capability to articulate is generated from a series of continual learning and experimenting activities on new healthcare information technologies. These activities were identified by exploring how AH can transform its technology and process experiments to set a new standard in the healthcare industry. From the internal trial-and-error practices, the hospital devoted continual efforts to examine potential mismatches between the work processes and patients expectations. In addition, different experimentation methods were used to explore, investigate and collect stakeholders’ feedbacks for articulation purposes before formulating appropriate requests into a business project.

The next phase suggests that the innovative capability processes are involved to find out how continual experimenting and learning activities can evolve in response to feedback and stimuli from the external environment (for example, government and patients). Such external involvement is inevitable, especially, in the volatile industry, where government and practitioners are trying to explore and set right the blurring healthcare IT structure (including standard and systems). To innovate, hospitals ought to conduct operational articulation that identifies discrepancies or gaps for modifications and improvements from trial-and-error processes (Winter, 2003). The innovative achievement in reducing patient waiting time by 50 percent demonstrated that the capability to articulate is crucial for new ideas and knowledge to emerge.

5.2 The evolution of capability to integrate in the process model of innovative capabilities development

Innovative capability refers to a firm’s ability to integrate and reconfigure resources within (Eisenhardt & Martin, 2000), and outside the firm (Eisenhardt & Martin, 2000; Wang and Ahmed, 2007) for a novel idea. The development of innovation capability for AH hospital is achieved from the integration of resources and technologies, both internally and externally (Aranda & Molina-Fernandez, 2002). AH leverages on its internal resource integration via the coordination of cross functional team members (operational team member, doctors, nurses, top management, and patients) to draft AS-IS problems and TO-BE statements (refer to Figure 1) AH’s Kaizen flow to articulate lessons learned from experimental activities undertaken in the redesigning of the emergency ward workflow. Thus, the soft technologies such as knowledge, skills and expertise from members of the AH’s cross functional teams were critical for innovation in services to integrate and create streamlined business processes.
The notion that innovative capability is embedded in distinct ways of tactical partnership and coordination helps to explain how and why process and technological changes can have tremendous impact and influence on a firm’s services. This capability enables firms to select, develop exchange relationships with units and/or individuals beyond the traditional organization boundaries, such as networks and alliances of customers, suppliers, competitors and other non-market participants which are the key source of innovations (von Hipple, 1988; Hagedoorn & Duysters, 2004) to improve its services. Besides, the existing literature also suggested that the user-producer partnership and coordination would be fruitful for creating a complex and new-to-market innovation (Koch & Strotmann, 2008).

Not surprisingly, as evident in the case, AH demonstrated its capability to integrate knowledge and expertise from tactical partnership. In the initial stage, AH first partnered with a local polytechnic to develop a prototype healthcare information system that is articulated through experimenting continuous learning by the operation team members. After the test-run of the prototype to enrich the innovation process, AH invites other professional industry partners (e.g. FS Corporation) to take part in its success model. Thus, to innovation, a hospital ought to conduct strategic industry partnership that could integrate and leverage internal and external expertise, knowledge and technology (Teece, et al., 1997; Hagedoorn & Duysters, 2004) as evident in the case. All in all, conscientious and strategic integration planning not only tactically positions AH to innovate, but also benefit its partners to creating a win-win scenario.

6 CONCLUSIONS

The purpose of our paper is to conceptualize the generation of innovative capability from innovative services literature. A service innovation goes beyond pure technological innovation to a creative combination of hard and soft technologies and processes innovation. As evident in AH case, innovative capability development in service industry includes innovative workflow and system improvement. By drawing on the AH case study, we have developed an innovative capability model that depicts innovation as a complex process, whereby an organization’s key resources are decomposed into strategic integration and operational articulation that are reconfigured iteratively during the process of capability development.

The innovative capability development model in this paper has significant implications for researchers and practitioners. Previous studies contributed to the understanding of innovative capability as a subset of dynamic capability but not much information was provided as to how innovative capabilities can be developed and achieved. Our model proposes and clarifies that there are certain activities and processes involved in achieving innovative capability. This study contributes to providing some answers in responding to the lack of innovation studies in service firms (Amble & Palombarini, 1998, Aranda & Molina-Fernandez, 2002, Martin & Horne, 1993) and also serving the basis for further investigations.

In addition, this model proposes a process model that describes and analyzes the ways how innovative capabilities are implemented and managed during the innovation of healthcare IT in the hospital backdrop with the aim of setting a new standard for the healthcare industry in Singapore. This innovative development process model presented in this paper is grounded on the activities and processes revealed in the case study. Aspects of this model could be generalized into other cases of dynamic and volatile environments. However, more empirical work is necessary to test the applicability of the model in other settings.

For the practitioner perspectives, this study provides useful insights into the ways to manage innovation capability in order to design and create new industry standards. Taking the initiative to innovate itself into a hassle-free hospital, the AH case underscores the needs for hospital management to understand the concept of innovative capability as well as the approaches to build innovative capability by using organizational intelligence to cross reference with the strategic integration and
operational articulation processes. This study suggests that hospital management could adopt a more systematic innovation process by focusing more on soft technologies while handling services related challenges.

7 REFERENCES


EXAMINING THE EFFECT OF USER EXPECTATIONS ON SYSTEM USE ACTIVITY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0106.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Adoption, Behavioural theories, IT/IS management, Structural Equation Modeling</td>
</tr>
</tbody>
</table>
EXAMINING THE EFFECT OF USER EXPECTATIONS ON SYSTEM USE ACTIVITY

Kwahk, Kee-Young, Kookmin University, Graduate School of Business IT, 861-1, Jeongneung-dong, Seongbuk-gu, Seoul, Korea, 136-702, kykwahk@kookmin.ac.kr
Oh, Song-Woo, Kookmin University, Graduate School of Business IT, 861-1, Jeongneung-dong, Seongbuk-gu, Seoul, Korea, 136-702, osw8040@hotmail.com

Abstract

It has been argued that the promising benefits of Enterprise Systems (ES) implementations are overshadowed by their high failure rate. One of the commonly cited reasons for ES implementation failures is the end-user’s unwillingness to adopt or use systems. Considering that the appropriate management of expectations may play an important role in forming positive behavior toward newly implemented systems, this study examines the effect of outcome expectations on the system use activity in the mandatory-use context of information systems (IS) from the perspectives of Social Cognitive Theory and Coping Theory. Structural equation model analysis using LISREL 8.7 provides significant support for the proposed relationships. The empirical results suggest that outcome expectations and user satisfaction have positive effects on system use activity conceptualized by immersion, reinvention, and learning. Theoretical and practical implications of the study shed some light on how to improve system use activity in the mandatory-use context of IS.

Keywords: Outcome expectations, IS use activity, Social Cognitive Theory, Coping Theory.
1 INTRODUCTION

Although enterprise-wide information systems (IS) such as Enterprise Systems (ES) have received massive investment and have promised major strategic benefits, they have suffered from a high failure rate and difficulties in realizing the anticipated performance. The factors involved in ES failures are not limited to technical issues, but include various causes arising from the interactions among people, task, environment, and technology.

One of the commonly cited reasons for ES implementation failures is the end-user’s unwillingness to adopt or use systems. Despite the successfully developed system, the lack of positive user acceptance can lead to simple rote use rather than sophisticated system use, leading to dissatisfaction in the organization as the expected performance gained by the system’s introduction is not realized.

The introduction of new information technology (IT) generates numerous expected and unexpected consequences in the user’s environment. As these consequences are subjectively interpreted and understood by users, they respond in a variety of ways to the new system. According to the Coping Theory, the individual appraises what he/she can do to cope with disruptive events based on the potential consequences of events with subjective value, and then performs different cognitive and behavioral efforts (Lazarus and Folkman, 1984). This means that the user’s efforts to achieve the expected benefits or cope with the feared changes of the IT implementation depend on the user’s appraisal of the expected consequences of a variety of events which are induced by IT changes. Bandura (1986) also suggests that outcome expectations are a significant cognitive factor for individuals to control their behavior. Similarly, Thompson et al. (1991) proposed that the long-term consequences are important and effective factors related to new IT usage. Particularly, Chau (1996) showed that the long-term consequences (e.g., increasing the flexibility to change jobs, increasing the opportunities for more meaningful work, enhancing the user’s prestige among colleagues or professional peers) play a more important role in an individual’s IT acceptance by directing the individual’s focus on future benefits rather than short-term consequences (e.g., usefulness). In addition, successful system implementations are more closely related to successful management of user expectations than to user involvement or management support, and expectation management is considered highly important in determining the user’s attitude toward IT acceptance from the initial stage of system development through to the adaptation stage (Hoffer et al., 1999). These study results imply that inappropriate management of expectations may have a negative impact on successful system implementation by failing to meet expectations, despite the intention for a positive system contribution to the organization.

Prior IS research on the new IT adoption has been conducted in the context of voluntary adoption of the new IT by users and either the usage or the intention to use is exploited as a dependent variable. However, most of the current enterprise-wide IS are used in non-voluntary contexts where users must use the system to perform their jobs. Therefore, debate has continued over whether usage or intention to use is an appropriate surrogate variable for actual system use behavior (Rawstorne et al., 1998, 2000). Moreover, the concept of frequency or duration used as a surrogate of system usage does not appear to be appropriate in a mandatory context. Therefore, some researchers suggest that a new concept as a dependent variable should be considered instead of the concept of intention to use or simple system usage (Barki et al., 2007; Burton-Jones and Straub, 2006).

This study aims to understand the role of outcome expectations for successful IS use and suggest a new surrogate variable for system use, based on the Social Cognitive Theory (SCT) and the Coping Theory. The study is conducted with two research objectives. First, from the IS acceptance perspective, we examine how outcome expectations affect the IS use activity of the organization members. Second, we suggest and examine a new outcome variable describing the actual acceptance behavior of the user in the mandatory-use contexts.
2 THEORETICAL BACKGROUND

2.1 Social cognitive theory (SCT)

According to SCT, environmental influences such as social pressures or unique situations, cognitive affects and other personal factors, including personality and behavior, are reciprocally determined as interacting determinants that influence each other. This relationship is characterized by “triadic reciprocal determinism” (Bandura, 1986). It is not that the individual simply reacts to the environment events, but that the individual acts for positively creating and changing his/her environments. Furthermore, individual cognition and the way of changing his/her environment are affected by either positive or negative feedback from behavior; that is, the determined behavior influences the individual’s cognitive process and environment. While SCT can approach many dimensions, this research is particularly concerned with the role of outcome expectations that form and control individual behavior. Outcome expectation that reflects expectations about the level of consequences from their behavior has been considered a significant factor predicting future behavior. While individuals are more likely to engage in behavior with expected favorable or positive consequences, outcome expectations affect their behaviors by letting them avoid expected behaviors that result in negative consequences. Therefore, outcome expectation is considered to play an important role in explaining individual behaviors in the IS research from the SCT perspective (Compeau and Higgins, 1995). Expectations about the net benefits of future IS use are continuously modified based on the IS use that is experienced, and modified expectations change the level of IS use, which further changes the perceptions of the IS performance again. Several recent studies have reported that outcome expectations affect such dependent variables as computer use, knowledge sharing, and user’s organizational commitment (e.g., Stone and Henry, 2003). These research results imply that outcome expectation is a very important cognitive factor to control the user’s behavior. Thus, outcome expectation that controls behavior in the SCT frame can play an important role in predicting and explaining the acceptance behavior of various users in the IS use context.

2.2 Coping theory

Lazarus and Folkman (1984, p.141) defined coping as “the cognitive and behavioral efforts exerted to manage specific external and/or internal demands that are appraised as taxing or exceeding the resources of the person.” It deals with the adaptation acts that the individual performs in response to stimulation or disruptive events and circumstances in his/her environment. That is, the individual either adapts to the subjective meaning of the event or makes coping efforts to seek additional information and evidence for the purpose of changing or avoiding the situation (Lazarus and Folkman, 1984). Individuals apply two subordinate processes (appraisal and coping efforts) to cope with the disruptive events (Lazarus and Folkman, 1984). First, the individual evaluates the potential consequences of an event (primary appraisal and secondary appraisal). The individual evaluates specific events in relation to his/her welfare, i.e., whether it is related to him/her or it is beneficial and positive to him/her. While the primary appraisal focuses on the events that give some kind of meaning to the individual, the secondary appraisal evaluates available social and personal resources required to cope with the events and considers available behavioral options (Lazarus and Folkman, 1984). Second, the individual makes problem-focused coping and emotion-focused coping efforts to deal with the situation (Lazarus and Folkman, 1984). The former aims to solve the problem (e.g., disruptive issue) by changing the environmental pressures, barriers, and resources or by changing oneself (e.g., learning new skills or procedures and finding new channel of satisfaction). The latter aims to reduce or manage the negative emotional distress. It doesn’t change the situation itself, but changes one’s perception of the situation (Lazarus and Folkman, 1984). The kind of efforts that the individual makes depends on how the given situation is appraised.
In an IS research drawing on the Coping Theory, Beaudry and Pinsonneault (2005, p.496) introduced adaptation as a similar concept with coping and defined it as “the cognitive and behavioral efforts exerted by users to manage specific consequences associated with a significant IT event that occurs in their work environment.” The user assesses the expected consequences of the IT event and how the changes affect himself/herself and his/her work at the stage of primary appraisal, so that the user will be aware of the opportunity or threat about the potential consequences of the significant IT event (Beaudry and Pinsonneault, 2005). Such potential consequences are a very similar concept to the long- and short-term consequences of Thompson et al. (1991) and to the outcome expectations of Compeau and Higgins (1995). For instance, if a user evaluates the fitness between work and IS to be high, the changing of the new IS would be recognized as an opportunity, otherwise it would be recognized as a threat. Based on such primary appraisal, the user appraises how much he/she modifies work, controls himself/herself with the new environment (Lazarus and Folkman, 1984), and controls the characteristics and functions of new IT at the stage of secondary appraisal (Beaudry and Pinsonneault, 2005). After such cognitive appraisal, the user makes emotion- and problem-focused adaptation that is similar to the coping efforts of the Coping Theory. Problem-focused adaptation aims at dealing with the issues of IT changes directly by adapting oneself, work, and technology (Rice and Rogers, 1980). Self-adaptation involves adjusting one’s habits, learning new skills, and adjusting commitment to the work to meet the requirements of the technology. Adapting the work changes procedures and routines, and adapting the technology changes functions and features. Emotion-focused adaptation aims at changing one’s perception of the consequences of the newly introduced IT or reducing emotional distress. Emotion-focused adaptation is oriented toward avoiding reality though self-deception or psychologically minimizing negative consequences, selective attention, and positive comparison to other users (Beaudry and Pinsonneault, 2005). It is especially emphasized that the user’s adaptation efforts depend on the user’s appraisal. In other words, the user’s adaptation efforts are changed by cognitive evaluation of the IT event (i.e., the awareness and expectation of potential consequences). Thus, the Coping Theory indicates that outcome expectation plays a significant role in explaining the behaviors of various users in the IS environment.

2.3 Conceptualization of system use activity

It has been argued that predicting behavior based on intention is inappropriate in a mandatory-use environment (Rawstorne et al. 1998, 2000). The link between intention and behavior only applies when the behavior is under volitional control of the person (Fishbein and Ajzen, 1980). That is, the connection between behavioral intention and actual behavior lacks consistency in mandatory contexts.

The mandatory IS use environment means that the user has no choice but to use the system to perform his/her work, so that prior measurement of system usage based on voluntary contexts is inappropriate. The effectiveness of technology usage is decided by the organizations’ aims and objectives in mandatory contexts (Adamson and Shine, 2003). Thus, simple usage behavior of IS may be inappropriate as an indicator of success for the system implementation. The behavior of the user is more complicated than simple IT use in the organizational setting (Adamson and Shine, 2003). Thus, we propose positive “system use activity” related to IS use as a surrogate variable for new IS acceptance instead of intention to use or mechanical use behavior under the premise of mandatory environment. We conceptualize system use activity into three dimensions based on the studies of Burton-Jones and Straub (2006) and Barki et al. (2007): user’s immersion, reinvention, and learning.

First, Burton-Jones and Straub (2006) proposed that system usage is an activity among a user, a system, and a task, and they defined the individual user’s level of system usage as the extent of his/her employment of one or more features of a system to perform a task. To measure the activity between the user and the system in their research, they adopted focused immersion, which is among the five dimensions in the cognitive absorption construct by Agarwal and Karahanna’s (2000). Immersion means the extent to which the individual can set aside other concerns related to their ability to focus on
and perform the specific task required. The user’s immersion activity related to system use activity can be defined as the extent to which the user concentrates on the task while using the system.

Second, Barki et al. (2007) conceptualized IS use-related activity based on task-technology fit and activity theory. IS use-related activity was classified into three behaviors: technology interaction behavior, task-technology adaptation behavior, and individual adaptation behavior. They empirically verified the effects of these activities on the individual and organizational benefits. Among them, task-technology adaptation behavior includes all behaviors like modifying or changing IT and determining how it will be used in the organization. This concept is based on Rice and Rogers’ (1980) notion of reinvention, which reflects the extent to which an adopter changes an innovation following its original development. That is, reinvention can be measured as the amount of effort the IS user puts into the enhancement of the fit between the task and the system to improve his/her performance.

Third, individual adaptation behaviors suggested by Barki et al. (2007) represent behavioral changes that the individual makes to himself/herself in order to adapt to IT. Such self-modification behaviors include learning activity and interaction between individuals and the IT system (Beaudry and Pinsonneault, 2005). Information acquisition activities reflect the coping strategy of users to reduce uncertainties in their work. As users learn how to use the new IT, they can apply new ways to perform their tasks and exchange information with each other. Barki et al. (2007) categorized individual adaptation behaviors into communication behaviors (interactions with other users or IS professionals to exchange information about an IT) and independent exploration behaviors (information search behaviors undertaken independently to improve one’s knowledge and mastery of an IT). So learning can be defined as the degree to which users communicate with each other and search for information in order to improve their knowledge and IT skill.

3 RESEARCH MODEL AND HYPOTHESES

The research model is shown in figure 1, which is based on previously discussed research motivation and theoretical background. We here discuss the conceptual background about the hypotheses in the research model.

![Figure 1. Research model.](image)

According to Bhattacherjee (2001), the comparison evaluation between expectation and system performance is a continuous process, so users’ expectation can change while using a system. Szajna and Scamell (1993) suggested that unrealistically high or low user expectations change into realistic levels, and that low expectations are associated with low satisfaction of users based on the Cognitive

Proceedings ECIS 2009
Dissonance Theory (CDT; Festinger, 1957). According to CDT, when cognitive dissonance occurs, the perceived performance is assimilated toward the initial expectation. It has been reported that consumers’ expectation based on CDT has a positive influence on perceived product performance in marketing research (Szymanski and Henard, 2001). Kim et al. (2004) suggested that the user’s high expectation highly assesses system performance in IS research. Hence, we hypothesize the following:

H1a: Outcome expectations-performance has a positive effect on perceived system quality.
H1b: Outcome expectations-performance has a positive effect on perceived information quality.
H2a: Outcome expectations-personal has a positive effect on perceived system quality.
H2b: Outcome expectations-personal has a positive effect on perceived information quality.

Outcome expectation is considered an important cognitive factor in determining individual behavior, according to SCT (Bandura, 1986). Individuals do not simply respond to the environment, but behave positively to create and change it. Similarly, when the user primarily assesses the expected personal and professional consequences of a newly-introduced IT event in the organization as an opportunity, he/she recognizes and exercises control over the situation, and makes problem-focused adaptation efforts (Beaudry and Pinsonneault, 2005). That is, the user tries to increase his/her knowledge of and improve on his/her ability in using the system and makes additional efforts to communicate with other users in order to seize opportunities offered by IT changes. This behavior consequently modifies prior work procedures and even changes the function of the system (Beaudry and Pinsonneault, 2005).

Thompson et al. (1991) also suggested that perceived consequences, which are composed of three dimensions (complexity, job fit, and long-term consequences of use), strongly influence the degree of computer utilization. The most valuable effect proved to be job fit, which is defined as an individual’s belief that the system enhances his/her job performance. Compeau and Higgins (1995) also reported that outcome expectation has a significant impact on usage by demonstrating that both outcome expectations-performance and outcome expectations-personal affect usage. An individual’s experience, environment, and use behavior affect his/her cognitive factor as suggested by SCT, which then feeds into one’s behavior again by forming a feedback mechanism. Hence, we hypothesize the following:

H3: Outcome expectations-performance has a positive effect on system use activity.
H4: Outcome expectations-personal has a positive effect on system use activity.

The IS success model suggested by DeLone and McLean (1992) and Rai et al. (2002) reported that user satisfaction is affected by the quality of system and information. It was also verified that IS performance is positively associated with user satisfaction. According to Roca et al. (2006), system quality and information quality have positive effects on satisfaction of Web and e-learning users in which perceived system performance is measured in terms of quality. An empirical study on Web-based decision support systems also verified that system quality and information quality have significant impacts on decision-making satisfaction (Bharatia and Chaudhury, 2004). Incidentally, this relationship has already been proved in many studies in the IS field including the IS success model. Hence, we hypothesize the following:

H5: Perceived system quality has a positive effect on user satisfaction.
H6: Perceived information quality has a positive effect on user satisfaction.

Individual user satisfaction with IS use reflects a status of either positive (satisfied) or negative (dissatisfied) feeling (Bhattacharjee, 2001). It can be an impression of delight or disappointment often derived from a comparison of perceived IS performance and its expectation (Bhattacharjee, 2001). It has been widely recognized that an affect associated with a user (e.g., finding delightful about the system) is a significant factor for system use behaviors (Compeau and Higgins, 1995) and that an individual’s feeling for particular behaviors significantly influences that behaviors (Bandura, 1986). According to a study on innovation acceptance of system developers (Chau, 1996), although system developers are mandated to use the innovation (as the adaptation decision about the software development tool is made at an organizational level), whether or not the user enjoys using it (i.e., satisfaction with the innovation) plays a critically important role in productivity. When the
introduction of new system forces users to change tasks, techniques, attitude, and efforts, user satisfaction can play an important role as a way to reduce the resistance of the users, particularly when the system is closely associated with the users’ work (Melone, 1990). We can infer that the user who is satisfied with system use may spontaneously apply the system to his/her work to achieve benefits provided by the new system. In doing so, he/she will exert great effort to learn the new system, modify his/her work procedures, and concentrate on the task during system usage, thereby reducing the negative aspects of the new system. Hence, we hypothesize the following:

H7: User satisfaction has a positive effect on system use activity.

4 RESEARCH METHODOLOGY

We performed a questionnaire survey at the individual level. The data used to test this research model were collected from ERP system users who use the system to perform their organizational tasks. We selected ERP system users because they are supposed to use the system in the mandatory context. Four hundred questionnaires were distributed by mail and 225 responses were received. After the exclusion of 17 incomplete responses, 208 usable respondents were analyzed in the study. The validated questions used to make measurements in this study were mainly adopted from the relevant literature whenever possible and were modified for adjustment to the current research context. All items were measured using a seven-point, Likert-type scale with answers ranging from strongly disagree to strongly agree. The instrument was reviewed by three IS researchers and feedback results were strongly disagree to agree. The instrument was reviewed by three IS researchers and feedback results were reflected in the question items. Measurement instrument used in this study is shown in table 1.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Item</th>
<th>Measure</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome</td>
<td>If I use ERP system, I expect</td>
<td>If I use ERP system, I expect</td>
<td>Compeau and Higgins 1995</td>
</tr>
<tr>
<td>expectations-performance</td>
<td>PFO1 I will increase my effectiveness on the job.</td>
<td>PFO1 I will increase my effectiveness on the job.</td>
<td></td>
</tr>
<tr>
<td>(PFO)</td>
<td>PFO2 I will spend less time on routine job task.</td>
<td>PFO2 I will spend less time on routine job task.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PFO3 I will increase usefulness of performing the task.</td>
<td>PFO3 I will increase usefulness of performing the task.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PFO4 I will increase the quality of output of my job.</td>
<td>PFO4 I will increase the quality of output of my job.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PFO5 I will increase the quantity of output for the same amount of effort.</td>
<td>PFO5 I will increase the quantity of output for the same amount of effort.</td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td>If I use ERP system, I expect</td>
<td>PSO1 My co-workers will perceive me as competent.</td>
<td>Compeau and Higgins 1995</td>
</tr>
<tr>
<td>expectations-personal</td>
<td>PSO2 My sense of accomplishment will be increased.</td>
<td>PSO2 My sense of accomplishment will be increased.</td>
<td></td>
</tr>
<tr>
<td>(PSO)</td>
<td>PSO3 I will be seen as higher in status among colleagues.</td>
<td>PSO3 I will be seen as higher in status among colleagues.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PSO4 My reputation will be better among my colleagues.</td>
<td>PSO4 My reputation will be better among my colleagues.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PSO5 My opportunities of getting a promotion or changing my job will be increased.</td>
<td>PSO5 My opportunities of getting a promotion or changing my job will be increased.</td>
<td></td>
</tr>
<tr>
<td>System</td>
<td>SYQ1 ERP system is possible to exchange data with other systems.</td>
<td>SYQ1 ERP system is possible to exchange data with other systems.</td>
<td>Bailey and Pearson 1983, Wixom and Todd 2005</td>
</tr>
<tr>
<td>quality (SYQ)</td>
<td>SYQ2 ERP system stably operates without interruption or errors.</td>
<td>SYQ2 ERP system stably operates without interruption or errors.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SYQ3 ERP system is flexible by new conditions, processes, structure of organization, and circumstances.</td>
<td>SYQ3 ERP system is flexible by new conditions, processes, structure of organization, and circumstances.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SYQ4 ERP system returns answers to my requests quickly.</td>
<td>SYQ4 ERP system returns answers to my requests quickly.</td>
<td>Bailey and Pearson 1983, Rai et al. 2002</td>
</tr>
<tr>
<td>Information</td>
<td>IFQ1 ERP system provides the precise information you need.</td>
<td>IFQ1 ERP system provides the precise information you need.</td>
<td>Bailey and Pearson 1983, Rai et al. 2002</td>
</tr>
<tr>
<td>quality (IFQ)</td>
<td>IFQ2 ERP system provides sufficient information to enable you need.</td>
<td>IFQ2 ERP system provides sufficient information to enable you need.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IFQ3 ERP system provides up-to-date information.</td>
<td>IFQ3 ERP system provides up-to-date information.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IFQ4 The provided ERP system’s output is presented in a useful format.</td>
<td>IFQ4 The provided ERP system’s output is presented in a useful format.</td>
<td></td>
</tr>
<tr>
<td>User</td>
<td>How you feel about your overall experience of ERP system use.</td>
<td>How you feel about your overall experience of ERP system use.</td>
<td>Bhattacharjee 2001</td>
</tr>
<tr>
<td>satisfaction (USF)</td>
<td>USF1 Very dissatisfied/Very satisfied</td>
<td>USF1 Very dissatisfied/Very satisfied</td>
<td></td>
</tr>
<tr>
<td></td>
<td>USF2 Very displeased/Very pleased</td>
<td>USF2 Very displeased/Very pleased</td>
<td></td>
</tr>
<tr>
<td></td>
<td>USF3 Very frustrated/Very contented</td>
<td>USF3 Very frustrated/Very contented</td>
<td></td>
</tr>
<tr>
<td></td>
<td>USF4 Absolutely terrible/Absolutely delighted</td>
<td>USF4 Absolutely terrible/Absolutely delighted</td>
<td></td>
</tr>
<tr>
<td>Immersion</td>
<td>When I use ERP system, I am able to block out all other distractions.</td>
<td>When I use ERP system, I am able to block out all other distractions.</td>
<td>Agarwal and Karahanna 2000, Burton-Jones and Straub 2006</td>
</tr>
<tr>
<td>(IME)</td>
<td>IME1 I feel totally immersed in what I am doing.</td>
<td>IME1 I feel totally immersed in what I am doing.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IME2 I feel completely absorbed in what I am doing.</td>
<td>IME2 I feel completely absorbed in what I am doing.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IME3 My attention does not get diverted very easily.</td>
<td>IME3 My attention does not get diverted very easily.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IME4 I am able to block out all other distractions.</td>
<td>IME4 I am able to block out all other distractions.</td>
<td></td>
</tr>
<tr>
<td>Reinvention</td>
<td>When I use ERP system, I exert myself to</td>
<td>When I use ERP system, I exert myself to</td>
<td>Barki et al. 2007, Rice and Rogers 1980</td>
</tr>
<tr>
<td>(RIV)</td>
<td>RIV1 Find improvements in the system’s functionalities and interface.</td>
<td>RIV1 Find improvements in the system’s functionalities and interface.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RIV2 Modify my tasks so that it better fits the system.</td>
<td>RIV2 Modify my tasks so that it better fits the system.</td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Measurement instrument

<table>
<thead>
<tr>
<th>Learning (LEN)</th>
<th>Items</th>
<th>Factor loading</th>
<th>CR</th>
<th>AVE</th>
<th>Cronbach’s α</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEN1</td>
<td>I exert myself to have opportunities to communicate with colleagues or specialists in order to better understand how ERP system operates on my own initiative.</td>
<td>0.857</td>
<td>0.894</td>
<td>0.681</td>
<td>0.886</td>
</tr>
<tr>
<td>LEN2</td>
<td>I exert myself to increase my knowledge and mastery of ERP system on my own initiative.</td>
<td>0.877</td>
<td>0.955</td>
<td>0.808</td>
<td>0.954</td>
</tr>
<tr>
<td>LEN3</td>
<td>I exert myself to learn ERP system on my own initiative.</td>
<td>0.681</td>
<td>0.849</td>
<td>0.652</td>
<td>0.848</td>
</tr>
<tr>
<td>LEN4</td>
<td>I invest much effort in order to better use ERP system.</td>
<td>0.840</td>
<td>0.849</td>
<td>0.652</td>
<td>0.848</td>
</tr>
<tr>
<td>PFO1</td>
<td>0.857</td>
<td>0.894</td>
<td>0.681</td>
<td>0.886</td>
<td></td>
</tr>
<tr>
<td>PFO2</td>
<td>0.870</td>
<td>0.894</td>
<td>0.681</td>
<td>0.886</td>
<td></td>
</tr>
<tr>
<td>PFO3</td>
<td>0.877</td>
<td>0.894</td>
<td>0.681</td>
<td>0.886</td>
<td></td>
</tr>
<tr>
<td>PFO5</td>
<td>0.681</td>
<td>0.894</td>
<td>0.681</td>
<td>0.886</td>
<td></td>
</tr>
<tr>
<td>PSO1</td>
<td>0.899</td>
<td>0.955</td>
<td>0.808</td>
<td>0.954</td>
<td></td>
</tr>
<tr>
<td>PSO2</td>
<td>0.870</td>
<td>0.955</td>
<td>0.808</td>
<td>0.954</td>
<td></td>
</tr>
<tr>
<td>PSO3</td>
<td>0.943</td>
<td>0.955</td>
<td>0.808</td>
<td>0.954</td>
<td></td>
</tr>
<tr>
<td>PSO4</td>
<td>0.919</td>
<td>0.955</td>
<td>0.808</td>
<td>0.954</td>
<td></td>
</tr>
<tr>
<td>PSO5</td>
<td>0.860</td>
<td>0.955</td>
<td>0.808</td>
<td>0.954</td>
<td></td>
</tr>
<tr>
<td>SYQ2</td>
<td>0.783</td>
<td>0.849</td>
<td>0.652</td>
<td>0.848</td>
<td></td>
</tr>
<tr>
<td>SYQ3</td>
<td>0.844</td>
<td>0.849</td>
<td>0.652</td>
<td>0.848</td>
<td></td>
</tr>
<tr>
<td>SYQ4</td>
<td>0.795</td>
<td>0.849</td>
<td>0.652</td>
<td>0.848</td>
<td></td>
</tr>
<tr>
<td>IFQ2</td>
<td>0.809</td>
<td>0.853</td>
<td>0.658</td>
<td>0.852</td>
<td></td>
</tr>
<tr>
<td>IFQ3</td>
<td>0.797</td>
<td>0.853</td>
<td>0.658</td>
<td>0.852</td>
<td></td>
</tr>
<tr>
<td>IFQ4</td>
<td>0.828</td>
<td>0.853</td>
<td>0.658</td>
<td>0.852</td>
<td></td>
</tr>
<tr>
<td>IME1</td>
<td>0.838</td>
<td>0.940</td>
<td>0.797</td>
<td>0.937</td>
<td></td>
</tr>
<tr>
<td>IME2</td>
<td>0.952</td>
<td>0.940</td>
<td>0.797</td>
<td>0.937</td>
<td></td>
</tr>
<tr>
<td>IME3</td>
<td>0.951</td>
<td>0.940</td>
<td>0.797</td>
<td>0.937</td>
<td></td>
</tr>
<tr>
<td>IME4</td>
<td>0.823</td>
<td>0.940</td>
<td>0.797</td>
<td>0.937</td>
<td></td>
</tr>
<tr>
<td>RIV1</td>
<td>0.857</td>
<td>0.940</td>
<td>0.821</td>
<td>0.948</td>
<td></td>
</tr>
</tbody>
</table>

5 DATA ANALYSIS AND RESULTS

5.1 Confirmatory factor analysis

First, we checked the unidimensionality of each construct to test the convergent validity. Following proposed methodological procedure, we conducted revisions on the measurement model by removing one at a time, those items which shared a high degree of residual variance with other items. The measurement model showed a suitable fitness level after removing five items (PFO4, SYQ1, IFQ1, LEN1, and USF4). The normed $\chi^2/n$ to degree of freedom was 1.532, which was a good fit as it was below the desired maximum cut-off of 3.0. The root mean square error of approximation (RMSEA) also showed a good fit at 0.051, which was below the maximum desired cut-off of 0.06. The root mean square residual (RMR) was 0.044, which was lower than the desired maximum cut-off of 0.05 (Hair et al., 1998). The goodness-of-fit index (GFI) was 0.849, which was above the recommended threshold of 0.8 (Hair et al., 1998). The other fit indices were also satisfactory: comparative fit index (CFI)=0.989, normed fit index (NFI)=0.972, and non-normed fit index (NNFI)=0.987. Thus, we concluded that the measurement model had adequate overall fitness.

As shown in table 2, most standardized path loadings were greater than 0.7 and significant ($t$-value > 1.96) with only one exception: PFO5. Composite reliability (CR) and the Cronbach’s $\alpha$ for all constructs were larger than 0.7. All the average variance extracted (AVE) for each factor exceeded 0.5. Thus, the constructs used in this study had convergent validity.
Next, we accessed the discriminant validity. As shown in table 3, the square root of AVE for each construct exceeded the correlations between the construct and other constructs. Hence, the constructs used in this study had discriminant validity. In addition, we examined common method bias (CMB) using Herman’s single factor test (Podsakoff et al., 2003). A significant difference in the $\chi^2$ statistic of the original ($\chi^2=534.6, \text{df}=349$) and single factor model ($\chi^2=3814.4, \text{df}=377$) revealed that this research was not affected by CMB.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Mean(SD)</th>
<th>PFO</th>
<th>PSO</th>
<th>SYQ</th>
<th>IFQ</th>
<th>IME</th>
<th>RIV</th>
<th>LEN</th>
<th>USF</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFO</td>
<td>5.140(1.118)</td>
<td>0.825</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSO</td>
<td>4.022(1.347)</td>
<td>0.420</td>
<td>0.899</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYQ</td>
<td>4.397(1.365)</td>
<td>0.474</td>
<td>0.549</td>
<td>0.807</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IFQ</td>
<td>4.780(1.251)</td>
<td>0.557</td>
<td>0.509</td>
<td>0.790</td>
<td>0.811</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IME</td>
<td>4.337(1.213)</td>
<td>0.504</td>
<td>0.696</td>
<td>0.505</td>
<td>0.687</td>
<td>0.893</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RIV</td>
<td>4.435(1.259)</td>
<td>0.382</td>
<td>0.580</td>
<td>0.430</td>
<td>0.496</td>
<td>0.658</td>
<td>0.906</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEN</td>
<td>4.697(1.298)</td>
<td>0.453</td>
<td>0.562</td>
<td>0.450</td>
<td>0.552</td>
<td>0.630</td>
<td>0.660</td>
<td>0.925</td>
<td></td>
</tr>
<tr>
<td>USF</td>
<td>4.863(1.184)</td>
<td>0.609</td>
<td>0.641</td>
<td>0.723</td>
<td>0.710</td>
<td>0.699</td>
<td>0.549</td>
<td>0.615</td>
<td>0.873</td>
</tr>
</tbody>
</table>

Table 3. Results of discriminant validity testing.

5.2 Second-order confirmatory factor analysis

System use activity consists of immersion, reinvention, and learning activity. Thus, the second-order factor, system use activity, was assumed to be indirectly measured through the lower-order factors. The results of second-order confirmatory factor analysis showed a satisfactory fit level (Normed $\chi^2=1.671$, RMR=0.025, GFI=0.943, NFI=0.984). Therefore, the first-order factors and the second-order factor were verified to have strong relations. We carried out item parceling by averaging the first-order factors to measure the latent variable of system use activity, the second-order factor, in the structural model.

5.3 Hypothesis testing

The test of hypotheses using the structural model was performed with LISREL 8.7. We first evaluated the goodness of fit indices. Considering the overall fit indices ($\chi^2/\text{df}=2.07$, GFI=0.846, CFI=0.977, and NFI=0.960), we concluded that the research model provided a good model fit to the data. The standardized path coefficients for the research model are presented in figure 2. All three variables (outcome expectations-performance, outcome expectations-personal, and user satisfaction) were significantly related to system use activity and explained 72.7% of the variance in system use activity: outcome expectations-performance ($\beta=0.125$, p<0.05), outcome expectations-personal ($\beta=0.457$, p<0.01), and user satisfaction ($\beta=0.434$, p<0.01). Outcome expectations-performance and outcome expectations-personal showed significant relationships with perceived system performance (system quality and information quality) and explained 44.1% and 47.3% of the variance respectively: outcome expectations-performance ($\beta=0.338$, p<0.01; $\beta=0.453$, p<0.01) and outcome expectations-
personal ($\beta=0.447$, $p<0.01$; $\beta=0.360$, $p<0.01$). Perceived system performance had a significant relationship with user satisfaction and explained 61.8% of the variance in user satisfaction: system quality ($\beta=0.466$, $p<0.01$) and information quality ($\beta=0.457$, $p<0.01$). These results supported all hypotheses.

Figure 2. Results of research model (*$p<0.05$; **$p<0.01$; One-tail test).

6 DISCUSSION AND IMPLICATIONS

The study results suggested that as the user uses the system, both the expected usefulness related with job (outcome expectations-performance) and personally expected usefulness (outcome expectations-personal; e.g., recognizing my competence, accomplishment, or obtaining a promotion) have a significant effect on positive system use activity of users. In contrary to Compeau and Higgins (1995) who reported that outcome expectations-performance has a more significant impact on computer usage than outcome expectations-personal, this study showed that improving job performance through IS may not be a strong motivation for IS use on the individual level in the mandatory-use context. Rather, in this context, a good assessment of outcome expectations-personal like image as a constant driving force is an important factor in positive system activity beyond a simplistic view of system use such as the amount or frequency of IS use. Similar to the present study result, Tampoe (1993) suggested personal growth as a knowledge worker’s motivation, and argued that, in particular, professional and personal accomplishments are seen to be the driving force of ongoing motivation. Two types of outcome expectations and user satisfaction explained 70.5% of the variance in system use activity. This might imply that the user assesses the potential consequences derived from IT changes as an opportunity. This translates to positive system use activity (e.g., concentrating on the job, efforts to fit between task and system, and efforts to understand the system activity) as users try to cope with these kinds of challenge in terms of the Coping Theory. The results of this research also supported the importance of user satisfaction as a way to reduce the resistance of the user, when the system is closely associated with the user’s work. The cognitive process from user’s outcome expectations to positive system use activity shows that the user’s high expectations induce a high evaluation of perceived system performance (system quality and information quality), which is consistent with CDT. In addition, the results of this research supported those of existing IS research showing that perceived system quality and information quality have a significant effect on user satisfaction with system use.
The findings of the present study have several implications for theory and practice. From the theoretical perspective, first, we developed a theoretical model to explain the effect of outcome expectations on IS use activity of organization members. This proposed theoretical model provides an understanding of the role of outcome expectations for successful IS adaptation, especially outcome expectations-personal as intrinsic motivation in the perspective of expectancy management. In addition, the suggested research model can present the cognitive process from outcome expectations to positive IS use activity through the IS success factors (system quality, information quality, and user satisfaction). Second, this research proposed a new variable composed of three dimensions (immersion, reinvention, and learning) as a surrogate variable of system use or intention to use in mandatory IS use context. This may reflect a user’s actual system acceptance behavior in the condition when the introduced system use is decided by management. Most of the existing IS studies suffer limitations in representing various responses of users as they have focused the user’s system acceptance behavior simply on the system use or the intention to use without considering the IS use environment (voluntary or non-voluntary environment). Thus, the research results facilitate an understanding of the varied responses and behavior of users in IS acceptance studies by suggesting a new dependent variable. Third, the study finding implies that user satisfaction and positive system use behavior are significant factors in individual impact to achieve the goals of introducing a new system in an organization. Igbaria and Tan (1997) suggested that user satisfaction is a more important factor in individual impact than system usage and the influence of user satisfaction on an individual’s performance is partially mediated by system usage. A satisfied user shows a high-level of system use activity because he/she does not simply use the system, but tries hard to use it better. Eventually, the user shows a high-level individual performance.

From the practical perspective, the study results support two findings. First, the results represent how outcome expectations enforce the system use activity of users in the perspective of expectancy management. Szajna and Scamell (1993) showed that expectation can change as a user uses an IS, i.e., the user can have realistic expectations and form a positive user attitude about the newly introduced IS while performing appropriate expectations management. In other words, positive system use activity can be reinforced, thereby supporting successful implementation of the system through the user’s positive employment of the system and appropriate expectations management. This is required for systematic evaluation and reward systems. Systematic evaluation and reward are the easiest ways for the management to inform the members of the organization of worthwhile activities and perceived results from these activities. The management should focus on preparing a systematic plan to properly manage job performance and expectancy at the individual level. Second, outcome expectations-personal is enhanced when an individual’s colleagues encourage system use. A user will expect, among other things, that his/her behavior will be recognized and will please other members of the organization (Compeau and Higgins, 1995). The results of Compeau and Higgins’s (1995) study show that a user’s image can be enhanced as the user follows his/her colleagues’ requirements or recommendations to develop the initiative to use the system. Consequently, when the user has opportunities to discuss aspects of the IS changes with his/her colleagues, it may lead to positive system use activity. Thus, the management should be interested in creating a friendly job atmosphere that is conducive to discourses regarding IT changes among members of the organization.

References
THE EFFECTS OF ONLINE COMMENTARY ON USERS’ INFORMATION PROCESSING IN THE CONTEXT OF ONLINE DISCUSSION FORUMS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0428.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Information Processing Theory, Virtual community, Process theory, E-Marketing</td>
</tr>
</tbody>
</table>
THE EFFECTS OF ONLINE COMMENTARY ON USERS’ INFORMATION PROCESSING IN THE CONTEXT OF ONLINE DISCUSSION FORUMS

Chung, Sunghun, KAIST Business School, Heogiro 87, Supex Building, 2307, Seoul, Korea, cshhm1@business.kaist.ac.kr

Yoon, Yeosun, KAIST Business School, Heogiro 87, Supex Building, 2564, Seoul, Korea, yyoon@business.kaist.ac.kr

Han, Ingoo, KAIST Business School, Heogiro 87, Supex Building, 2306, Seoul, Korea, ighan@business.kaist.ac.kr

Abstract

Online discussion forum, which plays an important role in online criticism, provides useful information such as online commentaries generated by other users. The paper uses regulatory focus theory to explain how online commentaries are processed differently depending on the user’s information processing style and how each self-regulatory mode moderates the impact of online commentaries on one’s overall evaluation of information. The study produces three major findings: (1) Promotion-focused users are more likely to distort online information than prevention-focused users do, (2) With hedonic information, information distortion will be stronger for promotion-focused users as compared to prevention-focused users, (3) With utilitarian information, information distortion will be stronger for prevention-focused people as compared to promotion-focused users. These finding have implications for online discussion forums in terms of how to manage users effectively and also how to prevent unintended criticism.

Keywords: online discussion forum, online commentary, regulatory focus, bias information processing

1 INTRODUCTION

Online discussion forums are widely available computer-supported communication technologies that facilitate virtual interaction on the Internet. Online discussion forums are open, internet-based forums that anyone can join to discuss topics of mutual interest (Desanctis, Fayard, Roach, and Jiang, 2003). By supporting extensive online interactions to complement conventional debating, online discussion forums can effectively enrich the process of acquiring, sharing and exchanging knowledge among users, thereby improving learning performance and outcomes (Leidner, Jarvenpaa, 1995). In online discussion forums such as internet news sites or online communities, we commonly face with online commentaries generated by others. Using these commentaries, online discussion forums can be a many-to-many communication space where users can post a new topic and reply to an existing one (Cunha and Orlikowski, 2008). According to “2008 Online Journalism Award” by Online News Association, large online discussion forum sites such as Beliefnet.com, Chron.com, and NYTimes.com were selected for their effective management of online commentary. As evidenced by such examples, the importance of online commentaries is on the rise recently. Although there is a great deal of research investigating the impact of online discussion forums on organizational change (Cunha and Orlikowski, 2008), intention of participation (Yang, Li, Tan, and Teo, 2008), learning (Thomas, 2002), and corporate reputation (Clack, 2001; Park and Lee, 2007), there has been no systematic
assessment of online commentaries’ role in framing users’ information processing and experiences of bias processing.

In light of this, we propose that online commentary has a crucial impact on users’ information processing, and also that there are psychological differences in the characteristics of such process. This study will specify the role of online commentary, and it will also test this empirically. It seeks to answer following questions: (1) Is there any information distortion behaviour depending on the direction of online commentaries (positive vs. negative)? (2) What kinds of differences are there for information distortion by users’ motivational orientation? (3) Does the characteristic of information affects these biased information processing?

In what follows, we first consider some previous research on information distortion and selective exposure to information, particularly those studies examining the people’s psychological information processing. We then describe the setting of our research study and the methods we use to connect these cognitive processing to the context of online discussion forums. We next discuss the information processing style of users in terms of regulatory focus, which we identified as a moderator for the valence of online commentaries and informational characteristics. We also discuss our finding that user’s information distortion by online commentaries was induced differently depending on users’ motivational orientation and informational characteristics. We conclude the paper by examining the research implications of the impact of online commentaries to the information processing during information adoption by the user.

2 THEORETICAL BACKGROUND AND HYPOTHESES

2.1 Information distortion and selective exposure to information in online discussion forums

As noted earlier, we expected that users of online discussion forum may distort their preference toward information by adopting online commentaries. It is widely known that after a decision is made, people distort information in favor of the chosen alternative in order to reduce cognitive dissonance (Elliot and Devine, 1994; Fazio and Cooper, 1983). In online discussion forum, users may seek online commentaries that confirm their choice and depreciated encountered commentaries that oppose it. Investing these phenomena, we lift up online commentaries to the surface of theoretical context.

Information distortion is defined as a change in the forms, meaning and/or availability of information (O’Reilly, 1978; Huber, 1982). Larson and King (1996) mentioned that tendency to suppress or alter ‘bad news’, or to exaggerate and circulated ‘good news’. Also, systemic filtering of unfavourable information is referred to information distortion. More specific, three types of information distortion are identified. One type of information distortion occurs when the sender does not intend to distort information but, because of an inadvertent change in form, meaning and/or availability, information becomes inaccurate. This type of distortion is referred to as unintentional distortion. Information distortion can also be malicious in nature. Departments or individuals who do not get along may distort information to mislead or satisfy their own objectives at the expense of the other party. This type of information distortion is referred to as malicious intentional distortion. A third type of information distortion occurs when information is intentionally distorted, either for the benefit of the receiver or for the benefit of both the sender and the receiving party. This third type of distortion is referred to as well-meaning intentional distortion. Users in online discussion forums may have this third type of information distortion. They are likely to find online commentaries for benefit of them to reduce cognitive dissonance. What might cause users to distort information in online discussion forum? To answer this question, we examined the effects of valence of online commentary (positive or negative) on users’ selective exposure to information.
People often systematically prefer information that is consistent with their beliefs, attitudes, or decisions and, in contrast, neglect inconsistent information. This phenomenon is called selective exposure to information, and online discussion forum users are likely to select information which their preference matches with online commentary. A great deal of empirical research on selective exposure to information has been conducted in the context of dissonance theory (Festinger, 1957; Frey, 1986; Jonas, Schulz-Hardt, Frey, and Thelen, 2001). Dissonance theory suggests that subsequent to having made a decision, decision makers experience cognitive dissonance because of the salient negative aspects of the selected alternative and the salient positive aspects of the rejected alternative. One prominent means of reducing dissonance is selective exposure to consistent information (Festinger, 1957). Previous research has revealed that such biased information-search processes occur in different contexts, including attitudes (Lundgren and Prislin, 1998), self-serving conclusions (Frey, 1981; Holton and Pyszczynski, 1989), and online news (Best, Chmielewski, and Krueger, 2005). Most important to the focus of the online discussion forums, biased information processing has also been consistently observed following individual and group decisions (Jonas, Schulz-Hardt, Frey, and Thelen, 2001; Schulz-Hardt, Jochims, and Frey 2002). In the context of online discussion forums, we expected that selective exposure to information has also been consistently observed between positive-framed online commentaries and negative-framed online commentaries. Moreover, we proposed that the valence of online commentary (positive vs. negative) can affect users’ information preference, and these effects can be moderated by users’ motivational orientation and informational characteristics.

2.2 Regulatory focus theory and informational characteristics

For the past few decades, many researches in IS has been focused on users’ motivational factor which can explain their information processing in many contexts (Gill, 1996; Couger, Zawacki, Oppermann, 1979; Burton, Chen, and Grover, 1992). In this research, we focused on online discussion forum users’ information processing style, specifically motivational orientation factor: self-regulatory. Regulatory focus theory (Higgins, 1997) offers an insight into this motivational difference of informational processors. Regulatory focus posits two motivational orientations: a promotion focus and a prevention focus. Promotion-focused people emphasize aspirations and achievements and focus on the presence and absence of positive outcomes. Prevention-focused people are concerned with responsibilities and safety and focus on the presence and absence of negative outcomes (Shah, Higgins, and Friedman, 1998). In accord, Idson, Liberman, and Higgins (2000) found that promotion-focused people exhibited greater “eagerness” in working toward a gain than guarding against a nongain, while prevention-focused people displayed greater “vigilance” in preventing a loss than working toward a nonloss. In this regard, in online discussion forums, promotion-focused users display greater eagerness in striving toward a positive outcome (positive online commentaries) than away from its absence (negative online commentaries). Further, consistent with these users’ desire to secure hits and avoid misses, positive online commentaries in online discussion forums allows for an inclusion of consistent original information. In contrast, prevention-focused users are concerned about negative outcome (negative online commentaries) more intensely than they are concerned with its absence. They focus on identifying correct rejections and on avoiding incorrect “hits.” Considered in this framework, we address that these motivational differences can induce different bias information processing and selective exposure to information in online discussion forums.

As evidenced by their endorsement of multiple alternatives and placing lower weight on prior choices in making future decisions, promotion-focused users should display a lower tendency to allow their evaluation of one information to bias the other information’s evaluation. In contrast, based on their tendency to entertain fewer hypotheses and to place greater weight on prior choices when making future decisions, prevention-focused users should allow their evaluation of negative information to influence their evaluation of the other. Thus, we expect that the tendency of prevention-focused users to allow their rating of online commentaries to affect the overall information’s rating is due to the possibility that they have a greater preference for consistency than do promotion-focused users. Preference for consistency (PFC), the “desire to be and to be seen as consistent” (Guadagno, Asher,
Demaine, and Cialdini, 2001), leads users with a higher PFC to weight previous expectations or choices more when engaged in subsequent decisions. Further, low-PFC users seen open and oriented to the new, in ways that are relatively unconstrained by the established (Cialdini, Trost, and Newsom, 1995). Related to regulatory focus, Higgins (1996) suggests that prevention-focused users' tendency to narrow alternatives “raise the intriguing possibility that self-consistency motivation is linked to the prevention focus of self-regulation.” So, we expect differences in PFC associated with different regulatory foci to manifest different level of information distortion in online discussion forums by online commentaries.

**H1:** Promotion-focused users are more likely to distort online information than prevention-focused users.

In line with consumer behaviour literature that distinguishes between utilitarian and hedonic products (Hirschman and Holbrook, 1982) and applying literature to IS context (Heijden, 2004), we can classify these types of information in online discussion forums. There are two kinds of information in online discussion forums: utilitarian information and hedonic information. Consistent with previous research in informational characteristic, we use the term “utilitarian information” to refer to the functional, instrumental, practical benefits of informational characteristic, and we use the term “hedonic information” to refer to its aesthetic, experiential, and enjoyment-related information (Batra and Ahtola, 1990; Chitturi, Raghunathan, and Mahajan, 2007; Dhar and Wertenbroch, 2000). In the context of online discussion forums, for example, the economic theme and health information can be utilitarian information, whereas aesthetic appeal from entertainment news. The consumer behaviour literature demonstrates that what specifically determines intention to consume depends on the utilitarian or hedonic nature of the product (Babin, Darden, and Griffin, 1994; Holt, 1995). We develop a parallel argument that what shapes information distortion is dependent on the utilitarian or hedonic nature of information in online discussion forums.

There is converging evidence that the types of goals people to be fulfilled by the utilitarian dimension of a product are different form those they seek from the hedonic dimension (Chernev, 2004). Specifically, whereas people expect the fulfilment of prevention goals on the utilitarian dimension, they expect the fulfilment of promotion goals on the hedonic dimension (Chitturi, Raghunathan, and Mahajan, 2007). According to the regulatory focus theory, prevention goals are those that ought to be met, such as “behaving in a safe and secure manner” and “being responsible.” Fulfillment of prevention goals in the context of online discussion forums eliminates or significantly reduces the probability of painful information such as negative online commentaries, thus making overall evaluation of information that results form fulfilment of prevention goals. For example, in the case of economic news information, prevention goals might be served by its useful features or helpful online commentaries. Conversely, promotion goals are those that a person aspires to meet, such as “fun or enjoyable information.” Fulfillment of promotion goals in the hedonic information such as entertainment news significantly increases the probability of a pleasurable experience, thus enabling users to distort information that result from the adoption of online commentaries. Therefore, following hypotheses are proposed that the different level of information distortion depending on informational characteristic (hedonic vs. utilitarian) and users’ regulatory focus (prevention vs. promotion). Also, the proposed conceptual framework in Figure 1 captures the relationship among hedonic versus utilitarian information, the impact of online commentaries, and users’ regulatory focus, respectively.

**H2:** In case of hedonic information, information distortion will be stronger for promotion-focused users as compared to prevention-focused users.

**H3:** In case of utilitarian information, information distortion will be stronger for prevention-focused people as compared to promotion-focused users.
3 RESEARCH DESIGN AND METHOD

3.1 Design, Subjects, Experimental System, and Procedure

The present study uses a $2 \times 2 \times 2$ between-subjects experiment manipulating regulatory focus (promotion vs. prevention), valence of online commentaries (positive vs. negative), and information characteristic (utilitarian vs. hedonic). Two hundred seven college students participated in the study voluntarily. Their average age is 21.8 and 54.2% is male. Most of the subjects had participated in online discussion forum such as beliefnet.com or NY Times.com. Subjects were randomly assigned to each of the cells in the factorial design, and they participated in different experimental conditions in a single session. The original information used in the experiment were two news manipulating information characteristic. Consistent with previous researches, utilitarian information contains useful and helpful reports about economic phenomenon. In contrast, hedonic information contains interesting and exciting reports about entertainment gossip. For the experiment, eight-independent virtual online discussion forum sites were constructed for each condition. The online discussion forum sites contained both original news information and the set of online commentaries. This experimental system encapsulates the attributes of a real online discussion forums sites. In an experiment, the subjects followed instructions that address a scenario manipulating different regulatory foci. Then original news information was provided to each group, and measured their evaluation about news. After measuring prior evaluation, subjects received online commentaries together with original news information, and measure their evaluation again. Subsequently, the subjects gave an answer questions about preference for consistency, manipulation checks, in regular sequence.

3.2 Independent Variables

3.2.1 Information characteristics (Utilitarian vs. Hedonic)

Based on the news information from a real online discussion forum, six candidates for news information were collected. To increase face validity, twelve subjects participated in a pre-test that checks whether this information was perceived in the appropriated dichotomy (utilitarian vs. hedonic). In the main experiment, relatively different news information was used. Utilitarian information was manipulated with instrumental and functional topic such as economic news, whereas hedonic information provide fun, pleasure, and excite topic such as entertainment gossip. This relative hedonic or utilitarian nature of information is consistent with prior research (Batra and Ahtola, 1990; Dhar and Wertenbroch, 2000). Each online news contained an attribute of information such as instrumental or
pleasure topics, and the length of information was set at fifteen lines. The length of each news information was controlled to avoid the effects of information quantity, and to focus on the effects of information characteristic. Also, original news information provided neutral opinion to focus on the effects of the valence of online commentaries.

3.2.2 The valence of online commentaries (Positive vs. Negative)

Before manipulating the valence of online commentaries, forty candidates for online commentaries were collected from a real online discussion forum. Each online commentaries’ length was 2 lines and consist of 40 words. A focus group interview organized with ten people who frequently use online discussion forum was performed. In the FGI, subjects were asked the degree of strength that how much online commentaries agree to original news or not. After the FGI, we calculated strength of online commentaries with nine-point scale (where -4 = extremely opposite, 0 = neutral, and 4 = extremely agree), and among same strength-commentaries, finally five positive online commentaries and five negative online commentaries were selected. The positive valence of online commentaries was organized with seven positive commentaries and three negative commentaries, and negative valence of online commentaries was organized vice versa. Figure 2 shows example of experimental online discussion forum screen.

3.2.3 Regulatory Focus (Prevention vs. Promotion)

To manipulate regulatory focus, we provide a role-play for subjects in the introductory session. This manipulation is very much like that of situational regulatory focus in past studies (Jain, Lindsey, Agrawal, and Maheswaran, 2007; Shah, Higgins, and Friedman, 1998). Regulatory focus was manipulated first by requiring subjects to unscramble six jumbled words that were actually names of commonly used cosmetic brands. Prevention condition subjects were informed: “Each incorrectly unscrambled name loses you 2 point. If you do not get a name wrong, you won’t lose 2 points. Your goal is to lose as few points as possible by minimizing the number of names you get wrong. For every brand name that you wrong, you will lose 2 point. For every brand name that you don’t get wrong, you won’t lose 2 point.” In contrast, promotion condition subjects were informed: “Each correctly unscrambled name gains you 2 points. If you do not get a name correct, you will not gain 2 points. Your goal is to gain as many points as possible by minimizing the number of names you get right. For every brand name that you get right, you will win 2 points. For every brand name that you do not get right, you won’t win 2 points.” Due to the dichotomized regulatory focus, prevention-focused subjects processed the information both original news and online commentaries more carefully via their concern about the presence of negative online commentaries, but promotion-focused subjects processed the information both original news and online commentaries more carefully via their concern for positive online commentaries. This is consistent with regulatory focus theory.

3.3 Control Variables and Dependent Variable

Control variables: Experimental systems should consider other effects such as the characteristics of subjects and exogenous stimuli on all matters. It could be affected by the tendency of reliance on online commentaries, general attitude toward the online discussion forum. To focus our independent variables and increase internal validity, the following methods were employed to control confounding effects on evaluation of information. First, the characteristics of subjects comprising individual differences such as online searching styles and personality were controlled by allotting subjects to each condition at random. Second, this present experiment used imaginary online discussion forums’ names. If we consider the online discussion forums’ reputation, many compounding effects are revealed. An online discussion forum which has a strong reputation may tempt users, and they only regard the reputational cue as serious. In this case, since the online discussion forums’ role (i.e. informational role) cannot work, the work retailer is probably unsuitable. Finally, prior knowledge
about original news information, the degree of reliance on online commentaries was measured by survey items in experiments, and they were used as control variables in an ANOVA analysis.

**Dependent variable:** Preference change toward information used in dependent variable (Russo, Medvec, and Meloy, 1996). It indicates the level of information distortion, and was measured by the difference between before and after the evaluation of information. First, participants indicated their evaluation of online news on three scales anchored ‘favorable/unfavorable,’ ‘good/bad,’ and ‘desirable/undesirable.’ And then, participants were asked their evaluation on both online news and according online commentaries. Preference change toward information was developed by subtracting next evaluation to prior evaluation.

### Oil and Gasoline

In 1999, the price of oil hovered around $18 a barrel. In July 2008, it reached a peak of $147 a barrel. In the months that followed, as fears of a global recession grew, prices plunged to the $75 a barrel range, a roller coaster ride that left both producers and consumers confused and wrung out. Prices were still far higher than they had been a few years earlier, but oil-producing countries that had reshaped their economies around the huge influx of revenues faced a suddenly altered landscape. Many factors contributed to the long buildup between 1999 and 2008, including the relentless growth of the economies of China and India and widespread instability in oil-producing regions, including Iraq and Nigeria’s delta region. The triple-digit oil prices that followed appeared to redraw the economic and political map of the world, challenging some old notions of power. Oil-rich nations made enjoying historic gains and opportunities, while major importers—including China and India, home to a third of the world’s population—confronted rising economic and social costs.

![Image of online discussion forum site](image)

**Figure 2.** The experimental online discussion forum site (example of utilitarian information and the negative online commentaries condition)
4 RESEARCH RESULTS

4.1 Manipulation Checks

To check the manipulation of informational characteristic, we used the perceived measure adopted from prior research (Batra and Ahtola, 1990). Informational characteristic was measured using four seven-point semantic differential scales (the utilitarian component by the items useful/useless, and beneficial/harmful; the hedonic component by the items pleasant/unpleasant, and nice/awful). The ANOVA analysis indicated that there are significant differences between the utilitarian information condition and the hedonic information condition [for utilitarian component measure, M=4.68 vs. 2.41, F(1,206)=167.295, and for hedonic component measure, M=5.31 vs. 1.91, F(1,206)=370.69]. The subjects also checked two items designed to check neutrality of original news information (“This online news is neutral,” and “This online news contained both support arguments and counter argument.”). The ANOVA results show that the neutrality of original news information was also successfully manipulated [M(utilitarian information)=3.80 vs. M(hedonic information)=3.24, p>0.1]. Then, as a check on the valence of online commentaries, we measured two items: “the online commentaries about this online news are mainly supportive/opposite (reversed coding).” Using this method, the valence of online commentaries was manipulated successfully [for utilitarian news information, M(positive commentaries)=4.68 vs. M(negative commentaries)=2.41, F(1,206)=167.295; for hedonic news information, M(positive commentaries)=5.31 vs. M(negative commentaries)=1.91, F(1,206)=370.69]. Finally, the efficacy of the regulatory focus manipulation was assessed by asking subjects to indicate the “extent to which you focused on scoring more points when playing the brand name quiz” and “the extent to which you focused on not losing any points when playing the brand name quiz.” The ANOVA tests confirmed that promotion-focused (prevention-focused) subjects concentrated more on winning (not losing) points [M(promotion/score more)=5.09, M(promotion/not loss)=4.50, F(1,206)=6.96; M(prevention/score more)=3.39, M(prevention/not loss)=4.93, F(1,206)=53.05]. Thus, the manipulations were successful.

4.2 Hypothesis Testing

The preference change toward information (Cronbach’s Alpha = 0.84) was calculated as a difference between after and before evaluation for information (where -3 indicated “strongly disagree,” 0 indicated “neutral,” and 3 indicated “strongly agree”). Table 3 presents the mean and standard deviation of dependent measures for each cell. The dependent measures were analyzed in a series of 2 (Informational characteristic) × 2 (The valence of online commentaries) × 2 (Regulatory focus) ANCOVA. The ANCOVA results are in Table 4. The covariate variables were not significant. So, exogenous factors were controlled successfully.

<table>
<thead>
<tr>
<th>Informational Characteristic</th>
<th>Hedonic Information</th>
<th>Utilitarian Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulatory Focus</td>
<td>Promotion</td>
<td>Prevention</td>
</tr>
<tr>
<td>The Valence of Online Commentaries</td>
<td>Negative</td>
<td>-0.23 (0.71)</td>
</tr>
<tr>
<td></td>
<td>n=52</td>
<td>n=52</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>0.31 (0.65)</td>
</tr>
<tr>
<td></td>
<td>n=50</td>
<td>n=53</td>
</tr>
</tbody>
</table>

Table 3. Descriptive statistics of Preference change toward Information

<table>
<thead>
<tr>
<th>Effect</th>
<th>F-value</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior Knowledge about original news information</td>
<td>1.157</td>
<td>0.341</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
The analysis indicated the presence of a significant main effect of users’ regulatory focus \[F(1,206)=5.323, p<0.05\]. It means that promotion-focused users more distort information than prevention-focused users. As suggested earlier in this paper, the preference for consistency (PFC) of users can affect their bias information processing. To validate and understand the underlying mechanism of plausible explanations, simple regressions are conducted both promotion-focused users and prevention-focused users. Consistent with the expected regulatory focus-PFC relationship, analysis revealed that, promotion-focused subjects are likely to distort information with no effect of PFC on preference change \[t(1,101)=0.515, \beta =0.051, p>0.1\]. In contrast, prevention-focused led subjects to less distort information with marginal significant effect of PFC on preference change \[t(1,104)=1.673, \beta =0.163, p<0.1\]. Therefore, Hypothesis 1 was accepted.

Also, there is a significant main effect of the valence of online commentaries \[F(1,206)=5.711, p<0.05\]. It means that users in online discussion forum are affected their attitude toward information by online commentaries generated by other users, and form their preference based on these online commentaries. The set of positive-direction online commentaries lift up users’ evaluation on information, but the set of negative-direction online commentaries reduce users’ evaluation on information irrespective of informational characteristics.

![Figure 5. Three-way Interaction for preference change toward information](image)

We further explored the different effects of the manipulation of users’ regulatory focus and the valence of online commentaries under hedonic and utilitarian information conditions. Theses analyses were possible because the three-way interaction effect of informational characteristic \( \times \) regulatory focus \( \times \) the valence of online commentaries was significant \[F(1,206)=4.178, p<0.05\]. For hedonic information, promotion-focused subjects showed greater tendency that change their preference by
positive and negative online commentaries \[F(1,101)=7.652, p<0.01\]. But, for prevention-focused subjects, there are no significant preference change depending on the valence of online commentaries \[F(1,104)=0.290, p>0.1\]. These results support our predictions that information distortion can be stronger for promotion-focused users than for prevention-focused users in case of hedonic information (see Figure 5 on the left). Accordingly, Hypothesis 2 was accepted.

Under utilitarian information, in contrast, prevention-focused subjects showed greater change of preference by the direction of online commentaries \[F(1,104)=7.24, p<0.01\]. For promotion-focused subjects, there are also marginal preference change \[F(1,101)=2.58, p<0.1\]. But, their information distortion is less strong compared with that of prevention-focused subjects (see Figure 5 on the right). Such results might be interpreted as suggesting the difference pattern of information distortion in utilitarian information depending on users’ self-regulatory. Specifically, information can be stronger for prevention-focused users than promotion-focused users in case of utilitarian information. Thus, Hypothesis 3 was accepted.

5 CONCLUSION AND LIMITATIONS

The online discussion forum has increasingly been recognized as a useful space that facilitates virtual interaction and forms online criticism. We focused on the question of users’ information preferences change due to specific direction of online commentaries in order to investigate the influence of users’ motivational orientation and the informational characteristics in online discussion forums. The empirical results indicated that the informational characteristic and the valence of online commentaries were important because they change the preference to the given information, and these patterns were different depending on users’ regulatory focus. The primary insights that this research provides are as follows: (1) Overall, promotion-focused users distort their preference more than prevention-focused users do when adopting online commentaries, (2) Information that meets or exceeds users’ hedonic needs and fulfills promotion goals enhance users’ information distortion, and (3) Information that meets or exceeds users’ utilitarian needs and fulfills prevention goals enhance users’ information distortion in online discussion forums. Furthermore, the research finds that the online commentaries have significant impact on users’ preference change to information. For the two decomposed concepts of the regulatory focus, promotion-focused users react more to the valence of online commentaries in case of providing hedonic information. In contrast, prevention-focused users are more sensitive to the valence of online commentaries when utilitarian information is presented. We establish a causal asymmetric link between framing and focus that connects information distortion of utilitarian information and hedonic information. Furthermore, we identify the process underlying this different information distortion. The higher PFC associated with prevention-focus induces little preference change to information. But, the lower PFC associated with promotion-focus allows for more preference change to information. PFC may also play a role in information distortion phenomenon occurring in online discussion forum such as the prevention-focused users prefer to the status quo and the notion that promotion-focused users entertain multiple hypotheses.

For theoretical contribution, building on the work of Yang, Li, Tan, and Teo (2007), Cunha and Orlikowski (2008), we proposed that sophisticated information processing in online discussion forums. Our examination of the information distortion in online discussion forum by online commentary has implications for the online communicational and educational literature in that it provides evidence for users’ biased information processing. To the best of our knowledge, this is the first study to demonstrate such a correspondence among informational characteristics (hedonic vs. utilitarian), the valence of online commentaries (positive vs. negative), and users’ regulatory focus (prevention vs. promotion). The research findings would have significant implications for managers in online discussion forums. Practically, this study’s major results emphasize the importance of proper management of online commentaries. Because of negative aspects of unintended online criticism, managers and users are often compelled to process among various attributes of information. Since online commentaries are differently processed depending on users’ motivational orientation, managers
can use this result strategically as a new channel that provides a bridge between users and online information providers.

The present study has several limitations. First, the reputation of the forum itself was not considered. Since users did not have a holistic valuation of the reputation of the online discussion forums, the effect of online discussions’ reputation could be explained by additional study. Second, in this study, the quantity of online commentaries was fixed to avoid unintended effects. It could be possible that the quantity of online commentaries act as a signal for the popularity of the information. Further research considering this issue should be conducted. Finally, regulatory focus was manipulated by differing situational role-playing. Individual’s regulatory focus can be either situational or chronic. From this point of view, users’ chronic regulatory focus needs to be considered in the future study. In spite of these limitations, the result described in this study shows that online news and related online commentaries need to monitor closely the extremely negative or defamatory postings in their online discussion forums in order to minimize possible spillover of negative sentiments from online discussion forum users to others, and also provides a rationale for online discussion forums to offer venue for online public discussions.

References


THE APPLICATION OF A PHENOMENOLOGICAL FRAMEWORK TO ASSESS USER EXPERIENCE WITH MUSEUM TECHNOLOGIES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0601.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Phenomenology, Museums, Evaluation, Tourism</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
THE APPLICATION OF A PHENOMENOLOGICAL FRAMEWORK TO ASSESS USER EXPERIENCE WITH MUSEUM TECHNOLOGIES

Pallud, Jessie, CESAG Research Center, EM Strasbourg Business School, 61 Avenue de la Forêt Noire, 67085 Strasbourg Cedex, France, jessie.pallud@em-strasbourg.eu

Abstract

Providing visitors with a valuable experience of the past has become a crucial mission for cultural heritage institutions. The experience of the past is one where visitors understand the museum’s communications about the meaning of artefacts and where visitors undertake an active role in interpretation and reflection on the past. Several studies promote technologies as a good way for museums to reenergize their relationships with their visitors. But even as some research has concentrated more and more on visitor experiences, this work has neither particularly stressed on visitors’ experience of the past nor on their evaluation of museum technologies with respect to their potential for engendering a better experience of the past.

Monod and Klein (2005) elaborated a phenomenological framework with six criteria to evaluate IT used in the cultural heritage. Since it has not been empirically “validated” yet, the objective of this paper is to employ these criteria with samples of users of museum technologies and in the process determine whether these criteria can be met by IT.

Our field study indicates that technologies available in museums positively contribute to an experience of the past.

Keywords: Museum technologies, evaluation, experience, phenomenology, Heidegger, historicity.
1 INTRODUCTION

Museum studies in a wide range of journals have noticed individuals having an increased interest in the past in general (Liew 2005), and more specifically in heritage sites (Poria et al. 2006). The principal motivations for heritage site visits are the desire to connect with one’s heritage and the desire to learn more about history (Poria et al. 2006). According to Kimmelman (2001), “we go to museums to remind ourselves [about] who we are”. There are many benefits that visitors gain from heritage sites, from connections with the past and identity finding (Chronis 2005) to a “nostalgic bonding” linked to the recollection of things belonging to the past (Holbrook and Schindler 1994). The past is a valuable experience because of its civilizing effects and the extent to which it becomes a highly personal experience for the participant. It is one where visitors understand the museum’s communication about the meaning of artifacts and where they undertake an active role in interpretation and reflection on the past.

Consequently, providing visitors with a valuable experience of the past has become a crucial mission for cultural heritage institutions. Technologies have been presented as the panacea to enhancing visitor experience. Indeed, numerous studies promote technologies as a reasonable way for museums to reenergize their relationships with their visitors (Fopp 1997; Messham-Muir 2005; Vom Lehn and Heath 2005). For instance, virtual reality technologies can represent “items which have not survived, creatures which are extinct, the visions of great men which were never realized or even the imagination of artists and thinkers” (Fopp 1997, p. 146). By displaying events from the past, virtual reality helps people relive historical events in their mind. Furthermore, the “first-person perspective” (Lok 2004, p.50) enables visitors to feel more concerned by what they see and so to project themselves in history. Hybrid environments or mixed reality simulators, devices that enable visitors to look at real and virtual objects in the same time (Lok 2004; Sparacino 2004), reduce the distance between past and present objects by allowing visitors to directly compare objects belonging to two different eras. More common museum technologies like audioguides, interactive and multimedia kiosks generally provide historical background and tutorial aids to visitors.

However, even while some prior research has focused on visitor experience, there has been neither a particular stress on visitor experience of the past nor on an evaluation of the technologies that are proposed to enhance better experience of the past. In effect, museum professionals have few guidelines on how to assess visitor experience, particularly in the context of IT use (Institute of Museum and Library Services 2006; Pujol Tost and Economou 2007).

Monod and Klein (2005) have proposed a phenomenological framework to evaluate IT used in the cultural heritage. Nonetheless, this framework has not been empirically tested and the criteria have not been operationalized with subjects. Therefore, this research aims at examining these phenomenological concepts. The subsequent research questions that guide this study are the following:

1. To what extent do museum technologies contribute to an experience of the past? More precisely, to what extent does IT meet the criteria proposed by Monod and Klein (2005) and convey a phenomenological experience to visitors?

2 THE EVALUATION OF VISITOR EXPERIENCE

Researchers from different disciplines have developed a wide variety of frameworks to study visitor needs and to provide directions to museums (Anderson 2004; Kotler and Kotler 2000). It is noteworthy that these frameworks have remained conceptual since they have not been empirically tested within museums or with visitors. Furthermore, they do not explicitly address the role of technologies in museums. For instance, the marketing researchers Kotler and Kotler (2000), propose
three dimensions on which museum professionals should focus in order to improve visitor experience. They are: (1) the variety of visitor experiences, (2) the level and depth of visitor experiences and (3) the design and orchestration of visitor experiences. While Kotler and Kotler (2000) principally identify operational levers to enhance visitor experience, Anderson’s framework (2004) points out the organizational functions of museums serving as a foundation to develop a visitor perspective. More precisely, Anderson (2004) suggests four domains on which museums should rely in order to be visitor oriented: 1) governance, 2) institutional priorities, 3) management strategies and 4) communication style.

Certainly, a number of empirical frameworks do address visitor experience and some examine visitor experience with technologies (Falk and Dierking 1992; Peacock and Brownbill 2007). But even while these studies do focus on visitor experience, there has been neither a particular stress on visitor experience of the past nor on an evaluation of the technologies that are thought to lead to a better experience of the past for visitors. Additionally, museum professionals have few guidelines on how to assess visitor experience, particularly in the context of IT use (Institute of Museum and Library Services 2006; Pujol Tost and Economou 2007). According to a recent report,

“[Museums] would like information, training, and guidance on how to better assess user needs, including methods of collecting information about the characteristics of users, how they use an institution’s technology and digitization services and products, and for what purpose.” (IMLS, 2006, p.124)

Information systems evaluation is an important research stream for the IS discipline and it has given rise to several frameworks, such as the IS Success Model (Delone and McLean 1992; Delone and McLean 2003) and the Task-Technology Fit (Goodhue and Thompson 1995). This issue has also been addressed in several IS journals and in special issues (e.g., Irani and Fitzgerald 2002). For instance, Delone and McLean (1992) elaborated the IS Success Model in order to determine the factors that play a role in IS success. This model was refined in 2003 to include other variables and new linkages that appear to be important in the assessment of IS. Recently, Petter et al. (2008) in their meta-analytical review of the IS Success Model highlighted the fact that this research model has mainly been applied to utilitarian contexts.

“[M]what still remains to be discovered is if the D&M model is appropriate for hedonic IS. Some of the dimensions may no longer be relevant or may need to be measured differently for gaming, social networking, or other types of IS used for enjoyment.” (Petter et al. 2008, p. 258)

Consequently, prior research dealing with evaluation has mainly focused on a utilitarian perspective of IS evaluation, that is trying to determinate how IS can increase organizational or individual performance. Appropriate for this utilitarian context, users are often represented by managers or employees.

What is different about the cultural heritage context, however, is that IT users are best compared to visitors but visitor performance does not appear to be as important as in company settings. Rather, enhancing user experience by providing entertainment, education, and a good experience of the past are the crucial targets for museums. According to the well-known museologist Sola (1997), “a museum should assist people to understand the experience of the past. In its mutual relationship with its users, it should find in past experience the wisdom necessary for the present and the future” (Boylan 2002). Moreover, museum technologies tend to serve hedonic purposes. Therefore, the evaluation of IS deployed in cultural heritage institutions should be done in accordance with these cultural objectives.

Some researchers who have paid attention to hedonic information systems (Van der Heijden 2004; Wakefield and Whitten 2006) show that they require specific criteria for their evaluation. For instance, Van der Heijden (2004) highlight the fact that enjoyment and ease of use are more relevant in the
assessment of hedonic technologies. The Human-Computer Interaction subfield has also devoted part of its research attention to hedonic systems. User reactions, such as cognitive and affective reactions, have been conjointly analyzed to better evaluate the efficiency of information systems (Sun and Zhang 2006).

Even if cultural heritage research focuses more on visitor experience, there is neither a particular stress on visitors’ experience of the past nor on evaluations of the technologies with respect to their potential for engendering this better experience of the past. Furthermore, Monod and Klein (2005) explain that the frameworks that exist to evaluate IT in cultural heritage are mainly driven by technological determinism. Indeed, these models generally postulate that the implementation of IS in museums will positively impact visitor satisfaction and experience even while there is little verification of whether these technologies really achieve their goal (Monod and Klein 2005). In the next section we present a framework that addresses many of these gaps by evaluating user experience with IT from a phenomenological point of view.

3 THE VALUE OF PHENOMENOLOGY IN STUDYING VISITOR EXPERIENCE: INTRODUCTION OF THE CONCEPTUAL FRAMEWORK

3.1 Phenomenology: How this Philosophy Particularly Fits Museums

Phenomenology is the “science of phenomena” (Heidegger 1962, p.50). This is also a philosophical movement that appeared in the first half of the 20th century (Spiegelberg 1975). It focuses on the experiences of individuals. Indeed, it aims at studying “phenomena as consciously experienced” (Spiegelberg 1975, p. 3). This philosophical underpinning was initiated by Husserl (1936) and his student Heidegger (1962). They encouraged researchers and philosophers to turn “to the things themselves”. People should turn themselves “to the world as it is already experienced” (Ilharco 2002, p. 304). Other philosophers like Merleau-Ponty and Sartre also nurtured phenomenology through the concepts of self and embodiment (Smith 2003). While these other philosophers are important in the development of the philosophy of phenomenology, our research follows Heidegger’s view as it was developed in his book Being and Time (1962).

Phenomenology aims at studying individual experiences. Heidegger (1962) contended that human beings need action and praxis with objects (i.e., to engage with them) in order to feel closer to these things (Smith 2003). Therefore, individuals cannot see an object or imagine it in order to understand it because it is only a “representational form of intentionality” (Smith 2003). This argument leads to conclusion that being able to touch things or to manipulate them contributes to a better experience and to better interpretation.

As indicated in the title of his book, Heidegger (1962) addresses the question of time and its relation to being. Heidegger asserted that time has an ontological function since it constitutes being (Dastur 1993). Indeed, “we are temporal beings not because we exist in time but because time is really what composes our beings” (Dastur 1993, p. 301). Temporal beings are open because individuals are always turned towards the future and the past, and their self-meaning is not fixed (Lyotard 1992).

Additionally, it seems that history plays a role in people’s existence as it can shape their present and future (Monod and Klein 2005). In effect, historical objects represent remains of the past and consequently, they give to people their historical dimension. It is thanks to these remains that individuals know that something before them existed (Heidegger 1962). However, these historical objects have a secondary historicity; they are historical because they belonged to a past humanity and were created by historical beings (Lyotard 1992).
In that Heidegger’s phenomenology puts a special emphasis on time (history), human existence and experience seem to be perfectly appropriate to the study of cultural heritage institutions, whose goals are to display past heritage and focus on visitor experiences.

3.2  Phenomenological Framework: Presentation of the Criteria

Based on Heidegger’s concept of historicity, Monod and Klein (2005) elaborated a framework to evaluate technologies of cultural heritage institutions. The framework aims to determine whether technologies, by meeting user requirements, have interpretive characteristics, and whether IT contributes to a good experience of past. In their original framework, the authors included eight criteria: re-enactment, embodiment, context, self-projection, possibilities of being, historical self, inquiring being, and universality in uniqueness. These criteria were reduced to six in a more recent version, and the framework now focuses on context, embodiment, self-projection, re-enactment, possibilities of being, and historicity. These criteria are defined in greater detail hereafter.

- **Context**

  The first criterion proposed by Monod and Klein (2005) to provide IT users with a phenomenological experience is context. According to Monod and Klein (2005), context is represented by the shared values, overarching values and beliefs that contribute to meaning-making experiences. Indeed, without cultural and historical context it is difficult for individuals to have a comprehensive understanding of their personal history and of history in general. It is very frequently true that in cultural heritage sites, visitors do not understand the purpose of an object or even realize its historical importance. Indeed, Schärer (1996) contends that information provided within museums is generally more structural (some general indications) than cultural (information on the earlier context of use).

- **Re-enactment**

  Re-enactment is the second criterion identified by Monod and Klein (2005). “Re-enact” in a literal sense means to “perform again” or “to go through a second time” (The American Heritage Dictionary of the English Language 2000). Collingwood (1946), who studied philosophy of history and devoted a lot of his research to re-enactment issues, argues that the work of historians should be seen as an imaginative reconstruction. Indeed, re-enactment was first set forth as an important capability for historians since these scholars need to relive historical events in their mind in order to interpret history and to better convey it to people (Collingwood 1946). Consequently, re-enactment can be viewed as a methodology to produce historical knowledge (Nielsen 1981). It is noteworthy that this possibility of reenactment should also be at visitors’ disposal in order to enhance their experience of the past (Monod and Klein 2005). Indeed, if visitors can relive historical events in their mind, they will be projected into the past and are more likely to understand historical personalities, for instance, or way of life in the past.

- **Embodiment**

  The third dimension deemed to be important for IT user experience is embodiment. Embodiment is a notion that was developed principally by Merleau-Ponty (1962). The Cambridge Dictionary of Philosophy (1999) gives this definition of embodiment. It is “the bodily aspects of human subjectivity. Embodiment is not a concept that pertains to the body grasped as a physiological entity. Rather it pertains to the phenomenal body and to the role it plays in our object-directed experiences.”

  Mingers (2001) also examined the concept of embodiment and its implications for IS research. According to Mingers (2001), embodiment resides in the fact that “our basic attitude is always (except in pure contemplation) one of doing, acting, having some aim in mind, having some concern” (p. 108). His explanation sheds more light on the definition of The Cambridge Dictionary of Philosophy. From this observation we can conclude that embodiment designates the sensory experiences that an individual may have with objects encountered in the world.

- **Self-projection**
Monod and Klein (2005) proposed self-projection as a fourth criterion. Self-projection works by allowing one to put oneself mentally in the shoes of historical characters and by imagining what one could and would have done in another’s situation. This type of self-projection has both cognitive and emotional aspects. The cognitive aspects are linked to the deliberations that lead to decisions and actions actually taken whereas the affective aspects are related to emotions such as love, anger, surprise, joy, etc.

- **Possibilities of being**

Possibility of being is the fifth phenomenological criterion. According to Monod and Klein (2005), a phenomenological experience helps people realizing the constraints that have been created by the past and the impacts on their present life. This leads to the realization that the present could have been different, too, had the past been different. Reflecting on alternative pasts, individuals come to realize how the present could have been different, too. Monod and Klein (2005) argue that cultural heritage sites, and more precisely historical characters, represent an important vehicle for inspiring this process.

- **Historicity**

The last criterion, historicity, is the most conceptual and existential dimension. Historicity refers to the understanding that we are fundamentally historical beings and that the meaning of our action and of our existence, is linked to history. In Being and Time, Heidegger (1962) devotes the entire section “The Vulgar Understanding of History and the Occurrence of Da-Sein” to define history properly. Heidegger insists that the term “history” should be distinguished from the term “past” in that history represents the influence and consequences of past on the present and the future: “Thus history does not so much mean the ‘past’ in the sense of what is past, but the derivation from it” (Heidegger 1962, p. 347). Heidegger also uses words such as “move, rise, fall, connection, change and transformation” to designate history (p. 347). These words capture the dynamic nature of history. It is noted that history is a specific component of human beings, or “Dasein,” since it constitutes our lives. Therefore, to capture history in a meaningful way, we argue that individuals need to be confronted with a dynamic representation of history by understanding the influence of past events on their personal existence.

Even though not grounded in a phenomenological framework, other studies validate the importance of these six phenomenological criteria for IT users’ experience. For instance, Pujol Tost and Economou (2007) surveyed visitors of the Ename Museum (Belgium) about their favorite rooms and devices at the end of their museum visit. The applications that were designated by visitors correspond to the ones that were able to convey context, empathy, interactivity, and sensations. More precisely, Pujol Tost and Economou (2007) found that context is one of visitors’ most important expectations. Moreover, empathy contributes to visitor engagement and satisfaction. The dimensions of interactivity and sensations (described by the participants as the possibility to touch) also led to better learning.

Hence, the phenomenological criteria developed and identified by Monod and Klein (2005) appear to be of great importance to visitor experience. As Monod and Klein (2005) have not verified their framework in the field, the present research will extend their work empirically through a field study with museum visitors.

### 4 METHODOLOGY

We distributed questionnaires to visitors of a French museum, hence employing a field study methodology. This field study aimed at empirically examining the extent to which museum visitors believe the phenomenological concepts are being facilitated by the museum technologies. Even while we are utilizing a quantitative methodology, this study does not ground itself in positivism. We do not test a research model with a set of hypothesis, but rather view the questionnaires as an opportunity to obtain more qualitative data. Moreover, Mingers (2003) asserts that “the tendency to link quantitative
methods with a natural science (positivist) approach, and qualitative methods with a social science (interpretive) approach” corresponds to a “crude dichotomy” (p. 236). The questionnaire methodology was also more convenient to implement at our field study because visitors did not have much time to spend in interviewing.

4.1 Presentation of the Site

The research site was the National City of History of Immigration (NCHI), a public museum located in Paris, France. This museum, inaugurated in October 2007, deals with the history of immigration, tracing back its evolution in France from the nineteenth century to the present. NCHI exhibitions also show how immigration has contributed in shaping French society. In research conducted before the opening of NCHI, potential visitors expressed strong expectations about historicity and an experience of the past. As an illustration, Poli et al. (2007) showed that potential visitors expected that NCHI would: 1) deepen their knowledge about history of immigration, 2) offer new perspectives on their personal history, 3) contribute to identity building and 4) explain the consequences of immigration on the world of today.

In addition to dealing with history, this museum uses modern media to communicate with its public. NCHI is equipped with common museum tools such as televisions, RFID audioguides, but also more recent technologies such as computers, Webcam, and interactive kiosks with both audio and video content. Overall, NCHI appeared to be an adequate setting for examining the historical and phenomenological concepts, as well as to evaluate the contribution of IT to an experience of the past.

Furthermore, the choice of a history museum can be justified by the following reasons. First, we did not want to study types of museums that were highly specialized to the point of idiosyncrasy, as, for example, science museums which generally include a lot of technological equipment and hands-on activities. In such a case, we believed that perceptions of authenticity and historicity would be more difficult to measure; they could also be biased in such an environment. Furthermore, this type of museum appears less appropriate to apply the phenomenological criteria, which are related to the history of Being. We also deliberately eliminated art museums since they generally offer less technology to their markets. Actually, after visiting several art museums in the Paris locale, our region of residence, we realized that this type of museum was less equipped with technologies. Finally, the choice of the historical setting was instrumental for conducting good research since we were able to obtain a broad license to conduct our field study at the National Center of History of Immigration.

4.2 Set of Scales to Measure the Phenomenological Criteria

Based on the literature review and the phenomenological framework, we developed scales for each phenomenological concept. We followed advice for scale development (Lewis et al. 2005; Moore and Benbasat 1991; Straub et al. 2004). Generally, three steps are suggested for construct development: 1) review the literature to determine the content domain of the constructs, 2) formulate the items for each construct and 3) test the psychometric properties of the scales.

The literature was employed as a rigorous benchmark to evaluate the content of our scales. In the final analysis, we retained three sentences per concept. After defining the domain of our constructs and developing the scales, we pre-tested the scales with a group of college students, after which we refined some sentences. A few months later, a pilot study was also conducted with twenty visitors at NCHI. Visitor comments were taken into account and we modified some sentences for a second time in order to fit better the research context.

4.3 Sampling and Selection

We used the following procedure for surveying NCHI visitors. The researcher was positioned at the museum entrance where visitors borrow the audioguides in order to induce participation by the
maximum of persons. This positioning was strategic because visitors had to return to this desk at the end of their visit to return the audioguides. We took advantage of this time to gather visitor feedback regarding their interaction with the museum technologies. Finally, we collected 111 questionnaires over a period of one and a half months.

4.4 Data Analysis

The descriptive statistics of the sample were computed using SPSS 12.0. To conduct the analyses, we had to delete questionnaires that were unusable because of too many unanswered questions. Hence, we collected 111 questionnaires but our final sample was 106 questionnaires.

The descriptive statistics indicate a higher proportion of females than men, with 66% women and 34% men. This result is not surprising, however, because French statistics on museum attendance also report a higher percentage of women in museums (Cardona and Lacroix 2007). The mean age was 39, with a population aged between 18 to 74 years-old.

We analyzed visitor responses in order to establish the psychometric properties of the constructs. Five out of the six developed scales demonstrate good psychometrics properties, as indicated in Table 1. Except for self-projection with a Cronbach’s alpha slightly lower than 0.70, the other scales range from the acceptable 0.735 to 0.889. However, Nunally (1967) considers that the 0.6 level is acceptable for exploratory research.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Items</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Cronbach’s Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td>CON1</td>
<td>5.21</td>
<td>1.44</td>
<td>0.874</td>
</tr>
<tr>
<td></td>
<td>CON2</td>
<td>5.27</td>
<td>1.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CON3</td>
<td>5.42</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>Embodiment</td>
<td>EMB1</td>
<td>5.87</td>
<td>1.03</td>
<td>0.735</td>
</tr>
<tr>
<td></td>
<td>EMB2</td>
<td>5.38</td>
<td>1.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EMB3</td>
<td>5.21</td>
<td>1.49</td>
<td></td>
</tr>
<tr>
<td>Self-projection</td>
<td>SP1</td>
<td>5.64</td>
<td>1.44</td>
<td>0.678</td>
</tr>
<tr>
<td></td>
<td>SP2</td>
<td>5.51</td>
<td>1.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SP3</td>
<td>5.57</td>
<td>1.21</td>
<td></td>
</tr>
<tr>
<td>Re-enactment</td>
<td>REC1</td>
<td>5.32</td>
<td>1.18</td>
<td>0.782</td>
</tr>
<tr>
<td></td>
<td>REC2</td>
<td>5.40</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>REC3</td>
<td>5.33</td>
<td>1.34</td>
<td></td>
</tr>
<tr>
<td>Possibilities of being</td>
<td>POB1</td>
<td>5.10</td>
<td>1.54</td>
<td>0.860</td>
</tr>
<tr>
<td></td>
<td>POB2</td>
<td>4.91</td>
<td>1.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>POB3</td>
<td>5.46</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>Historicity</td>
<td>HIS1</td>
<td>4.71</td>
<td>1.66</td>
<td>0.889</td>
</tr>
<tr>
<td></td>
<td>HIS2</td>
<td>4.05</td>
<td>1.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HIS3</td>
<td>4.00</td>
<td>1.79</td>
<td></td>
</tr>
</tbody>
</table>

Data analysis also helps in determining the role of IT in visitor experience of the past. We asked NCHI participants to rate their experience with the museum equipment using Likert scales. Since NCHI offers several types of technologies, we decided to survey the museum equipment globally. Hence, we cannot assert which technology contributes the most to the phenomenological criteria.

The mean of each phenomenological construct is reported in Table 2. We also provide a ranking of these constructs from the one that is the best reached through IT to the one that is the less achieved.
<table>
<thead>
<tr>
<th>Construct</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-projection</td>
<td>5.57</td>
<td>1.05</td>
<td>1</td>
</tr>
<tr>
<td>Embodiment</td>
<td>5.47</td>
<td>1.04</td>
<td>2</td>
</tr>
<tr>
<td>Re-enactment</td>
<td>5.35</td>
<td>1.05</td>
<td>3</td>
</tr>
<tr>
<td>Context</td>
<td>5.30</td>
<td>1.24</td>
<td>4</td>
</tr>
<tr>
<td>Possibilities of being</td>
<td>5.16</td>
<td>1.30</td>
<td>5</td>
</tr>
<tr>
<td>Historicity</td>
<td>4.25</td>
<td>1.53</td>
<td>6</td>
</tr>
</tbody>
</table>

5 DISCUSSION AND CONCLUSION

Unexpectedly, it is self-projection that is the criterion that best describes the NCHI technologies. This means that the visitors who used the set of technologies found that these devices helped in projecting themselves into the shoes of the characters. More precisely, 13.2% visitors strongly agree with the fact that IT improves their self-projection, 34% visitors agree and 28.3% slightly agree with this assertion. Visitors likewise felt empathy for the different portraits of immigrants displayed at NCHI through the videos and interactive kiosks. IS research has already shown that IS can provoke emotional reactions (Messham-Muir 2005; Sun and Zhang 2006), such as flow (Koufaris 2002; Skadberg and Kimmel 2004) and enjoyment (Lin and Gregor 2006; Van der Heijden 2004). However, affective behavior like self-projection (or empathy) has been lightly studied in the context of IT use.

Embodiment comes second. Actually the sensory experience offered at NCHI was quite highly rated. Most of the devices convey visual materials (videos, pictures and texts), the audioguides and the TVs diffuse audio content, while the interactive kiosks give an opportunity to have a physical interaction. The visitors cannot really touch the museum artifacts, but very few objects are exhibited in this museum since the scenography relies on IT. Therefore, the possibility to touch is offered through computers and kiosk. Smell and tastes were the two missing senses but it seems that visitors appreciated the variety of sensory experiences. Visitors also agree with the fact that the technologies allow them to relive the historical events in their minds. The mean value for this construct is 5.35 and the levels of satisfaction are the following: 42.3% slightly agree, 27.9% agree and 9.6% visitors strongly agree that NCHI technologies contribute to re-enactment.

Context appears as the fourth construct achieved by the NCHI technologies. Surprisingly, we would have thought that context would be the easier criterion to be met by the NCHI technologies. Tellingly, one of the chief roles of Information Systems is to transmit information to users. But our empirical results contradict this assertion. Nonetheless, the results also indicate that visitors were not entirely satisfied with the cultural and historical background presented by the technologies. These findings support prior research that points out the lack of understanding and context presentation in museums (Hooper-Greenhill 2000; Schärer 1996).

When we scrutinize the ratings in Table 2, we clearly see that the most advanced (and conceptual) criteria, namely possibilities of being and historicity, are also the ones that are the most difficult to reconstitute with IT for museum visitors. Historicity is last in the minds of the visitors, but in the philosophical literature it also represents the ultimate step towards a phenomenological experience. Pointedly, 34.4% visitors are neutral and 30.1% disagree with the statements that NCHI technologies contribute to historicity. Therefore, roughly 65% visitors think that NCHI technologies do not convey a sense of historicity.

NCHI provides several technologies to its visitors, but the types of IT implemented do not provide an entire satisfactory experience of the past. More advanced technologies such as 3D or virtual reality systems are an option to address this lack. It may also be that technologies cannot replace the physical objects displayed in the museums and this explains why visitors have difficulties feeling historicity. The IT role should be to support the visit, enhance the appreciation of the artifacts, but not to enclose...
visitors in a virtual world (Ciolfi and Bannon, 2002). What is clear in the potential use of IT in museums is that technologies that support all the criteria, like those suggested in Monod and Klein’s (2005) framework, are rare. Furthermore, since few technologies have all the characteristics described in the phenomenological framework, museums need to combine multiple devices in order to meet the overall objective of full coverage of the criteria (Monod et al., 2006).

Several researchers have already pointed out the potential of phenomenology for user-driven research and information systems in general (Boland 1985; Introna and Ilharco 2004; Mingers 2001; Monod and Klein 2005). This study confirms this assertion of prior research by showing how phenomenology presents a viable perspective for assessing visitor experience with museum technologies.

Museums represent distinctive institutions and the technologies they provide to visitors generally aim at presenting history, creating reflection, enhancing knowledge and entertaining. Consequently, we decided to rely on Monod and Klein’s (2005) phenomenological conceptualization, a framework which proposes a set of criteria relevant for assessing e-heritage systems. But, even though we rely on this new framework for IS evaluation, links can be created to existing scales. For instance, the criterion “context” can be related to the IS variable “information quality”. Information quality includes notions such as “relevance, understandability, completeness and accuracy” (Petter et al. 2008, p. 239) and it represents a strong predictor of user satisfaction. Our results indicate that context provided by IT is an important factor in hedonic context as well. Therefore, this research can contribute to the evaluation of other hedonic systems (Van der Heijden 2004) and in future research our phenomenological items could be merged or inserted into existing scales that serve for the evaluation of hedonic information systems.

References


VALUE POTENTIAL AND CHALLENGES OF SERVICE-ORIENTED ARCHITECTURES - A USER AND VENDOR PERSPECTIVE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0530.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Service oriented architecture (SOA), Empirical study, Business value of IT, Economics of IS</td>
</tr>
</tbody>
</table>
VALUE POTENTIAL AND CHALLENGES OF SERVICE-ORIENTED ARCHITECTURES - A USER AND VENDOR PERSPECTIVE

Becker, Alexander, Technische Universität Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany, becker@is.tu-darmstadt.de

Buxmann, Peter, Technische Universität Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany, buxmann@is.tu-darmstadt.de

Widjaja, Thomas, Technische Universität Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany, widjaja@is.tu-darmstadt.de

Abstract

This article summarizes the results of 25 expert interviews, regarding the value potential and challenges of the Service-oriented Architecture (SoA) paradigm for users and software vendors.

On the user side, the SoA value potential most often mentioned by experts was agility, followed by process optimization. Cost reduction, through a more efficient development, or consolidation of redundant IS landscapes are also seen as a value potential. In contrast to the literature, reuse of services seems not that significant and is highly debated among the experts. In order to realize the value potentials, users have to overcome certain challenges. Here, the experts mention the financing of the investments, the setting up of a proper governance and the creation of a common understanding of the SoA paradigm as biggest hurdles.

For the software vendors, experts mention that SoA-based solutions create new market opportunities as lock-in-effects are lowered and offerings can be easier enhanced by third-party services. Regarding cost, there is potential to harmonize existing product portfolios or integrate acquisitions more easily. Development cost reduction is less evident at the current stage of adoption. It is assumed that lower lock-in effects also lead to increased competition, which could become a major challenge especially for established vendors.

Keywords: Service oriented Architecture (SOA), Business value of IT, Empirical study, Economics of IS
1 INTRODUCTION

IT departments are currently facing the challenge of determining whether they should adopt Service-oriented Architectures (SoA) as their key architecture paradigm. One reason why they are struggling with that decision is that it is still unclear whether SoA-based solutions can actually deliver the many promised value potentials. The basic idea behind the SoA paradigm is the support of business processes by IT systems consisting of services. Those services are clearly encapsulated, and loosely coupled entities, which deliver a defined business functionality. (Erl 2008, p. 290ff), (Papazoglou & van den Heuvel 2007, p. 389). The SoA concept itself is technology-independent; however, SoA is mostly seen in direct correlation with the Web-Service technology and associated standards.

A first analysis shows that the main research into SoA is focused on technical integration questions, as well as on the design and management of services (Kaczmarek & Wece 2008, Tab. 7, p. 56). So far there is a lack of independent and reliable studies on the actual value of SoA (Vitharran 2007 et al., p.7 f.). Whereas a large set of statements regarding the value potentials for SoA can be found in literature, correct theoretical deductions of the value arguments are very rare. In fact, most of the potentials are only non-revisable claims described at a very generic level. This lack of clear evidence leads to a critical discussion of SoA value in practice (e.g. Graegert (2007)).

The target of our research is to analyze which of the many SoA value potentials discussed in literature can be actually realized in practice. We do not aim for a complete return-on-investment or business case consideration, since such an analysis has to be case-specific (for a full framework to determine the RoI of SoA we refer to Beimborn & Joachim & Weitzel (2007)). The goal is rather to deliver more detailed findings for the value side of the equation, which is usually much harder to determine than the costs. We assume, that the costs for a SoA-based implementation can be quantified using classical IT cost estimation techniques, therefore we focus on the quantification of the benefits. In differentiation to the work of vom Brocke & Sonnenberg & Thomas (2008) who investigate the impact of SoA from a process perspective, we approach the question of SoA value from the perspective of added value of SoA to the entire enterprise. The underlying value concept is the economic value in terms of the EBIT impact, which focuses on the quantifiable effects in terms of revenue increase or cost reduction for the whole company. Effects that contribute only indirect to these two quantitative value drivers are considered as "qualitative" value.

It is important to distinguish two types of companies that can benefit from this value: On one hand, there is the rather small group of software vendors, meaning companies who develop software based on SoA in order to sell it. On the other hand, there are the companies who use these SoA-based software products or develop their own individual solutions while applying the concept (further on named "users"). As there are many similarities in the actual usage (e.g. both groups try to reuse services while developing a system) but also different motivations for the use of SoA, we incorporated both views in our study and try to gain an additional explanatory value from comparing the two sides.

Based on a literature analysis, an explorative expert-study was conducted from end of May till beginning of August 2008. A broad focus was chosen in order to determine as many relevant potentials as possible. After a short introduction into the methodology, this article summarizes the major results of the literature review and 25 expert interviews. Chapter 3 shows the categorization of value potentials based on the literature review. The next chapter then presents the results of the expert interviews by showing the order of mentions and describing the categories of SoA value potentials in more detail. Wherever possible, supporting examples from the interviews are shown. To emphasize that realizing the SoA value potentials requires significant effort, a paragraph on the challenges facing users and software vendors deducted from the expert interviews completes this chapter.

Due to the qualitative method used here, the results are not fully representative. Therefore, the article concludes with a critical discussion of the results and the prospect for the further research opportunities aiming for fully representative answers.
2 METHODOLOGY

Our overall research, targeted to identify the SoA value potentials observable in practice, is structured in three major steps: 1. State-of-the-Art analysis (literature review), 2. Qualitative survey (expert interviews) and qualitative evaluation, 3. Quantitative empirical survey (planned for early 2009).

This article summarizes the first and second step with a focus on the qualitative survey. The expert interviews followed an explorative approach, in order to determine a broad spectrum of value potentials. Semi-structured expert interviews were chosen, as they are a frequently used and well-suited method for this purpose (Denzin & Lincoln 2000, p.653). In this context, an expert is defined as a person, who possesses special knowledge on the subject. This knowledge results from a day-to-day occupation with the concept of SoA in practice. It has been rumoured that Europe, or Germany in particular, would be lagging behind in SoA adoption; however, we have not found any scientific evidence to support that claim. We therefore assumed that the geographical location had no influence on the realization of SoA value potentials. In light of this, the survey was limited to the German-speaking regions for reasons of cost and practicability. The interviews were mainly conducted by phone. As Sturges and Hanrahan (2004) have shown, there are no significant differences between a telephone and a face-to-face interview regarding the quality of the gathered data.

The interviews were targeted at 4 groups (users, software vendors, IT providers and IT consultants), and the structure was based on a modularized interview guideline. In addition to 10 core questions, which were identical for all the groups, a specific set of questions was formulated for each individual group. Note that IT consultants and IT providers were only asked about the SoA value potential for users and software vendors, as they are seen as unable to profit directly from SoA itself. Vendors were asked for the impact of SoA in their own business as well as for the value on the user side. Although statements of vendor representatives bear the risk of being biased for marketing reasons, their opinion on the user value was considered relevant, as they usually have insight in several SoA projects and can therefore provide information which is more representative. The experts were chosen based on internet research and contacted via mail or telephone. The expert status was guaranteed by choosing individuals who were either known through multiple contributions to practical oriented SoA conferences and/or were engaged in several specific SoA implementation projects. 28 experts were contacted and 25 agreed to participate in an interview. The interview sample included eight users, six IT consultants, six software vendors, five IT providers. The users came from five different industries (banking, insurance, telecommunications, retail and logistics). This supported the broad explorative focus of the approach and ensured the diversity needed to create meaningful results. The interviews were recorded digitally before being transcribed. In 10 of the 25 interviews a recording was not possible, either because the interviewee declined to be recorded (six times) or due to technical problems (four times). In these cases a protocol was made on the basis of notes that were taken during the interview (Yin 2002, p. 92). The evaluation was performed according to the methodology of qualitative content analysis according to Mayring (2000).

3 "SOA VALUE" IN THE LITERATURE

An analysis of 20 literature sources revealed 145 text passages describing SoA value potentials. These potentials were extracted by paraphrasing before being aggregated into 16 categories (see Table 1). Although some sources stated different arguments belonging to the same category, they were counted as 1 nomination only. SoA value potentials mentioned only once among all sources were grouped into the category "other". The literature analysis considered the major English books on SoA (including scientific and practical oriented literature), as well as papers from leading IS conferences (01/2000-
05/2008) and IS journals\(^1\) matching the keywords "SoA" and "value"/"benefits". It has been interesting to note that there are only a few sources dedicated to the question of SoA value. Most sources are limited to a small list of potentials with a short explanation of each, however, no explicit discussion or logical deviation can be found. The number of mentions per source differs between three and 16 potentials. Only three of the analyzed sources ((Erl 2008), (Krafzig & Banke & Slama 2007), (Müller & Viering & Riempp 2007)) discuss the subject in a broad manner over several pages. Note that, the method of the qualitative analysis assigns statements to a certain category based on defined rules. This ensures that every category contains at least one clearly defined and therefore distinctive core argument, which minimizes overlaps. However, we found in our explorative analysis that there are dependencies between the categories (e.g. reuse can be seen as facilitator for agility). In our further research, we aim to examine possible dependencies more deeply (rf. chapter 4.1 for first approaches).

Table 1 Ranking of SoA value potentials according to the results of the literature analysis

<table>
<thead>
<tr>
<th>Description of value potential</th>
<th>Mentions</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group of users</strong></td>
<td>N*=20</td>
<td></td>
</tr>
<tr>
<td>Agility (Flexibility and speed for the business)</td>
<td>16</td>
<td>3,4,5,6,8,9,10,11,12,13,14,15,16,18,19,20</td>
</tr>
<tr>
<td>Reuse (Reduction of development investments)</td>
<td>15</td>
<td>1,2,3,4,5,8,9,11,12,13,15,16,18,19,20</td>
</tr>
<tr>
<td>Process optimization (Automation and management of processes)</td>
<td>11</td>
<td>4,5,6,11,12,14,15,16,17,18,19</td>
</tr>
<tr>
<td>Strengthening of IT/Business alignment</td>
<td>9</td>
<td>3,4,6,9,10,11,13,20</td>
</tr>
<tr>
<td>Simplified third party integration (E.g. along Supply Chains)</td>
<td>9</td>
<td>1,2,3,8,9,11,14,15,17,18</td>
</tr>
<tr>
<td>Facilitation of software development (Via standardization)</td>
<td>8</td>
<td>1,4,8,11,13,15,16,19,20</td>
</tr>
<tr>
<td>IT-landscape consolidation (Harmonization of IT applications)</td>
<td>6</td>
<td>8,11,14,16,18,20</td>
</tr>
<tr>
<td>Facilitation of maintenance (Via transparency and capulation)</td>
<td>6</td>
<td>1,4,7,11,14,16</td>
</tr>
<tr>
<td>Enablers of new functionality and business models</td>
<td>4</td>
<td>13,14,16,19</td>
</tr>
<tr>
<td>(Software) vendor independence</td>
<td>4</td>
<td>2,8,11,14</td>
</tr>
<tr>
<td>Improved information quality and availability</td>
<td>4</td>
<td>6,7,11,20</td>
</tr>
<tr>
<td>Management of IT complexity</td>
<td>3</td>
<td>3,10,19</td>
</tr>
<tr>
<td>Simplified Outsourcing</td>
<td>3</td>
<td>2,3,14</td>
</tr>
<tr>
<td>Simplified execution of M&amp;A activities</td>
<td>3</td>
<td>1,4,18</td>
</tr>
<tr>
<td>Risk reduction (Via evolutionary IT modernization)</td>
<td>3</td>
<td>10,13,16</td>
</tr>
<tr>
<td>Improved project management</td>
<td>2</td>
<td>6,11</td>
</tr>
<tr>
<td>Other value potentials with reference to business aspects</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Other value potentials with reference to IT aspects</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>


As there was a very limited discussion of value potentials for the software vendors, this aspect was not included in Table 1. However, it should be mentioned that one source named two main advantages, while discussing the impact of modularization, standardization and loose coupling on the software value chain. Firstly, it enables vendors to outsource the development of services thereby creating new ways of collaborative software development. Secondly, by integrating third party services from niche players, the existing offers from large vendors can be enhanced, leading to a restructuring of the value chain (Messerschmitt & Szyperski 2003, p. 172 ff.).

Concerning the challenges of realizing value by applying SoA, the situation was similar to the question of vendor potentials; explicit literature sources were very rare. Only three challenges could be found consistently in more than one publication: the issues related to governance, complexity and

\(^1\) Note that no relevant articles were found in journals, this is in line with a recently published overview on SoA value papers from Beimborn & Joachim & Weitzel (2008)
financing. Governance is about adapting the organization and the processes in order to efficiently manage a SoA. Doing this right was mentioned by 11 sources as a major hurdle to leveraging the SoA value potentials. (A good overview on the aspects of governance can be found in (Bieberstein 2006)).

The challenge of complexity, mentioned twice, addresses the problem that the fine granularity and distribution of services require significant additional effort in the design phase as well as in the run phase, compared to monolithic systems. Two sources state the challenge of financing the investment in SoA, however with little detail. All other challenges were technology related (e.g. limited data throughput performance of SoA-based solutions). As they were only mentioned sporadically and described very heterogeneously, they were aggregated in one category "Technology issues" (seven mentions). Several sources refer to "success factors", which could be interpreted as the logical counterpart of challenges. However, they are mostly generic to IT projects and not SoA specific, therefore they could not be related to the question of SoA value and are not considered in our study.

It is interesting to note that several of the examined sources on the topic have at least one chapter on SoA benefits (e.g. "Advantages and Motivation" in (Krafzig & Banke & Slama 2007) chapter 11). However, apart from (Erl 2008) none of these sources have an explicit chapter on "challenges" and most work merely addresses the issues implicitly while describing how to implement SoA. This observation, in combination with the fact that several books have very positive titles such as "Succeeding with SoA" (Brown 2007) while only one subtitle (McGovern & Jain & Little 2006) speaks of "Challenges", strengthens the hypothesis that there is a lack of critical and profound discussion of the actual value of the SoA paradigm.

4 RESULTS FROM THE INTERVIEWS REGARDING VALUE POTENTIALS AND CHALLENGES OF THE SOA PARADIGM

The following paragraph shows the main results from the expert interviews concerning the question of what kind of value potentials users and software vendors can draw from applying the SoA concept. It will conclude with a discussion of the challenges which complicate the realization of those potentials. Regarding the current hype about SoA and the marketing activities of software vendors and consultants, it is important to mention that all the experts’ arguments were critically reviewed. The experts were consistently asked to provide examples during the interviews. Most of the interviewees (21/25) could base their argumentation on personal practical experience. Many of them were also clear about the fact that the value of some potentials is hard to quantify or to measure - which is why the value discussed here is only "potential" or "expected" value rather than actually "realized".

4.1 Value potential for users of SoA based software systems

Table 2 provides an overview of the interview results sorted by the number of positive mentions of the SoA value potentials for users in the expert interviews. The rightmost column indicates whether the experts’ opinion was homogenous or not on this potential. Potentials marked with a "+" were not doubted among the experts and based on several arguments and at least one specific example. If there were no specific examples and many weak arguments, the potential was marked with "+/-". A "-" shows that the potential was debated among experts and there were even some critical voices negating this potential. As you can see in table 1, the categories were grouped by two dimensions: a) the type of value they are providing according to our value model: revenue, cost (logically split into business- and IT-related cost) and qualitative effects; b) the "layer of impact" which classifies the layer of the company the value is created in. The layer structure follows the model of Österle (1995), who states that the design of information systems has impact on the strategic layer, the process layer, and IT layer of a company. This structure helped to minimize redundancies (e.g. all potentials with a effect on IT-cost have to consist of mutually exclusive arguments) and allow to explain dependencies between categories, which mainly exist between layers (e.g. reuse on the IT-layer can be seen as a precondition for agility on the strategy layer). Due to space limitations, only those potentials mentioned by more than 10 experts will be discussed in the following.
All interviewed experts mention the improved agility of an enterprise as a SoA value potential. The arguments and examples thereby are centred around two main aspects: The first aspect is speed: agility is expressed by a faster adoption of the IT to business changes (measured by a reduction of the so called "time-to-market" of change requests). Several experts report a significant reduction of the implementation time of new requirements. Four interviewees (three consultants and one user from the industry) mention the telecom industry as a good example in this context. The introduction time of new tariff models, can be reduced by applying the SoA concepts. Starting with the completed concept in product management to the actual positioning in the market, a time saving of between 20% and 50% is stated by three experts, compared to the old architecture. Two interview partners phrase the advantage: "We came down from a question of months to a question of weeks". In this case the effect can be explained by the decoupling of customer centric from network- and infrastructure-related systems by applying the concepts of SoA. Following this, a new product, e.g. a so called bundle-offering, is maintained on only one system and can be changed independently of the complex systems administering the telecom network infrastructure. The second aspect of agility besides speed is an increased flexibility in IT systems. Most interview partners explain this increased ability to change, by the architectural separation of dynamic logic (process) from the static logic (service). The dynamic logic is more likely to be subject to changes and can be adapted more easily. Although agility can be separated into the two aspects mentioned it was contained as one category. As agility can be seen as an overarching term for several aspects, according to the definition of Yusuf et al. (1999): "Agility is the successful exploration of competitive bases (speed, flexibility, innovation pro-activity, quality and profitability)". During the interviews most interview partners did not differentiate between the terms agility, speed and flexibility and even used them synonymously.

The majority of experts agree that reuse is a value potential of SoA (20/25), however, there are also very critical voices. Two interview partners express their concerns that the design of reusable services is very difficult and they never saw a real value from reuse in practice. Especially in the early phases

---

2 This and all following citations are translated from the German transcript into English.
of SoA maturity, it would be very difficult to achieve a consensus on the functionality among the stakeholders, as there is no experience about how a reusable service would look like. Another expert points out that reuse is not a new subject in computer science. He claims that earlier approaches - like modularization or object orientation – did not fulfil the expectations of high reuse either. There are also limitations among those experts who were positive regarding reuse. Four experts state that reuse exists and is a value potential, but it is limited and can only be achieved after a long period of SoA experience. They explain that there are only a few services in an enterprise that can be reused frequently (e.g. services for the management of customer or product data, or basic services like printing or calendaring). Three experts cite analyses that measured a reuse factor (i.e. average reuse-of one service in different contexts in the same period of time) between 1 and 2. One software vendor reports on his customer, an Austrian bank, which measured a reuse factor of 1.3. An IT-consultant cites an internal analysis measuring a reuse-factor of 1.5 among several companies. One user states that, in his company, a few services were frequently reused (up to 14 times), however, the majority, and therefore also the average, had a reuse factor below 2. Contrastingly, there are interview partners who strongly believe in reuse. An IT provider reported the development of a new core banking system able to be finished ahead of the planned time because many services (approx. 10 %) could be reused. He states the example of a service for transfers which is reused for the web portal as well as for the ATM and the teller application. Two interview partners from the user side point out the business value of reuse in this context: Reusing an existing service allows standardizing functionality; for every process using the same proven business logic is applied enterprise wide. One reason for the obviously contradicting opinions about reuse might be the variety of interpretation of the concept. Most experts understand reuse to be the usage of the same service in multiple processes at the same time ("multiple usage"). Others also mean the reuse of one service in a different context at another time, which means that the service is "used again" but only by one user. Another group of experts describes a "reuse" of legacy systems by encapsulating existing functionality and providing it via a Service interface. This approach allows the prolongation of existing systems’ lifecycles by using their functionality in new processes with new user interfaces. Some experts also phrase this value potential explicitly as "prolonged usage" instead of "reuse".

The vast majority of experts (17/25) report that process optimization in particular is the main motivation for users wanting to adopt SoA. Especially vendors and consultants report that there is a strong demand among their customers to integrate and automate process flows, especially across line of business or even enterprise boundaries. SoA is seen as the best basis for Business Process Management (BPM) systems, as it allows the process-oriented integration of IT systems, leading to a more automated IT support of processes. The value can be measured by known process key performance indicators, e.g. shorter cycle times or reduction of required manpower. One user reports that process optimization was the only reason why the insurance company he works for adopted SoA. By automation of a claims settlement process across several silo systems, the cost could be cut to 1/50 of the old manual process. It is obvious that not the full amount of this savings potential can be assigned to the SoA-based IT solution, however the interviewee pointed out that SoA was a necessary precondition to realizing this value, as other attempts to realize the potentials failed before. It is also worth mentioning that according to three experts, SoA is especially valuable when it comes to the harmonization and standardization of processes in a globally acting enterprise.

17 experts state the value of facilitation of software development driven by the standardization of interfaces which leads to less mapping efforts. Standardization of tools and methods for model-driven development reduces coding. However, there are, as yet, no concrete examples or data points available for this potential and some experts are uncertain whether this potential could be actually captured.

The potential of IT-landscape consolidation is directly tangible in the context of systems operation. Implementing a certain functionality only once in a IT-landscape is a core concept of SoA and allows

---

3 Note that this category does not include efficiency effects from reuse as reuse is a category of its own.
the identification of redundancies. Two interview partners point out that this value potential can already be leveraged by the mere introduction of SoA Management tools, such as the development of a domain map, without actually implementing SoA at a technological level.

The subject of improved information quality and availability addresses the typical "Master-Data Management" questions. Through consolidation of customer data in a central service, for example, the integrity and consistent availability of that information within the enterprise is improved.

The increased IT/Business Alignment is debated, which is partly due to the different interpretations of that term. If a better IT/Business Alignment is defined as a better fitting of IT solutions to the business problems, the majority of experts agree. This is explained by the process orientation of SoA, which leads to a business proximity. Some experts interpret the IT/Business Alignment as the cooperation between IT and business departments. The picture regarding this interpretation is more unclear. Some experts state that SoA enables IT to make architecture questions more understandable for people outside the IT department and creates "a common language". However, some users also state that there has been no improvement in IT/Business cooperation through the introduction of SoA; on the contrary, the already existing problems only became more obvious.

The idea of creating new functionality via SoA is valued more in practice than in literature. Interestingly it is not so much about providing functionality as a service externally, but rather about offering new products by integrating third party services, or even building up new business models on that idea. One consultant reports the case of a retail company that integrated the Web-Services of a bank in order to offer the customer a loan at the point of sale. Two other experts name the Hypoport AG as an example of an SoA based business model. Hypoport acts on the technological basis of a SoA platform as a broker for financial services. (For a more detailed description of Hypoport’s business model see (Hypoport 2008)). These examples could also been realized in a different way. However, the SoA paradigm and the underlying technology ease the integration and make it cheaper. This lowers the integration hurdles and new, collaborative business models become more attractive.

Concerning the value potentials on the strategy and process layer, it can be stated that most potentials identified in the literature analysis can be realized in practice. Regarding the SoA value potentials on the IT layer, the discussion is much more controversial. The majority of IT related potentials are marked with a "+/−" or "−" in table 2, meaning that their validity is low in average.

4.2 Value potential for software vendors

Four different value potentials were identified for software vendors among eight interviewees (Consisting of the six vendors and two consultants).

All vendors, in particular smaller ones, state that the standardization allows them to integrate in existing product landscapes from other vendors more easily (Two already leveraged this potential). It is also easier to sell software as a service (SaaS) when software is based on SoA principles. These two arguments were summarized in the value potential category of new market opportunities.

Most experts (5/8) also agree that an integration of acquisitions, or already existing heterogeneous portfolios, is much easier when software is based on SoA principles. The standardized interfaces and clear capsulation eases the combination of different products to an integrated solution and helps to reduce redundancies. Two vendors have already actually realized specific value from this potential.

Vendors can also enhance their product offerings by integrating third party services. This results in a more complete and therefore more attractive product for the customer. Based on capsulation and standardization this enhancement can be done quickly and the vendor does not need to invest in the development of the new functionality himself. Some experts also claim that the capsulation leads to the effect that software offerings could be more modularized (i.e. sold in smaller pieces), which allows vendors to offer customers a more individual product instead of a big "all-in-one"-solution. This could attract new customers that are interested only in a dedicated part of the existing solution. Specific
statements concerning revenue or cost effects of this potential could not be garnered from the interviews. However the arguments were consistent among all experts and therefore the validity is "+".

<table>
<thead>
<tr>
<th>Description of value potential</th>
<th>Impact Layer</th>
<th>Impact Type</th>
<th>Expert Mentions</th>
<th>Validity in interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>New market opportunities</td>
<td>M/S</td>
<td>R</td>
<td>6</td>
<td>+</td>
</tr>
<tr>
<td>Simplified integration of acquisitions</td>
<td>PD</td>
<td>C</td>
<td>5</td>
<td>+</td>
</tr>
<tr>
<td>Enhanced offering</td>
<td>M/S</td>
<td>R</td>
<td>3</td>
<td>+</td>
</tr>
<tr>
<td>Facilitation of software development</td>
<td>PD</td>
<td>C</td>
<td>2</td>
<td>+/-</td>
</tr>
</tbody>
</table>

**Table 3 Ranking of SoA value potentials for software vendors according to expert interviews**

Two of the six software vendors also explained that they had realized a facilitation of software development through SoA. Others said it was too early to measure that, but they would also expect benefits in this area, one expert however was sceptical. Therefore this potential was marked with "+/-".

Comparing the results from the qualitative analysis of the vendor potentials and the results from the user side, only the SoA value potential "facilitation of software development" can be found on both sides. The judgement on this potential is consistent, as both groups expect an improvement, whereas there are only few proof-points yet. It is interesting to see that reuse was not mentioned as value argument from the vendors themselves. This supports the hypothesis from the user side that this value potential is – at least to the current stage of adoption – very hard to realize. Although the consequences are different, the basic arguments for SoA value are the same, as they relate to the same core aspects of SoA (especially capsulation and interface standardization). Overall, one can summarize that according to the number of mentions on the vendor side SoA value is seen to be more about creating new opportunities (in terms of revenue) than improving efficiency (in terms of reducing cost). This pattern could also be identified on the users' side. The matching results of the analysis confirm the value of comparing both views. However, due to the small number of statements on the vendor side, a more comprehensive and structured data gathering is needed to derive further results.

4.3 Challenges for vendors and users in realizing the potential values of the SoA concept

During the interviews, it became quite clear that SoA does not only bring new value to a company, it also introduces new challenges, or at least different ones than typical challenges of new IT technologies. Table 4 gives an overview of the challenges related to the concept of SoA, identified during the interviews from the user perspective. As the discussion of challenges was less controversial then the one on value potentials, all statements can be seen equally valid and therefore no validity was assigned. Some of the statements were not directly focused on the question of value realization, but on the question of challenges with SoA adoption in general. However, it is obvious that a challenge hindering a company to adopt SoA is a challenge to get value out of SoA too.

According to the experts, the most important challenge is the question of financing of the investment.

---

4 Note that the stated "Impact Layers" differ from the user side. As the model of Österle (1995) does not apply for software vendors, we assign the identified potentials to the steps of the software business value chain which consists in our definition (based on Messerschmitt & Szyperski 2003) of 8 steps: 1) "Research", 2) "Product Development", 3) "Documentation", 4) "Marketing and Sales", 5) "Consulting", 6) "Implementation", 7) "User Instruction", 8) "Maintenance".

5 Related to SoA means in this context, that the challenge has its root cause in one of the core aspects of SoA according to our definition in chapter 1 ("process orientation", "loose coupling", "capsulation", "(Interface) standardization"). Statements on general project challenges that are not SoA specific (e.g. "Find a strong project manager") were not included in the analysis.
In the past, the different systems could easily be assigned to one owner on the business side. Only a rather small part of shared applications were financed by a collective budget. SoA incurs a much bigger part of collective infrastructure. In addition, the services (especially when there is reuse) do not have explicit owners anymore either. The SoA enablement of legacy software, as well as the infrastructure investments, requires significant investment and it is, as yet, unclear in many companies how this cost should be divided among different stakeholders.

Table 4 Challenges for users in realizing the potential values of the SoA concept

Another big challenge to realizing SoA value is the set up of a proper business case because of the long-term nature of the value: As most of the potentials are only expected in the long run, and therefore are hard to calculate, it is difficult to show a credible calculation proving the investment is worth it. According to the experts, many companies that have so far adopted SoA did it because they "believed" in it, rather than having a solid business case calculation. It is obvious that if nobody really thought about the specific potentials upfront it is hard to capture the value afterwards.

As already explained in chapter 1, the adaptation of the organization to SoA is difficult, 11 experts mention aspects concerning this challenge, which were categorized by the term "governance".

As one can see from the issues described with the definition of the value from IT/Business Alignment or reuse (rf. chapter 4.1), there is still much uncertainty about what SoA actually is. The experts report that this effect results in a significant effort for a company to define what is actually meant by the term SoA, and to align all people, especially on the business side, with one common view. This is why 10 experts name "comprehension" as a big challenge in realizing value from SoA.

Although integration is eased by common syntactic standards, there is still the issue of missing semantic standards. Until there is a common semantic model available, integration will still require project effort, and the IT world is far away from "plug-and-play" scenarios. Compared to the issues before, this challenge seems rather small, as it is mentioned only a few times. The same is true for the design of a service or sourcing a service out of the box from a software vendor. Regarding technological issues like security and performance, there are also voices that explicitly do not see them as issue anymore. Software vendors in particular claim that nowadays the issues regarding technology have been solved, and most scenarios can be supported via SoA securely and with high performance. Interestingly "complexity", a challenge identified in literature, was not mentioned in the interviews, in contrary nine experts claim that SoA helps the users to manage complexity better (rf. Table 2).

Regarding the software vendors, data about challenges could not be collected. In the interviews the vendor experts were quite reluctant to talk about challenges for themselves. Although the results were guaranteed to be anonymous, they might have been afraid of undermining their reputation as leading SoA vendors by talking about these issues in public. This hypothesis of hidden agenda in positive marketing is supported by another phenomenon that occurred during the interviews. Regarding the question "What is your unique and differentiating value proposition in the SoA market?", five out of six vendors answered that their solution is the only one that would offer the complete range of tools. Whereas four out of eight users asked for vendor strategy answered that they would source a best of breed solution, as no vendor could deliver a full tool portfolio. Despite this lack of data, it is possible
to logically derive some hypotheses from the results on the user side: Apart from the question of internal accounting, the challenges are assumed to be the same for vendors and users of software based on the SoA paradigm. It is imaginable that the vendors also are facing the challenge of justifying the architecture shift in their products economically. They might have struggled in adapting their governance in order to manage a service rather than a product portfolio. And it seems possible that they are also struggling with missing semantics, technology and the difficulties of service design. Creating new market possibilities for one vendor because of standardization, and therefore lower integration hurdles, means a direct threat for those vendors that have profited from a lock-in effect so far. The big, established vendors in particular could face increasing competition when it becomes easier for niche players to integrate and enhance existing applications with their services.

5 CONCLUSION AND FURTHER RESEARCH

Summarizing the results discussed, it can be stated that the SoA value potentials postulated in literature are, at least partly, evident in practice. On the user side, the value potentials "process optimization" and "agility" seem to be most evident in the current phase of the SoA adoption process. Nevertheless, some of the value potentials remain highly debated. Generally, our interview series identified more value potentials on the user than on the vendor side. On the user side, the provable potentials are mainly found on the strategy and process layer, IT (cost) potentials are less evident. This matches the results from the vendors, where the value arguments were centered around new revenue opportunities instead of development cost reduction. The experts pointed out that the majority of the promised SoA value potentials, especially those regarding cost reductions on the IT side, are only achievable in the long run. As the current experience horizon with SoA is too short for most users, those value potentials are still only "expected".

As the expert interviews were not designed to gather specific and comparable data points about the absolute economic impact of the potentials, it is hard to bring them into a relation. It is imaginable that the economically quantifiable and measurable benefits from process optimization might be bigger than the potential from facilitation of software development, although both can be seen as equivalently important potentials according to table 2. Concerning this question, a broad empirical study could deliver further insights.

It also became transparent that most value potentials of the SoA concept can only be realized with significant investments in infrastructure and service enablement of applications. In addition, governance questions have to be solved and a common understanding about SoA in the respective company has to be achieved to make the realization of value possible. The discussion with the experts showed that it is very hard to generalize the value of SoA. Therefore there is no generic answer and every user company and software vendor has to do its own calculation concerning the individual profitability of a SoA adoption. It strongly depends on the individual usage scenario, such as the industry or the functional context which the architecture is used in. Based on the interviews, we developed first hypothesis on such influence factors, which will be tested in further research.

We are currently working on an even more detailed structure of the SoA value potentials, allowing to identify dependencies and to derive clearly defined and representative indicators to measure the potentials. As the study presented here was not designed to be a representative analysis, a quantitative empirical study evaluating such indicators could be a promising field for further research. We hope that the proposed categorization of value potentials and challenges is a first step towards a more precise and fact-based value argumentation in the controversial discussion about the value of SoA.

6 ACKNOWLEDGEMENT

We would like to thank the anonymous reviewers and Helge Buckow for their many useful suggestions on our research. Thomas Widjaja’s research is supported by a research grant from the FAZIT-Stiftung.
7 REFERENCES


FACTORS INFLUENCING THE ALIGNMENT OF SOA DEVELOPMENT WITH BUSINESS OBJECTIVES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0645.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Alignment, Service oriented architecture (SOA), Information Systems Development (ISD), Exploratory framework</td>
</tr>
</tbody>
</table>
FACTORS INFLUENCING THE ALIGNMENT OF
SOA DEVELOPMENT WITH BUSINESS OBJECTIVES

Jari Antikainen, Tampere University of Technology, PO Box 541, 33101 Tampere, Finland,
jari.antikainen@tut.fi
Samuli Pekkola, Tampere University of Technology, PO Box 541, 33101 Tampere, Finland,
samuli.pekkola@tut.fi

Abstract

Service-Oriented Architecture (SOA) has gained focus as a driver for bridging and aligning business
and IT-oriented views in information system development. The critical aspect of successful SOA is
aligning technology and business; without proper alignment, the full potential of SOA will not be
achieved. The current academic literature includes only few empirical studies on business-IT align-
ment aspects related to SOA adoption. In this exploratory study we explore factors influencing suc-
cessful SOA implementation. We interview IT and business people from nine organisations appearing
as SOA forerunners in Finland. Our findings indicate that successful SOA adoption is affected by sev-
eral factors varying from organisation culture to processes and methods, communication and technol-
ogy. The findings form a basis for future SOA research and a set of guidelines for practitioners.
1 INTRODUCTION

Aligning information technology (IT) with business, i.e. “applying IT in an appropriate and timely way, in harmony with business strategies, goals and needs” (Luftman & McLean 2004, p. 90) has been the main objective of IT executives for years (Luftman & McLean 2004, Luftman & Kempaiah 2008). This is not surprising as IT has been considered as a means of gaining competitive advantage – if used appropriately (c.f. Ragowsky & Gefen 2008; Yoon & Carter 2007). However, as Luftman & Kempaiah (2007) argue, achieving appropriate alignment is not straightforward but involves consideration of interrelated components such as communication, value, governance, partnership, scope and architecture and skills.

Business-IT alignment has often been seen as a strategic asset (Bergeron et al. 2004; Peppard & Breu 2003; Cumps et al. 2007). Under these circumstances, it is evidently interpreted as a duty and as an obligation by IT executives (Silva et al. 2007). Alignment is consequently every day business for organisations (Cumps et al. 2006). Because of dissimilar strategies, processes, practices and cultures alignment is perceived as the unique, inimitable competence of an organisation where learning from others and exploiting their experiences, models or practices is difficult. As Silva et al. (2007) pointed out, this emergent phenomenon indicates that different alignment models and methods are not necessarily known by IT executives.

However, service-oriented architecture (SOA) has gained a lot of interest lately (c.f. Ren & Lyytinen 2008; Baskerville et al. 2005) since service oriented organisations and SOA are seen as revolutionising methods for managing information and streamlining operations (Vitharana et al. 2007). Yet SOA architecture itself seldom provides significant benefits. Baskerville et al. (2005) stated that SOA is a means for attaining organisational agility. In other words, SOA development is a continuous process where business-IT alignment needs to be invariably ensured. In their words: “SOA is only one ingredient in a recipe for a much larger meal” (ibid.). Although SOA is usually perceived as technical component (Antikainen et al. 2009), other non-technical stakeholder groups also have to be involved for successful business-IT alignment. This necessitates shared knowledge and mutual understanding about the dimensions of alignment between IT unit and business department (Beimborn et al. 2007). Achieving this kind of reciprocal understanding is difficult indeed. Silva et al. (2007; see also Ciborra 1997) even showed that there is no mutual understanding about the alignment concept between IT executives and IS researchers.

Understanding business-IT alignment in organisations is our focus. We try to identify the main factors influencing SOA development from the business-IT alignment viewpoint. Understanding these factors provides a basis for further research and theory building, and helps researchers to articulate their findings to (IT) executives and practitioners in SOA development initiatives. The data for the case study on SOA development and deployment was derived through twelve interviews conducted in nine large organisations. Next we take a tour of related business-IT alignment research from several perspectives. Then we present our research methods and the findings. The paper ends with discussion and concluding chapters.

2 RELATED RESEARCH AND THEORETICAL BACKGROUND

Business-IT alignment has been defined as the degree to which the mission, objectives and plans articulated in business strategy are shared and supported by IT strategy (Reich & Benbasat 1996). Henderson and Venkatraman (1993) follow this by stating that alignment is the degree of fit and integration among business strategy, IT strategy, business infrastructure, and IT infrastructure. These definitions claim that alignment is static. Campbell (2005) took a different approach by defining alignment as business and IT working together to reach a common goal. Yet this approach does not explicitly point out the strategic aspect of alignment. Chan and Reich (2007) identified several dimensions of alignment apparent in the IS literature: strategic and intellectual dimensions, structural dimensions,
informal structure, social dimension and cultural dimension. These dimensions emphasise the need for alignment on several levels: organisational, system, project and individual and cognitive. Alignment is correspondingly a multidimensional subject ranging from strategies to practices, from processes to activities bridging business to IT, and vice versa.

Business-IT alignment has been studied extensively in the IS literature. Some recent studies discuss the strategic level of alignment (Bergeron et al. 2004; Griffiths & Finlay 2004), the evolution of strategic alignment (Peppard & Breu 2003), business process alignment (Cragg et al. 2007; Beimborn et al. 2007), cultural aspects of alignment (Silvius 2008, Bieberstain et al. 2005), aligning users, organisations and IT goals (Singh et al. 2008), the role of informal networks (Haried & Ramamurthy 2006), measuring the success of alignment (Denford & Chan 2007) and several models and methods (e.g. Mueller et al. 2007; Griffiths & Finlay 2004; Chen 2007; Dertzi 2008; Klose et al. 2007; Henderson & Venkatraman 1993). Yet these studies do not correspond with the call by Chan & Reich (2007) to examine the practical processes of alignment; i.e. how alignment really takes place, what factors frame it, and how managers may benefit from these studies.

Different factors of alignment have been studied. Henderson and Venkatraman (1993) argued that an organisation’s external fit is influenced by IT scope and system characteristics (reliability, flexibility and IT governance), and internal fit by the IS architecture, IS processes and skills required to manage and operate information systems. Chan and Reich (2007) reviewed the challenges in attaining alignment as being related to knowledge (unknown corporate strategy, lack of awareness or belief of the value of alignment, lack of industry or business knowledge), related to locus of control or the status of IT, and related to organisational change (see also Ciborra 1997). An influential study by Sabherwal & Chan (2001) discussed the dynamics of alignment. Although they identified several factors and challenges, their focus was rather on organisation level issues and not on the process or operation level factors. In fact, this kind of focus on theoretical issues is common in the IS literature (Gutierrez et al. 2008). Baskerville et al. (2005) took analysis to the SOA context. They identified the challenges of alignment as application integration, value reconfiguration, value preservation and meagre use of agile development approaches.

Individual factors influencing business-IT alignment have also been examined. The role of management and their commitment and involvement has been recognised as a key factor for successful alignment (e.g. Cumps et al. 2007; Henderson & Venkatraman 1993; Brown & Magill 1994; Benson et al. 2004; Earl 1989; Kean 1993). Yet management’s involvement is not enough. Communication and knowledge sharing have also been identified (e.g. Reich & Benbasat 2000; Chan et al. 2006; Luftman et al. 1999; Campbell 2005; Sledgianowski and Luftman 2005). These studies usually emphasise a close relationship and continuous cooperation between business departments and IT unit. The influence of IS architecture, its governance and technology, and social skills have also been acknowledged (c.f. Henderson & Venkatraman 1992; Brown & Magill 1994, Luftman 2000) in addition to the ability to justify up-front ICT investments as they inevitably have an impact on alignment (Cumps et al. 2006; Mueller et al. 2007).

SOA has been considered as a means for alignment although it has been approached widely, ranging from technical solution and implementation methods to enterprise architecture and IT governance tools (Lyytinen & Rens 2008). Legner and Heutschi (2007) argued that SOA adoption can most beneficially be targeted as standardised integration infrastructure, for decoupling application domains, or for flexible user/business process integration. Yoon and Carter (2007) perceive SOA as IT strategy, where it requires extremely high levels of business-IT alignment to achieve benefits. Other key factors to successful SOA implementations include SOA registries, governance, effective pilot projects, top management support, trust among business units, personnel training and change management (ibid.). Chen’s (2007) BITAM-SOA framework attempts to consider all these aspects by combining SOA, engineering alignment methods, enterprise architecture, and business performance management. She suggested that the impact of SOA adoption can be divided into three different factors: communication, architecture and governance, each being a critical success factor for alignment. Alignment via architecture is influenced by architectural adaptivity, misalignment detection, misalignment prevention,
enterprise modelling, and integration (of workflow management application portfolio management). Alignment via governance necessitates regulatory compliance. Alignment via communication stresses management asking for support and implemented vision.

The studies merely focusing on SOA success do not provide adequate results that would explain how business-IT alignment is realised in SOA development. From a technological perspective, a successful SOA project may not fulfil the business objectives while still being successful in technical excellence. In such a situation, proper alignment is not achieved. Bridging the business-IT gap requires 1) a structured view of a business that facilitates its strategic and operational analysis and is a familiar representation to IT professionals, 2) a rigorous method to translate this structured business view to the appropriate (service-oriented) IT layer, and 3) new build and runtime technologies suited to the new IT layer (Cherbakov et al. 2005). Shiskov et al. (2006) argued that adequate business-application alignment can be achieved only if the initial business model is a valid reflection of the real world, and provides a suitable basis for future models. As bridging the gap between business and IT is not trivial, both the architecture style and development approach should facilitate the alignment. Particularly the architecture approach emphasises the importance of service-orientation in establishing enterprise-wide perspectives on organisational resources (Nuffel 2007). This way SOA adoption improves architectural agility, which further positively impacts on IT’s abilities to align with the changing business environment. Shiskov et al. (2007) summarised three properties for desirable SOA development. First, application architecture must allow the usage of SOA infrastructure, second, ‘loose coupling’ should be enforced and third, application architecture must fit within the business context. This is in line with Klose et al. (2007) and Yoon and Carter (2007), who supplemented the list by the need for systematic identification of IS functions to be implemented as SOA services.

## 3 RESEARCH METHOD

This case study is based on a set of semi-structured theme interviews, where the interviewees’ interpretations of different SOA related issues were discussed. The questions were developed by the authors according to their interests. After a pilot interview in an organisation, unclear or incorrectly understood questions (3 out of 20) were reformulated. The results from the pilot interview were not included in the study. Four groups of questions were used. First, background information about the interviewee, his/her organisation, and the project where SOA was applied was gathered. The second set focused on the expectations and challenges to SOA adaptation in general. The third set was about a certain SOA adoption project, while the last set was about business process modelling and its role in SOA projects.

<table>
<thead>
<tr>
<th>Case</th>
<th>Branch of industry</th>
<th>Interviewee</th>
<th>Job title</th>
<th>IT/Business</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case A</td>
<td>Public administration</td>
<td>I1</td>
<td>CIO</td>
<td>Business</td>
</tr>
<tr>
<td>Case B</td>
<td>Metal industry</td>
<td>I2</td>
<td>Portal manager</td>
<td>IT</td>
</tr>
<tr>
<td>Case C</td>
<td>Telecom industry</td>
<td>I3</td>
<td>Account manager</td>
<td>Business</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I4</td>
<td>ICT architect</td>
<td>IT</td>
</tr>
<tr>
<td>Case D</td>
<td>Financial industry</td>
<td>I5</td>
<td>Development manager</td>
<td>Business</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I6</td>
<td>CIO</td>
<td>IT</td>
</tr>
<tr>
<td>Case E</td>
<td>Financial industry</td>
<td>I7</td>
<td>IT architect</td>
<td>IT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I8</td>
<td>Project manager</td>
<td>Business</td>
</tr>
<tr>
<td>Case F</td>
<td>Logistics</td>
<td>I9</td>
<td>Integration architect</td>
<td>IT</td>
</tr>
<tr>
<td>Case G</td>
<td>Public administration</td>
<td>I10</td>
<td>Director</td>
<td>IT</td>
</tr>
<tr>
<td>Case H</td>
<td>Public administration</td>
<td>I11</td>
<td>Chief Architect</td>
<td>IT</td>
</tr>
<tr>
<td>Case I</td>
<td>Retail trade</td>
<td>I12</td>
<td>Chief IT-architect</td>
<td>IT</td>
</tr>
</tbody>
</table>

Table 1. Interviewees and their background

We decided to focus on multiple organisations instead of one. This approach was believed to provide a broader understanding of the phenomenon and more generalisable results. Potential organisations for
The study were identified according to their public presentations in several national practitioner-oriented SOA conferences where they presented themselves as having a significant amount of SOA experience. SOA consultation houses and tool vendors were excluded as we wanted to avoid their market speech, and tool and supplier dependence. After potential organisations (12) were identified, they were contacted for two sets of interviews: one from the business side, the other from the IT side. This follows the suggestion by Silva et al. (2007) that in order to gain insights both stakeholder groups need to be involved. Unfortunately not all organisations responded or granted two interviews. The set of organisations, persons interviewed and their backgrounds are presented in Table 1.

Each interview followed the same order of questions, which were sent in advance to the interviewees by email. The interviews, each lasting 60 to 90 minutes, were audio recorded and transcribed. Data analysis was conducted according to the principles of grounded theory methodology (Urquhart 2001). First, without collaborating with any of the other four members of the research team, each researcher made preliminary coding using varying coding schemes. The coding process continued until theoretical saturation of the data was considered to have been achieved. Next, the researchers held two workshops where the results of the first round of analysis were discussed and refined into a final result set containing harmonised presentation of the previous results. Finally, the final outcome was reflected against the findings of a literature review and the sample data to ensure the validity of study.

4 FINDINGS

The analysis of empirical data revealed 11 factors related to the business-IT alignment of SOA development. These are discussed in detail.

Organisation culture

Organisation culture has various impacts on the business alignment of SOA development. For example, when business and IT had traditionally operated in a close cooperation, SOA development seemed to be better aligned with IT objectives (as expected) than in organisations where such cooperation was not a part of organisational culture. A developer manager described this as follows: "Traditionally we have had strong IT and business. They have never been that much separated... In our organisation business-IT collaboration is happening on a daily basis."

Furthermore, business unit oriented organisational structures had an inhibiting influence on the realisation of business-IT alignment in SOA development. The existence of business silos facilitates a culture where individual business units have to take ownership of specific data, processes and even IT systems. In such situations services, data sharing and open interfaces are not favoured. This was believed to be due to the fear of losing power and influence within the organisation. A CIO from the same finance industry organisation also described their successful SOA process as an outcome of good business-IT alignment. Yet he emphasised organisation culture, not only good alignment. "Our way of working has been on a common good basis for a long time. Our organisation structure also supports it. We have organisation wide services like finance, human resources, Internet services and customer interface services."

As each service unit was producing services to others, the transformation to SOA thinking was minimal.

Another problematic area was data confidentiality. A public sector interviewee demonstrated that business unit people consider the data in their IT systems to be so confidential that it could not be exposed even within the organisation. "[Data confidentiality was] the biggest problems. It was caused by change resistance... For example, a few people from [unit X] had realised that the data in their systems was secret and valuable that it could not be shared with others."

---

1 The questions were the same for both business and IT personnel.
2 Translations from original Finnish data by the authors.
A similar phenomenon was also identified on the IT side. Openness and the sharing of data and IT services were perceived to decrease the importance of the IT unit making them “replaceable”. The same interviewee continued that “in the IT department the valuable asset was the technology. They felt that along with more open communication environment, their power would be taken away... For some, for this reason, information sharing is really difficult.” [I1]

Competences

The skills needed in the design phase, such as business process modelling and service design, were considered particularly important for SOA application development. This was underlined, especially in the business process oriented projects. An architect argued for the importance of the business process modelling in SOA development for all participants. However, he emphasised that this was critical for business people as it is the main competence they bring into the project. “They [business people] should be totally unaware of the technical aspects related to information systems but they should have an understanding of processes and process development in general.” [I11].

Yet business process modelling is not only the business people’s responsibility. A director from a public sector organisation described the shift in thinking that “in the past, it was good enough that the one who made the specification got something out of it. Now, the IT must focus on business services and understanding processes.” [I10] This view, shared with the other interviewees, points out that the focus in SOA development should be on understanding business services and processes. According to the interviewees’ experiences, putting focus on services and processes is not easily achieved.

Service identification and finding the right level of service granularity were identified in several case organisations. Service design was recognised by many organisations to be one of the most challenging parts in SOA development process. Several interviewees stated that understanding the essence of business services, what they really are, and then finding a suitable balance between abstraction, modularity and reusability required skills and knowledge beyond what the organisation currently had. The need for technical, domain specific skills was pointed out by several IT side interviewees. This is in line with Shiskov et al. (2006).

SOA teams

Multidisciplinary teams, with members from both IT and business domains, were seen as crucial for SOA development. In all cases it was common that the main role of SOA team was to endorse cooperation between business and IT, and furthermore to act as a centralised steering body for the organisation’s SOA development projects.

“An SOA team was established. It had a few people from business and few people from IT. The business people were working in positions where they were able to gain a holistic view to the customers. They had the capabilities to think what services would be needed.” [I6].

Following Baskerville et al. (2005), SOA development is rather a means than an end. Such thinking as acknowledged governance of SOA compatibility was seen to be a part of the SOA team’s responsibilities:

“There were participants from business units, the CIO’s office and the IT. The purpose was to establish SOA policies and practices. There we looked at things from the exploiter point of view... That way we started to develop our SOA platform and the governance process.” [I5]

Business driven SOA development

It seems that the business driven SOA projects yield better business-IT alignment than other projects. The fundamental difference between the approaches was that the requirements were based on real business cases in business driven projects, while in non-business driven projects the requirements typically originate from a technical viewpoint without clear links to business objectives. Business driven projects were using a process-oriented development approach with intensive use of business
process modelling for capturing functional requirements. Interestingly, this led to a heightened awareness of SOA in general in comparison to IT driven projects.

**Governance**

In the five case organisations, SOA development was directed by defined policies and practices. This improved the quality of the SOA implementation by enabling a better fit with business needs. Yet SOA governance was not initially considered but necessitated some degree of adoption and use. Moreover, SOA applications had to be business critical. This implies that the role of SOA governance is emphasised in an organisation where SOA is used in applications that have a direct impact on the organisation’s business. There SOA is needed as a value adding component that can be further invested in.

The following quotes emphasise SOA governance in the organisational context:

“[Now] we [IT] can challenge business to start thinking things in a new way... Now we get all development needs from the business units to our [IT] headquarters, where we [the business solutions team] analyse them before any development takes place.” [I2]

“[It was hard to set up a governance infrastructure. The job is still ongoing, but we have already taken some parts into production... Governance with a wider scope, when [service] reuse will happen, services will be located and everything running smoothly [that’s still to happen]” [I7]

**Business stakeholder participation**

Having a common forum for business and IT people is an important factor in achieving business-IT alignment through SOA development. In addition to organisation level SOA teams, business domain stakeholder participation in SOA development was also identified as a key factor, having a positive influence on alignment on the project team level. Business domain knowledge and expertise in certain business cases and business requirements were considered beneficial for SOA development projects. Further, although business process modelling was used to specify and illustrate the project’s business requirements, business people had an important role in interpreting and clarifying the models to the IT developers.

“We have business people participating in development projects. Those who can do business process modelling, their role is really important because they can interpret and read these kinds of generic models.” [I6]

Moreover, business stakeholder participation had a positive influence on project execution and outcome as both change management and exception management were jointly handled. For example, an IT architect from finance industry explained the benefits of having business people participating in the development project: "Business has their own developers, who are creating these processes. After they have finished their work, the focus shifts to the IT side. These business people are still taking part in project implementation, and monitoring it as members of a steering group, and, if necessary, giving support to the project team.” [I7]

**Common language**

Mutual understanding between business and IT people is achieved through continuous communication. SOA provided several means for this. First, SOA terminology and concepts formed a basis for IT people to comprehend the current system architecture and articulate it through common language.

“The greatest benefit of SOA has been that our own understanding of how things should be done is improved. Our [IT] vision has been improved, and our competence to communicate with business about IT as a dynamic enabler has been gently enhanced.” [I2]

However, the business people disagreed. They did not consider that SOA had improved their communication capabilities or increased understanding about the systems and architecture, because "every-
body has heard about SOA. To really understand what it means, that’s difficult because it gets so tech-
nical so fast” [13]

These quotations indicate that SOA is mostly a technical concept. To understand SOA terminology
requires an in-depth understanding of the technical concepts. This implies that SOA, as a software
architectural style, does not provide an adequate base for business-IT communication.

Second, using business process modelling was seen to be useful in gaining a common understanding
of the business requirements. An ICT architect described how business process models contributed
business-IT cooperation: “A common language was established when process flow diagrams were
walked through together [with business people]” [14].

Third, business process models were also seen to be useful in validating the implementation of the
development project. The concepts and terminology provided a common language for business and IT.
Process-oriented design approach, combined with SOA development, thus seems to be able to enable
an appropriate environment for business-IT alignment.

However, there was a major problem in communication. Business people and IT people have a ten-
dency to conceive of the concept of service differently: both groups comprehend service from their
own perspectives. “The problem is that in business, we have difficulties to identify what is really a
service. The level of abstraction is either too high or too low. Even in our IT unit they are not able to
identify services properly or to support business units in identifying services. IT people focus too much
on technical details.” [16]

This gap is because of the ambiguity of the terminology. From the business process viewpoint the
service is simply a business service. Similarly, from the technical perspective, the same service is per-
ceived as a technical, application-level service.

Strategy

Strategy level alignment was identified in only two interviews with the IT people, the business people
did not address it at all. This may be because the development projects were still only discrete applica-
tions or aggregates of applications rather than organisation-wide systems. Yet SOA was recognised as
a long-term business plan, valuable for IT development. Knowledge about business strategy was seen
as a useful input for designing reusable services and architectures that would fit future business needs
without major re-implementation efforts. Two architects explained their problems when evaluating
whether to develop reusable services.

“It [potential for service reuse] should be interpreted from the business perspective. To be able to do
it, business should have a future vision that reaches far enough” [14].

Enterprise architecture framework

The utilisation of enterprise architecture framework was considered to improve business-IT alignment.
With the framework, IT’s capability to involve business people in systems development was in-
creased. On the other hand, business people were able to gain a better awareness of the organisation
architecture and of SOA in the system architecture landscape at large. However, enterprise architecture
frameworks originated from IT as “we, in the IT unit, came the conclusion that we need a [language]
for communicating to the business side people what sort of development process we should have in
order to identify architectural requirements from the business point of view” [111]. “Even in practice,
the business people may like the report [that describes what systems are related to certain business
processes]… We get onto the right level in the discussion faster… The idea was that modelling is done
so that IT and business are already aligned there in the models.” [112]

Rapid development tools

Baskerville et al. (2005) argued for agile development methods in SOA development. Our findings
concur, and indicate that the utilisation of such methods for SOA development enabled faster delivery
cycles by increasing development efficiency and enabling agility in adjusting to changes in requirements.

**Process automation**

The process engines technology was scarcely used in the production of SOA services. They were only used for process automation purposes, yielding additional business value for the existing SOA implementation. The biggest challenges in process engine utilisation were claimed to be immature tools and technologies, and interoperability issues.

5 DISCUSSION

In this paper, we explored the potential and realised benefits of SOA utilisation in an organisational context. The interviews with key SOA people provided us with an understanding of the factors essential for SOA development for achieving business objectives. We categorise eleven factors into the 4 themes, discussed in below.

**Organisational culture and human resources.** Resistance to change and siloed organisation structures are creating barriers to SOA utilisation (c.f. Bieberstein et al. 2005). This impedes the achievement of alignment objectives. Yet, by creating common ground and shared understanding, it is possible to overcome cultural barriers. A close relationship with business and IT was found to be an indicator for a positive organisational culture. This positive atmosphere is essential for fruitful business-IT alignment. To achieve it, the importance of continuous communication between IT and business needs to be acknowledged, as this has been one of the main reasons for numerous project failures (e.g. Chen 2007). The solution is to have an SOA team, composing of members from IT and from business domains, and to use shared models and terminologies. Business process modelling and service identification contribute to organisations’ capabilities to develop business process driven SOA projects.

**Processes and methodologies.** Business driven SOA projects seemed to yield better results in terms of business-IT alignment than IT driven projects. We believe this is because business driven SOA projects meet mutual needs better. There business process modelling is used as a means to capture functional requirements. Yet translating the models into application layer artefacts is very difficult. Although it is claimed that business process modelling is suitable for both IT and business users, current modelling notations are still too complex to be efficiently used by average business users. To correctly comprehend the models, business stakeholder participation is needed throughout the SOA projects in order to achieve the business goals. Business stakeholders may contribute to this by interpreting and clarifying the models for the IT developers, participation in the project’s change management and exception control, and ensuring the congruence of SOA implementation in the long run. This emphasises the importance of accommodating SOA development with pre-defined rules and policies, which further improve the quality of the SOA implementation and enable better alignment with business needs in a sustainable manner.

**Communication and artefacts.** The importance of communication for business-IT alignment has been identified (e.g. Reich & Benbasat 2000; Chan et al. 2006; Luftman et al. 1999; Campbell 2005; Sledgianowski and Luftman 2005). Our findings show that earlier findings are also applicable in SOA development. The greatest difference is with SOA terminology, which is comprehensible only to people with a technical background. It can thus be argued that SOA is a technical concept requiring an in-depth understanding. This implies that SOA, as a software architectural style, does not provide a feasible terminology for business-IT communication.

Another significant finding is the tendency to comprehend services differently. This could be explained by the duality of the service concept. Although the concepts and process model terminology offer a basis for a common language between business and IT, they value strategies differently. While business people do not seem to evaluate SOA implementation from the business strategic viewpoint,
IT people perceive it to be valuable as a basis for technical advancement. The business strategy knowledge is thus a useful input for service and architecture design.

**Technology.** Lyytinen and Ren (2008) and Baskerville et al. (2005) discuss the importance of rapid development methods for SOA development. Our findings confirm this. The methods shorten delivery cycles and increase development efficiency, thus enabling a better fit with changing requirements.

These factors influence SOA development from the business-IT alignment viewpoint. Their identification provides us with a better understanding of the complex nature of SOA development. Although many of the findings are already recognised in the literature, the factors provide a complete, practically oriented approach to business-IT alignment. In other words, we try to respond to Chan & Reich’s (2007) call for **practical** studies on how alignment really takes place, what factors frame it, and how managers may benefit from these studies. Researchers may utilise our findings as follows. First, the factors provide a basis for understanding different factors and their causalities regarding successful SOA development. Second, this study complements earlier studies by producing empirical evidence on the phenomena. Third, new research themes are introduced to the landscape of service-oriented systems research e.g. Kontogiannis (2007). Practitioners and managers may benefit from the list of factors by using it as a checklist in their SOA initiatives. By so doing, practitioners may be able to increase the business value obtained from SOA utilisation. Therefore, we suggest that managers should evaluate their current SOA development environment in the light of our results.

The study has several limitations. First, the small number of organisations and interviews is an obvious limitation. Having more organisations would have enabled better generalisations of the results. Second, although we collected the data using key informants at each organisation, a wider set of stakeholders might have provided additional insights to the research theme. Third, only one person per organisation was interviewed in most cases. Taking these weaknesses into consideration, we argue that we have been able to gain an in-depth understanding of the phenomenon in general. We acknowledge that the findings do not provide an exhaustive or widely generalisable explanation for SOA development, but, following the tradition of qualitative and exploratory research, we hope such a defect is pardonable. All in all, we believe that these findings provide a basis to understand the complexity of the phenomenon.

### 6 SUMMARY

In this paper we gained a better understanding of factors influencing SOA development from the business-IT alignment viewpoint. We illustrated how business-related SOA projects are carried out in organisations and what main factors affect their success. We used a qualitative empirical research approach in the study. Data was collected through semi-structured theme interviews and the data analysis was made by applying the grounded theory method. Our findings can be summarised as follows. We identified four themes, including altogether 11 different factors that are related to business-IT alignment of SOA development. The themes and factors are: Organisation culture and human resources: organisational maturity, competences and SOA team. Processes and methodologies: business driven development, governance and business stakeholder participation. Communication and artefacts: common language, strategy and EA framework. Technology: rapid development tools and process automation. These results can be used as a starting point for new research avenues and as guidelines for practitioners.

### ACKNOWLEDGEMENTS

Our sincerest thanks to Tarja Systä and Timo Kokko for their cooperation in this research project.
REFERENCES


IMPACT OF SERVICE-ORIENTED ARCHITECTURES (SOA) ON BUSINESS PROCESS STANDARDIZATION – PROPOSING A RESEARCH MODEL

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0662.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Service oriented architecture (SOA), Business Process Improvement, IT Business Alignment / Value, Standards</td>
</tr>
</tbody>
</table>
IMPACT OF SERVICE-ORIENTED ARCHITECTURES (SOA) ON BUSINESS PROCESS STANDARDIZATION – PROPOSING A RESEARCH MODEL

Beimborn, Daniel, University of Bamberg, Feldkirchenstraße 21, 96052 Bamberg, Germany, beimborn@is-bamberg.de

Joachim, Nils, University of Bamberg, Feldkirchenstraße 21, 96052 Bamberg, Germany, joachim@is-bamberg.de

Muenstermann, Bjoern, University of Bamberg, Feldkirchenstraße 21, 96052 Bamberg, Germany, muenstermann@is-bamberg.de

Abstract

What is the impact of a Service-oriented Architecture (SOA) on the efficiency and effectiveness of business process standardization (BPS)? The contribution of this paper is the development of a research model around the impact of SOA on BPS in terms of achieving fundamental efficiency and flexibility potentials while covering both the business layer and the IT layer of the firm. Drawing on an accepted and widespread enterprise architecture model, we derive propositions that explain why and how SOA’s characteristics help to standardize business processes and how the interplay between SOA and BPS leads to an increased overall business value. Additional moderator arguments, such as the level of service granularity, the centrality of SOA governance, or Business IT alignment, are added to the research model as critical success factors of achieving business value of SOA.

Keywords: Service-oriented Architecture, SOA, business process standardization, BPS, business value, research model.

1 Authors in alphabetical order of names.
1 MOTIVATION

In 2008, the use of Service-oriented Architecture (SOA) in North-American, European and Asian-Pacific companies increased from 44% to 63% (Heffner 2008). This comes along with a paradigm shift from a strongly process-oriented value chain perspective to a service-oriented perspective, viewing the firm as a set of encapsulated but interrelated services (Demirkan and Goul 2006). This shift, together with the corresponding change in the IT architecture, offers various tremendous potentials such as identifying redundant activities in the firm, supporting insourcing vs. outsourcing decisions and subsequent transitions, identifying core competencies, or changing and managing business processes in a more flexible and consistent manner.

In Business Process Management, one particularly important instrument for enabling and achieving efficiency potentials is standardization of business processes (BPS) (Hadfield 2007; Muenstermann et al. 2009a; Muenstermann and Weitzel 2008; Venkatesh 2006). In this paper, we aim at particularly merging the concepts of SOA and BPS and therefore propose a conceptual research model about the impact of Service-oriented Architectures on the standardization of business processes. There exist many works on how SOA should be designed, implemented, and managed, but there is still a huge gap on what the actual benefits of SOA are, or, whether and how SOA delivers business value in terms of Melville et al. (2004). Therefore, we want to tackle this essential question and restrict our focus on the role SOA plays for achieving efficiency potentials from BPS.

In this paper, we will develop a theoretical research model covering the following research questions:

1. What is the impact of Service-oriented Architecture (SOA) on business process standardization (BPS)?
2. How can the interplay of SOA and BPS increase both business process performance and organizational performance?
3. What are the critical success factors required as moderator for an effective interplay of SOA and BPS?

The result of the research paper at hand is a theoretical research model consisting of a set of propositions which describe the interrelationship between SOA and BPS and the role of critical success factors. This conceptual model is intended to form the basis for future empirical research.

Our paper is structured as follows: section 2 provides the theoretical foundation and a concise terminology which allow a mapping of business processes to SOA and introduce the concept of BPS. Subsequently, we develop a set of propositions explaining the interplay of SOA and BPS as well as both the consequences and necessary success factors of the interplay in section 3. The paper concludes by interpreting the results and discussing potential limitations as well as the planned following research steps in section 4.

At this point, we want to clarify about the following restrictions regarding the focus of this paper: (1) Regarding process standardization, there is often the question about the optimal trade-off between standardization and individualization (Hall and Johnson 2009; Lampel and Mintzberg 1996); nevertheless, this is not our research focus. (2) As formulated in the research questions, we focus on the role of an (already implemented) SOA for process standardization. The reverse question whether or how standardized processes facilitate the introduction of an SOA, is not tackled in this paper.

2 THEORETICAL FOUNDATION

The goal of this section is, firstly, the introduction of a concise terminology to describe organizations and their business activities to then, secondly, allow proceeding with an analysis of the interplay of SOA and BPS in the next section.

The first subsection (2.1) introduces "business process standardization" and shows exemplary value drivers. The second and third subsections (2.2 and 2.3) provide the terminology to describe enterprise
architecture and the respective building components. The fourth section (2.4) finally aims at showing
the conceptual relationship between SOA and BPS at the different enterprise architecture layers.

2.1 Business process standardization (BPS)

Standardization of business processes is increasingly receiving attention from academics and
practitioners alike. Since Venkatesh (2006) identified process standardization as one of three broad
future research directions, numerous papers dealing with BPS have been published (Bala and
Venkatesh 2007; Hall and Johnson 2009; Sánchez-Rodríguez et al. 2006; Wüllenweber and Weitzel
2007). On the practitioner side, companies hope to save significant costs by means of BPS (Hadfield
2007). However, in both academia and industry BPS has hardly been treated as discrete and concisely
defined object of research. Even very recent papers claim a lack of research on a concise definition
great attractiveness, academics' and practitioners' work on [business] process standardization is
conspicuously absent". Shaw et al. (2007, p. 92) indeed define BPS as a means to change business
processes from where they are to a standard business process and "focus on a meta process: the
process of changing a process", but completely leave open the question, how a standard business
process should look like and how a given business process at hand can be transformed into a standard
business process.

In contrast to that, Muenstermann and Weitzel (Muenstermann and Weitzel 2008) propose a definition
of BPS as well as an according conceptualization. They define BPS as a two-staged approach:

(a) **internal BPS:**
Unification (homogenization) of multiple existing business process variants to either one
single variant among the existing or to a newly designed target business process, which itself
is composed out of selected tasks of the existing business process variants

(b) **external BPS:**
Alignment respectively adaption of unified/homogenized business process variants to an
externally available reference business process or an externally available best practice
business process

BPS pursues different goals: In the short and medium term, internal BPS focuses on the
unification/homogenization of the business processes considered, foremost aiming at eliminating
inefficiencies and thereby increasing the business process performance along the dimensions business
process time, cost, and quality. In the medium and long term, external BPS aims at aligning/adapting
the previously internally standardized business process variants (to an externally available reference
business process2) to reach and sustainably guarantee cross company validity, reusability, and
adaptability. The closer a company adheres to an externally available reference process, the easier it is
to react e.g., to market and external change in accordance with the way and speed of competing
companies that also have adopted the reference process. Hence, the described alignment/adoption
within the context of external standardization on the one hand increases business process performance
and on the other hand expands organizational flexibility.

---

2 Examples of industry specific reference processes are the "Supply Chain Operation Reference Model (SCOR)" or the
"enhanced Telecom Operations Map (eTOM)". See Kindler and Nuettgens (2005) for an overview of reference processes or
Malone et al. (1999) for how to design reference processes. Such industry specific reference processes have mostly been
adopted by a large amount of companies. Hence, adhering to them can be a necessary condition to allow for easy realizable
process changes and consequently expand organizational flexibility within specific industries.
2.2 Enterprise architecture: enterprise plan, business process model and resource model

We base our terminology to describe enterprise architectures on the "Semantic Object Model" (SOM) following Ferstl and Sinz (Ferstl and Sinz 1995; Ferstl and Sinz 1997). According to the SOM methodology, an enterprise architecture comprises three layers: the enterprise plan, the business process model and the resource model as well as the respective interdependencies and relationships between the layers:

The enterprise plan describes the global tasks of the enterprise, a strategy to cope with that global task as well as surrounding conditions relevant to the solution. The business process model specifies main and service processes as solutions to realize the enterprise plan. The resource model finally provides resources, e.g. personnel, application systems, and machines/equipment necessary to follow the solution path described in the business process model to realize the enterprise plan (Ferstl and Sinz 1997).

Business process models constitute the central layer within the SOM methodology. According to Ferstl and Sinz (1997) they depict a model of the inside view of a business system. They can be interpreted as solution approach for the realization of the enterprise plan. Hence, they combine the enterprise plan – being aware of the business context of the business process at hand – with the resource model and thereby form the basis for the subsequent analysis of the impact of SOA on BPS.

2.3 Components of the business process model and resource model

Figure 1 below displays the building components of the business process model and resource model as well as the respective interdependencies between the building components. The business process model comprises the three building components "processes" (layer 1), "tasks" (layer 2) and "sub-tasks" (layer 3). The resource model (on layer 4) provides "applications systems", "personnel" and "machines/equipment".

According to Davenport a business process is a "a set of logically related tasks performed to achieve a defined business outcome" (Davenport and Short 1990, p. 12) or the "specific ordering of work activities across time and place, with a beginning, an end, and clearly identified inputs and outputs" (Davenport 1993, p. 5). Lee et al refer to business processes – according to the SOM methodology – as connecting business goals and resources realizing them as "a vehicle to build and materialize organizational capabilities" (Lee et al. 2004, p. 645). Ferstl and Sinz (1997) summarize prominent features of business processes as event driven chain of tasks" as well as the "mapping and use of resources.

In summary, on layer 1, we define a business process to be an event-driven, ordered chain of tasks pursuing the goal of reaching a previously defined business goal. Continuing, on layer 2, we define tasks as descriptions of distinguished goals that have to be reached, together with a preceding business event (input) and a resulting successing business event (output). On the next level of detail (layer 3), the SOM specifies sub-tasks, which are defined as business activities realizing parts of a task.

Now we leave the business process model and apply ourselves to the resource model. The resource model comprises the three building components "applications systems", "personnel" and "machines/equipment" and describes how and using which resource the sub-tasks are instantiated.

---

1 As alternatives to the SOM methodology, several practitioners and researchers offer further approaches to describe enterprise architectures. Among them, one example approach is the "Component Business Modeling" (CBM) approach promoted by IBM (Ramchandani and Harwood 2005). To our knowledge all available other approaches are based on an equivalent underlying internal structure and – along the parts relevant for our research approach – can directly be translated into the terminology used in the SOM approach. As a consequence within this article we can derive our terminology from the SOM approach. Moreover, SOM is a sound and completely integrated approach, which offers in contrast to other approaches the advantage to provide a consistent methodology (cf. Ferstl and Sinz 1998). This methodology guides the modeler through the entire process with concrete rules on how to transform strategic business goals into business process models and then into concrete building components (such as interfaces and objects) of application systems.

4 "variant of a business process" defines a business process similar to a given business process except for some minor tasks of the given business process left out.
As shown in Figure 1, the components of the business process model and the resource model depict consecutive layers of the enterprise architecture, each describing the preceding layer on the next level of detail.

**Figure 1.** Overview of building components of business process model and resource model.

### 2.4 Relationships between SOA and enterprise architecture

According to the presented enterprise architecture, SOA is a methodology for designing the application landscape (Siedersleben 2007). According to Erl (2005, p. 54) “SOA can establish an abstraction of business logic and technology that may introduce changes to business process modeling and technical architecture, resulting in a loose coupling between these models. SOA is an evolution of past platforms, preserving successful characteristics of traditional architectures, and bringing with it distinct principles that foster service-orientation in support of a service-oriented enterprise.” Therefore, we distinguish two interrelated paradigms associated with SOA: (1) a technical perspective, which interprets Service-oriented Architecture (SOA) as pure IT paradigm, as well as (2) an enterprise perspective, which models and organizes the entire organization according to a Service-oriented approach, which also includes business processes and workflows in contrast to the first perspective resulting in a so-called Service-oriented Enterprise (SOE) (Demirkan and Goul 2006). If solely applying the technical perspective, SOA serves the design and realization of the resource model (see Figure 1 layer 4).

Combining both paradigms, SOE/A becomes manifest in a three-stages model, consisting of process model, service model, and technology model (Siedersleben 2007). Whereas the process model corresponds to the previous business process modeling, the service model defines not only the services themselves but also the relationship to the business processes. The technology model defines the concrete implementation for realizing the services. Therefore, the service model decouples the business processes from the underlying technological realization in order to avoid that changes in the implementation will affect the business processes. Thus, the technology model enables an efficient implementation of services, whereas the service model facilitates the efficient realization of the business processes (Siedersleben 2007). In ideal circumstances, the upper models define the antecedents for the lower models in order to allow for an Service-oriented Enterprise.

However, there exist also backward links from the lower to the upper models, as organizations still use legacy systems and standard software which do not perfectly support existing business processes. Consequently, organizations make also use of bottom-up approaches for identifying services despite the pure top-down approaches. Whereas top-down approaches analyze existing business processes in order to identify adequate services, bottom-up approaches examine legacy systems and standard software for the identification of services. While organizations using the former approach may face problems realizing the identified services within an existing infrastructure, the latter approach can lead to too fine-grained (more data-centered than business-oriented) services, which are similar to an application programming interface (API). Therefore, a hybrid approach combining both methods is suggested in order to minimize the problems associated with these approaches (Arsanjani 2004).
Linking the model of SOA according to Siedersleben (2007) to the components of an enterprise architecture (see Figure 2) leads to the following relationships:

The process model corresponds to the first layer of the enterprise architecture.

Increasing the level of detail of the process model is the foundation for the service model, which is represented on the second and third layer of the enterprise architecture. The increased level of detail allows to derive the services of the service model directly from the sub-tasks (layer 3), which documents the demanded relationship between services and their business process within the service model (Siedersleben 2007).

Whereas both models mentioned before are independent from the technology used to realize SOA, the technology model describes the technology selected to implement and dispose services (Siedersleben 2007). The technology model originates from the assignment of sub-tasks (layer 3) to services, which are located at the fourth layer of the enterprise architecture.

Consequently, distinguishing between tasks, sub-tasks and application systems or services, which perform the sub-tasks supports a separation being compatible with the three-stages model of SOA. This clear separation of sub-tasks and performing services enables flexibility as well as reuse potentials of SOA (Sinz 2008).

Figure 2. Relationship between the three-stages SOA model (right side) and the components of business process and resource model of the enterprise architecture (left side).

3 SOA AND BUSINESS PROCESS STANDARDIZATION – A RESEARCH MODEL

After showing which layers and components of SOM are influenced by BPS activities (3.1), we develop propositions about the collaboration of SOA and process standardization within the layers of the enterprise architecture (3.2). The last subsection will expand the propositions by moderating success factors with respect to the effective interplay between SOA and process standardization (3.3).

3.1 BPS in the Context of the Enterprise Architecture

On the first layer (processes), internal BPS means reduction of process variants, i.e., different sequences of order and "call" of tasks. Resulting consequences for the second layer (tasks) are (1) the identification of tasks which are redundant, and thus not relevant anymore, and (2) the identification of tasks being elements of different business processes (leading to "reuse" of tasks). Moreover, rather similar tasks can be identified and homogenized as well. Regarding the third layer (sub-tasks), we can argue similarly.

Internal standardization ensures that the same sub-tasks within different tasks are performed from the same service within the resource model.
External BPS focuses on the first and second layer of the enterprise architecture and ensures that the choice, order and interface definition of tasks of processes or process variants correspond to an external reference process or best practice process.

From our point of view, the introduction of a precise terminology for the description of the enterprise architecture and the contained components allows us to point out the area of activity of BPS (internal as well as external).

3.2 Deriving the propositions

With respect to the three characteristic elements of SOA – component orientation (including standardization of interfaces), loose coupling and workflow (2007) – we develop propositions about the influence of SOA (in terms of the extent and the way of an already implemented SOA) on BPS.

As explained earlier, the first step to standardize business processes is to homogenize them internally. SOA supports this internal standardization due to the reuse of the same service (fourth layer of the enterprise architecture) for the same sub-tasks (third layer). However, organizations may face two problems when standardizing their business processes: (1) existing legacy systems may hamper the alteration of business processes, as IT is not flexible enough, or (2) the use of standard software may force an organization to change its processes according to this software (Merrifield et al. 2008; Siedersleben 2007). Due to the service model and the three characteristic elements of SOA, the latter offers the flexibility to support every possible business process. Therefore, IT does not longer hamper the business in their aim to internally standardize processes (process model) as the need to align them with the actual implementation (technology model) is not necessary anymore.

Proposition P1: SOA eases internal BPS due to loose coupling of business processes and of the supporting IT.

Introducing SOA requires a very detailed and precise documentation of business processes. Brahe (2007) shows that the implementation will hold up if tasks are not modeled sufficiently or information flows are not clearly specified. This very detailed process documentation is necessary for internal process standardization activities as well (Ungan 2006).

Proposition P2: Firms that have established an SOA will more easily conduct internal BPS since the necessary process documentation is already available.

Whereas component orientation focus on the design of a single application system, SOA focuses on the design of entire system landscapes (Siedersleben 2007). Therefore, an enterprise wide SOA aims at modularizing services of the technology model (Sanchez 2004) in order to implement them only once for reusing them at different locations, guaranteeing freedom of redundancy and minimizing maintenance costs of the IT landscape (Baskerville et al. 2005; Siedersleben 2007). For example, the use of WSDL (Web Service Description Language) promotes loose coupling and component orientation as it offers clearly defined interfaces for using a business functionality encapsulated in a service. Consistently implementing the three-stages model of SOA down to the technology model leads to a reduction of variants of tasks and sub-tasks within the service model, which in turn promotes the standardization of business processes. This results from a reduction of unplanned redundancies of business functionality within heterogeneous, historically grown system landscapes (Siedersleben 2007).

Proposition P3: Due to the concepts of modularization and reuse, SOA facilitates a reduction of redundancies within the technology model which in turn leads to enhanced BPS.

Moreover, SOA also affects external BPS (on the business process layer). While, from an enterprise perspective, the sequencing of services along a business process is primarily standardized by internal BPS (see Figure 2), the centralization of the process flow control by implementing an Enterprise Service Bus (ESB), which orchestrates the service calls, facilitates the integration of external best practices (external capabilities) – in form of services offered and provided by external providers – into firm-internal business processes. Consequently, the flexibility of the service model allows to optimize the business process model because tasks which are executed by external services can be integrated more easily. Here, adapting the business process model is not necessary; only the relationship between
the service model and the resource model have to be changed (Beimborn et al. 2008). Thus, the following proposition (P4) follows the same argumentation as P1, but applies it to the integration of external capabilities instead of firm-internal standardization.

**Proposition P4:** SOA facilitates external BPS by enabling the integration of firm-external capabilities (in form of services).

By contrast, we can also argue the opposite way: if business processes have been standardized before an SOA is introduced, redundant software can be eliminated and IT costs can be saved (Merrifield et al. 2008). But, the causality is difficult to determine because both BPS and SOA represent requirements for achieving a business value in terms of more efficient business processes and more strategic flexibility (e.g., faster and easier implementation of new business models etc.). Without revising the business processes, we cannot realize technical potentials as well (Davenport 1993; Hammer and Champy 1993). On the other side, BPS without an SOA in place can only be as broad as the used applications and legacy systems allow for. By separating the business process model, the service model, and the resource model, SOA enables step-wise evolutionary process optimization, which is different to the "classical" big bang business process reengineering (BPR) approaches proposed in the early 1990s (Davenport 1993; Hammer and Champy 1993). This flexibility advantage is even more important in times of frequently changing business and regulatory environments which require fast adaptations of the business. Consequently, we can argue that both elements, BPS and SOA, show high synergies when applied together.

**Proposition P5:** There exists a positive interaction effect between SOA and BPS, leading to super-additive benefits in terms of organizational efficiency and increased strategic flexibility.

### 3.3 Moderating effects

After deriving theoretical propositions which explain the interplay between SOA and BPS, additional moderator arguments which serve as critical success factors for achieving business value from SOA and BPS are developed. These are (1) the level of service granularity, (2) the centrality of SOA governance, and (3) Business IT alignment.

The first aspect focuses on one of the most important and most difficult questions of efficient and effective SOA design: determining the optimal degree of service granularity (Acharya et al. 2005). During service design and implementation, there are different options to determine the scope of tasks and sub-tasks and their services on the resource layer. If the "size" of the service is chosen too fine-grained, coordination and orchestration may get very complex (Baskerville et al. 2005). Moreover, the technical performance can be negatively affected since the network has to carry out higher workloads than the services themselves. In this context, several authors speak about "chatty services" (e.g. Woodley and Gagnon 2005). From a BPS perspective, fine-grained services would allow maximal standardization, theoretically, but the administrative complexity would, in many cases, prevent the identification of an existing and appropriate service within the firm’s SOA and thus lead to a redundant implementation. This will threaten the objective of BPS.

**Proposition PM1a:** If the functional scope of the implemented services has been defined too fine-grained, this will hinder internal BPS on the task layer.

By contrast, if the service size is chosen to coarse-grained, reusability potential will decrease. The higher the functional range encapsulated by a service, the more it is fitted to the specific demands of the related task on the business process layer (Baskerville et al. 2005; Brahe 2007). From a BPS perspective, the problem of specifying too large services is that slightly different process requirements from different contexts always result in new service implementations which, in turn, reduce internal standardizability on the task layer.

**Proposition PM1b:** If the functional scope of the implemented services has been defined too coarse-grained, this will hinder internal BPS on the task layer.

Thus, in both cases, there is a threat of repeatedly implementing similar business logic – which is contrary to the objectives of BPS (Woodley and Gagnon 2005).
As a further argument, we introduce the establishment of an SOA governance as moderator. Since introducing SOA often is a firm-wide change of the IT architecture and since it affects all layers of the firm, establishing a central and powerful organizational unit is a critical success factor (Yoon and Carter 2007). Similarly, standardizing business processes across the firm has to be managed by a central and superordinate instance as well (Muenstermann and Weitzel 2008). Moreover, this central unit has to be capable of administering the implemented services in order to ensure a maximum degree of reusability and reuse (Bieberstein et al. 2005). With sometimes more than thousand services implemented by a firm’s SOA (Brahe 2007), this represents a highly complex management task, which not only requires a central SOA governance, but also equipping it with adequate management, power, and documentation tools and mechanisms (Brahe 2007). Without a central SOA governance, SOA cannot be a driver of BPS. Moreover, the positioning of the SOA governance within the firm’s organizational structure will also be an adequate unit for setting up and managing firm-wide BPS activities. Since both SOA and BPS affect the whole enterprise, we propose a joint and central SOA and BPS governance to be best suited in order to capture the potential super-additive benefits from both initiatives. This joint governance can enable and ensure a consequent top-down procedure from business process via task to resource layer when implementing SOA (Vinoski 2005).

Proposition PM2: A centrally established SOA governance has a positive impact on the relationship between SOA and BPS.

Proposition PM3: Merging SOA governance and BPS governance facilitates achieving the goals of SOA and BPS (i.e., P5).

In order to implement an IT architecture which efficiently, effectively, and flexibly supports a firm’s business processes, the organization has to establish and to maintain a sufficient level of Business IT Alignment. Business IT Alignment is a multi-dimensional concept, not only covering the alignment of business and IT strategy, but also the alignment of structures, in terms of processes, capabilities, routines, and resources (Henderson and Venkatraman 1993) – which is also reflected by the SOM introduced above. Further, Reich and Benbasat distinguish an intellectual dimension and a social dimension of alignment. The first covers the congruence of explicated and documented knowledge, such as strategies, project plans etc., while the social dimension captures the "shared mindset" and a common understanding between the managers and staff from business and IT departments (Reich and Benbasat 2000). This social dimension can again appear on the strategic level (i.e., between executives or top management), such as measured by Reich and Benbasat, and on the remaining levels of the firm such as trust and mutual understanding between IT and business managers and employees involved in joint projects or even during daily business (Beimborn et al. 2006; Franke et al. 2005).

All decision problems discussed earlier, such as the optimal level of granularity or the implementation of an adequate SOA governance, ideally incorporating the BPS governance as well, emphasize the critical need for having sufficient alignment between business and IT both on the strategic and on the structural level, and along the intellectual and the social dimension. Without good alignment, which has to be adopted and carried out by each individual involved (social dimension), neither a successful SOA implementation nor effective BPS activities are possible. Exchanging knowledge and establishing a common understanding of "service" and the firm’s business processes is necessary for collaborating on process modularization, deciding about the level of service granularity, and standardizing business processes. Ideally, the firm establishes alignment engineers which show high expertise both on the business side (process layer) and on the IT side (resource layer) (Brahe 2007). This will greatly ensure the mapping between the sub-tasks of the process layer and the service implementations on the resource layer.

Proposition PM4: Strategic and structural Business IT Alignment is critical for successfully implementing an SOA and for its impact on BPS.

Figure 3 summarizes the derived propositions:
Figure 3. Conceptual research model about impact of SOA on BPS.

4 CONCLUSION AND NEXT STEPS

In this paper, we developed a research model consisting of propositions regarding the impact of Service-oriented Architecture (SOA) on business process standardization (BPS). Both concepts represent important factors for increasing a firm’s performance and strategic flexibility. Drawing on the main design concepts of SOA, we propose a positive effect on BPS resulting from loose coupling and modularity which helps to disconnect the business logic of the business processes from their supporting IT. Moreover, modularity and reuse help reducing unwanted redundancies of implemented business logic, which makes BPS easier as the same business logic does not have to be altered at different implementations. These relationships and the potential to integrate external capabilities are moderated by governance design as well as by the level of granularity, which has a direct effect on the potential to reuse services. Additionally, Business IT Alignment is expected to be a critical success factor moderating the direct effects as well as affecting the moderating effects mentioned before. Overall, we expect a combined positive influence of both SOA and BPS together on a firm’s performance and strategic flexibility.

As the next step, we intend to refine the theoretical model developed in this paper by conducting case studies. Based on interviews with managers responsible for SOA governance in large firms, we get insights into the different service design parameters and management mechanisms. Moreover, accompanying the whole process from specifying a certain demand for a new service on business side (process layer) to its final implementation and integration within the firm’s SOA (resource layer) will sharpen our understanding how reusability is determined and how reuse is finally achieved. Thus, the relation from SOA to BPS can be explored from an SOA-driven perspective (i.e., exploratory approach). In other case studies, we will explicitly interview managers involved in firm-wide process standardization activities and will examine – if an SOA is available – whether our propositions can be indicatively justified (i.e. confirmatory approach). After refining our theoretical model, we will transform it into an empirical model, consisting of testable hypotheses and measurable constructs. In this stage, it might be necessary to split single theoretical propositions into several hypotheses which can be quantitatively validated. Moreover, the rich concepts of SOA and BPS need to be unfolded to multi-dimensional constructs which are represented by measurement models each. For example, BPS has been operationalized as a two-dimensional construct, consisting of internal and external BPS, each being represented by a multi-item measurement model, in e.g., (Muenstermann et al. 2009b). Subsequently, we intend to conduct a survey in order to measure our constructs and to validate our hypotheses by using common SEM approaches.

Ultimately, this research project will clarify basic relationships between SOA and BPS and uncover essential success factors for realizing both of them. Thus, we hope to contribute to the research on the business value of flexible and modular IT architectures as well as to managerial reality.
References


THE EMERGENCE OF A NEW FORM OF IS OFFSHORE ENTERPRISE – THE MODERN HETERARCHY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0064.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Global service organisations, Global software development, Global information systems, Offshoring / Outsourcing</td>
</tr>
</tbody>
</table>
THE EMERGENCE OF A NEW FORM OF IS OFFSHORE ENTERPRISE – THE MODERN HETERARCHY

Gannon, Brian, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK, brian.gannon@vesime.com

Wilson, David, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK, dave@cs.bbk.ac.uk

Abstract

This paper describes how IS offshoring organisations are changing in response to increased globalisation of the practice of software development. It posits the emergence of a new form of multi-national enterprise (MNE), described in this paper as a ‘modern heterarchy’, from the construct of the heterarchy originally developed by Gunnar Hedlund in 1986. The paper draws on theoretical antecedents in the discipline of international business studies, and is supported by empirical data gathered from two extended case studies of offshore IS projects. The research uses grounded theory techniques for the collection and analysis of data, and has particular value for IS practitioners in offshore IS MNEs.

Key words

Offshoring; globalisation; heterarchy; grounded theory; multi-national enterprise; international business
1 INTRODUCTION

This research is concerned with a particular form of globalisation: the practice of offshoring Information Systems (IS) development. IS offshoring uses low cost labour in distant countries to provide IS products and services for use in developed economies.

IS offshoring is nowadays deployed extensively and is regarded by many as a mature and cost-effective approach to application development and maintenance (Lewin and Peeters, 2006). Suppliers of offshore IS services have graduated from simple sourcing models - such as providing individuals to do specific tasks - to complex and sophisticated cross-border contractual and resourcing arrangements with their customers (Soota, 2002; Murthy, 2004). New project and organisational structures are required to take account of the dislocation of staff, which in turn demands new ways of managing development activities. Cultural traditions are often disrupted, both for offshore practitioners who come to reside in an onshore location and for the onshore individuals who encounter them (Tsotra and Fitzgerald, 2007).

The rapid development of the IS offshore industry has also resulted in the creation of large multinational enterprises (MNEs). Some of these have originated in industrialised economies – recent manifestations of systems integration (SI) firms such as Accenture which typically provide offshore software development as part of a wider portfolio of ‘multi-shore’ consulting, technology and outsourcing services. Others - so-called ‘pure play’ offshore IS providers such as Wipro - have originated in newly industrialising countries (particularly India) and for the most part export labour and IT-enabled services to developed economies.

Research on offshoring is at a relatively early stage, although it is experiencing more interest in recent years, as evidenced by the special issue of MIS Quarterly (King and Torkzadeh, 2008). By contrast, globalisation, MNEs and international business have long been the focus of research, although to date, few scholars have applied MNE theory to offshore IS MNEs, or to this sector as a particular example of international business. Since MNE theory has proven valuable in explaining how aspects of traditional MNEs function, it may further the understanding of potential changes in the structure and composition of companies that provide offshore IS services, and the forms of distributed multi-national IS organisations that may emerge in the future.

The purpose of this research is therefore to apply an international business perspective to the field of IS offshoring. It asks the question: ‘are there antecedents in the field of international business that can inform the study of IS offshore practice?’ and thereby seeks to add to the body of knowledge on the IS offshore phenomenon. The conclusions from this research will have particular relevance for IS offshore MNEs, whose structure, work practices and perspectives are affected by this phenomenon.

The research comprises a multiple case study approach using grounded theory techniques to analyse respondent interviews. The epistemological approach is firmly interpretive. The researcher shares the view taken by Galliers (1992) that IS comprises computer systems embedded in a social context, and not just hardware and software. Moreover, it is often the social context that gives rise to the most interesting and problematic aspects of IS (Hirschheim and Newman, 1991). This applies particularly to phenomena like IS offshoring, which are mainly concerned with commercial, social and organisational arrangements of IS.

This paper is structured as follows. In section one, the nature and scope of the research is described; section two presents a brief review of the literature on offshoring and international business, including a description of the main organisational constructs used to describe MNE structures. Section three describes the research method, case studies and analytic framework used. Section four presents findings from the research, and is complemented by section five, which presents the conclusions.
2 LITERATURE ON IS OFFSHORING AND MNES

2.1 Literature on IS Offshoring

The perspectives on IS offshoring in the literature present a wide range of opinion, from Farrell (2005) who asserts that offshoring offers huge benefits to both organisations and the economy, to Levy (2005) who presents a more cautious view of the benefits of offshoring. It is possible to summarise the existing body of research on IS offshoring as falling broadly into four categories, determined by the main perspective of the researcher.

First, there is the economic perspective, which highlights such factors as the commercial drivers for offshoring, labour arbitrage opportunities, contractual implications and so on. Examples of this viewpoint include Ang and Straub (1998), Lacity and Willcocks (1995), Farrell (2005), and Venkatesh and Krishna (2004) amongst others. A second point of view is cultural, addressing risks and tensions inherent in distributed software development across political and geographic boundaries. Examples of research that takes this as its primary viewpoint include Carmel and Agarwal (2002), Edwards and Sridhar (2003), David et al (2007) and D’Mello (2005). The organisational perspective focuses on aspects relating to the skills, expertise and organisational structures required when application development is distributed. Research by Doh (2005), Tolentino (2002), Evaristo et al (2005) and Oshri et al (2007) offer examples of this orientation. Finally, the operational viewpoint is dominated by consideration of such elements as the processes, methodologies, tools and infrastructure involved in IS offshoring. Harmsen et al (2007), Gopal et al (2002) and Norbjer et al (1997) all provide examples of research from this point of view.

Gopal et al (2003) and Murthy (2004) look at IS offshoring from the perspective of the IS offshore provider. However, there is relatively little research on IS offshoring as a form of international trade, and on the strategic management, organisation and operation of IS offshore MNEs.

2.2 Literature on organisation of MNEs

International business research on global organisations ranges from early studies of the theory of the MNE (Hymer, 1960; Dunning, 1973; Perlmutter, 1969; Buckley and Casson, 1976; Teece, 1977; Hedlund, 1986; Bartlett and Ghoshal, 1998) to more recent work by Peng (2004), Knight and Cavusgil, (2004) and Fan and Phan (2007). Early research on MNEs tended to view international organisations simply. Buckley and Casson (1976), for example, define the MNE as “an enterprise which owns and controls activities in different countries.” Behrman (1974) identified three types of international organisation - the ‘classic investor”, the “international holding company” and the “multinational enterprise”. Porter (1986) examined firms in the context of their industries, which he categorised as ‘multi-domestic’ and ‘global’. Perlmutter’s (1969) ethnocentric and polycentric description of the MNE, essentially corresponding to centralised and decentralised operating models, offered a radically different perspective. Bartlett & Ghoshal (1998) define ‘multinational’, ‘global’, ‘international’ and ‘transnational’ businesses, characterised by the relative emphasis placed by the organisation on how it configures its assets and capabilities; by the role it assigns to its overseas operations; and by the way in which it exploits its knowledge and intellectual property.

Research from the 1990s onwards, such as Bartlett and Ghoshal’s, has tended to place less emphasis on a hierarchical view of the MNE (headquarters controlling subsidiaries directly) and more frequently takes the view of the MNE as a network of differentiated intra- and inter-firm relationships (Tolentino, 2002). This perspective assumes a distributed labour division among subunits of the MNE arranged in an integrated network configuration (ibid).
3 RESEARCH METHOD, ANALYTICAL FRAMEWORK AND EMPIRICAL DATA

3.1 Research method and design

This research comprises a multiple case study approach (Yin, 2002) using grounded theory techniques to analyse respondent interviews (Glaser and Strauss, 1967). It is inductive rather than hypothetic-deductive, an approach that is suited to grounded theory analysis. Because this approach facilitates a process-based description of change in its organisational context (Orlikowski, 1993), it seems appropriate to the study of the offshore phenomenon, which is heavily process-based and organisationally dependent.

Regarding theory, the approach in this research combines the use of theory as an initial guide to design and data collection and theory as part of an iterative process of data collection. The limited number of cases means that the output is a conceptual framework and related propositions (Eisenhardt, 1989), and does not comprise a nomothetic theory.

In this research, a set of guidelines has been formalised into what is called an analytical framework, which in addition to providing guidelines for analysis also offers a structure for data collection, description and presentation of results, and for allowing the prioritisation and assessment of the relative importance and impact of the results. In this respect, the metaphor of research framework as scaffolding seems particularly appropriate (Walsham, 1995).

The macro-level ‘actor’, or unit of analysis, is the organisation or firm – that is, the organisations that provide IS services, both onshore and offshore. Four primary dimensions of impact of offshoring were selected, comprising cultural, economic, organisational and operational impacts. These categories emerged from the data, and were selected as being the most comprehensive; others categories could equally have been used. The cultural dimension of the analysis covers those impacts of offshoring that have primarily a cultural interpretation or significance. The economic dimension address impacts of offshoring that affect the actors commercially and politically. The organisational dimension is concerned with how offshoring is affecting the structure of IS organisations, and the skills of practitioners onshore and offshore. The operational dimension is concerned with factors that pertain to how offshoring impacts the processes, methodologies and tools of the actors considered.

3.2 Case studies and empirical material

Two offshore IS projects are used to provide a body of data for analysis. These were conducted over a period of 18 months from 2005 to 2007 in separate organisations in the financial services industry: one a UK retail bank, a subsidiary of an international institution, and the other a global insurance broker headquartered in the USA with its European headquarters in the UK. The two companies differ in size, structure and culture. The bank is headquartered in the south-east of the England and has a growing, motivated and stable IT workforce. The insurance broker is located in the City of London, and exhibits some of the organisational volatility and pace of change typical in this environment.

In each instance the primary offshore outsourcing provider was Capgemini, a global systems integrator headquartered in Paris – a typical IS offshore MNE. Other IS firms were involved in more peripheral roles. One project (Project MARS) involved the development of a package-based system to support a new lending product and the other (Project EUROPA) was a custom development of an existing system used to provide retail brokerage for customers across Europe. Both developments were initially of a similar scale – over 10,000 days of development effort – and both used IBM’s Rational Unified Process (RUP) development methodology, although in different technology environments (Java for the bank; Assembler and COBOL for the insurance broker). On both projects offshore developers from Capgemini’s Indian operation were located on site in the clients’ offices in the UK and Belgium for at least part of the time. Thus the projects are philosophically similar (Orlikowski, 1993), drawing on the
same basic application development approach of use cases, separation of process and data, and iterative development phases.

The rationale for selecting two case studies is to allow the continuous comparison of evidence, and to control the conceptual level and scope of the emerging theory (Orlikowski, 1993). At a more basic level, observations made in one organisational context can be compared and contrasted with observations in the second site. The most striking difference between the two companies is in their organisational culture: the bank’s culture is one that has a balanced approach towards risk, and displays a ‘can-do’ attitude to business, reflecting its origin as a successful, marketing-driven start-up. The insurance broker, by comparison, operates on a much more traditional, hierarchically-sensitive basis, typified by extended lead times for decision making and a risk-aware approach to business.

Primary data sources in the form of semi-structured interviews were gathered from client and Capgemini staff directly involved in the selected projects, both onshore and offshore. In all, seventeen interviews were conducted over a period of eight months (October 2006 to May 2007), in various locations in the UK, and by phone with respondents in India. These were supplemented by additional phone calls to validate points of fact. In addition, written data sources—project reports, memos, emails and letters—were collected and analysed. Data collection was aimed at gathering information (loosely) identified by the categories of impact defined in the analytical framework.

Walsham (1995) notes the importance of ‘capturing’ people's interpretations effectively in the course of a normal conversation. To facilitate this, all interviews were recorded and stored electronically as digital files in a ‘wav’ format. These were supplemented by written notes. The data was later analysed manually (line by line analysis of data), and using nVivo software.

4 FINDINGS FROM THE RESEARCH

4.1 The emergence of a new organisational form for IS offshore MNEs

The research indicates that new forms of IS MNE are evolving that are characterised by:

- A blurring in the distinction between onshore and offshore, with allocation of resources from multiple locations rather than solely from offshore locations;
- Greater distribution of risk between onshore and offshore parts of the organisation, and the introduction of commercial models to support this;
- Introduction of methodologies and toolsets that accommodate distributed development across multiple sites (a global development model);
- Rebalancing of skills across locations rather than simply downsizing onshore staff numbers.

This form of organisation is best described by the construct of the heterarchy (Hedlund, 1986) which at the time of writing was viewed as a departure from the more traditional (hierarchical) structures.

4.2 Cultural observations

For both the MARS and EUROPA projects, it is clear that offshoring no longer means pure labour arbitrage. Rather, it is a consequence of an increasingly integrated corporate view of operational efficiency, from the point of view of the supplier (Capgemini) and the customer. This aspect of offshoring is highlighted in the interaction between third party software vendors on project MARS, where there was an inverted relationship between project members in Mumbai and Cheltenham (where one of the project components was developed). Since development was coordinated by Capgemini, the third party in Cheltenham was effectively treated as ‘offshore’ by the Indian development team, some of whom were located in Reading in the UK, and some in Mumbai. Despite the disparity in cost of labour at each
location, this perception seems entirely justified. For example, from the perspective of scale and sophistication, Mumbai is a world city, and Cheltenham a backwater, so it is legitimate to view Cheltenham as ‘offshore’ through this philosophical lens. Further, ‘offshore’ resources in India are just as likely to have a broad world view as their colleagues in Cheltenham.

In effect, as these firms become less location-specific, the distinction between onshore and offshore becomes less relevant, and the commonly accepted definitions of the words, which relate primarily to the physical location of the IS resources, become redundant. This applies generally to the traditional taxonomy in the literature: words like ‘subsidiary’, ‘host country’, ‘home country’, ‘headquarters’, are less relevant in the heterarchical construct, which is peer-to-peer, collaborative and mobile. This represents a fundamental (primarily cultural) shift in the perception of offshoring.

Doh expresses this viewpoint as follows:

“Moreover, as Levy (2005) notes, the development of communications technologies and the requisite mobility of labour have allowed for an accelerated internationalization of production that accords neither with the product life-cycle nor the sequential internationalization perspective. Indeed, some have argued that many firms are now ‘born global’ (Knight and Cavusgil, 2004) and that the notion of sequential internationalization – whether on a country, industry, or firm scale – is outmoded and anachronistic.” (Doh, 2005)

Buckley concurs:

“One issue is whether the firm should be divided into domestic and international divisions (in the era of globalisation now a rather redundant debate…)” (Buckley, 2002)

4.3 Economic observations

Global IS organisations are changing their business models fundamentally: in effect they are adopting a hybrid approach to offshoring that involves the use of joint onshore/offshore teams. There is a rebalancing of the development contact, with each part of the enterprise (onshore, nearshore and offshore) sharing risk and reward. This is different to the current environment, where typically the risk and reward is assumed disproportionately by either the onshore or offshore division.

This rebalancing of the development contract was illustrated in a discussion with the EUROPA delivery director, and concerned the extent of the risk assumed by the offshore division of Capgemini on the project. The traditional model is for Capgemini to use the Indian offshore business as a cost centre with a more sophisticated, risk-bearing onshore front end. The heterarchical model assumes that all development centres are equal, and capable of agreeing their own terms.

The EUROPA project delivery director proposed a different business model that involved sharing the risk – one that was readily accepted by the offshore organisation:

“Yes, India is still run as a cost centre, so the UK or France or the front office country takes all the risk. ..We were trying to resolve this for smaller projects, to transfer risks, and at the time it seemed to me that this was a new way of working but one that they (the Indian colleagues) were absolutely up for. It was an explicit conversation: “Look, guys, we're not going to take the risk on this because this is a fixed price deal - you guys will have to bear it. Are you happy and comfortable with that?” And their view was, well great, finally somebody's taking some notice of us who are actually doing things we want to do.”

This reflects a profound change in the way that offshore phenomenon is impacting IS organisations. It is independent of the pricing mechanism (the fact that the deal is fixed price is irrelevant here: what is important is how the risk is being shared between onshore and offshore components of the same organisation). This change represents a significant maturing of the offshore components and recognition on the part of the onshore part of the organisation that it can no longer dictate the terms of IS engagements with offshore colleagues.
4.4 Operational observations

The modern offshore MNE is adopting new tools and operational processes. However, these tools are basic for the most part, and often include software downloaded from the Internet. New processes - for example to conduct code reviews with developers - and new methodologies - for example to incorporate remote prototyping - are similarly being deployed. However, in the case studies in this research, these changes were basic, and were supported by web-based tools like Instant Messenger. The project manager on MARS describes the process of code review:

“The only way we all kept in contact was Yahoo! And it’s the only way to just maintain contact and you know, sometimes you’d just be cutting and pasting components of code and saying, ‘How do you think this looks like?’ or ‘What do you think?’ and it’s great.”

Similarly, few formal standards were in place, and those that existed were not strictly adhered to. Developers chose their own approach, with little apparent sensitivity around security, as described by a developer on the MARS project:

“…we now have a standard toolset that we’re supposed to use. All our J2EE components used Star Team (for change control) … and the guys in Mumbai just VPN’d in and used it … effectively, we got exemptions to do it our own way, which was maintain it on the client’s site and we would VPN in and do it.”

The rather informal use of methodology and tools on both the MARS and EUROPA projects hides the fact that all of the organisations involved in the development – users, onshore, nearshore and offshore – were closely networked and operated with a good deal of consistency and efficiency. The use of tools like Instant Messenger emphasises the immediacy of the interaction, and was complemented by the adaption of existing methodologies to cope with the new (distributed) environment: on the MARS and EUROPA projects, Capgemini had invested in building a distributed toolset and methodology to account for the fact that the operational impact of offshoring affects all aspects of the development life cycle. This illustrates a resourcefulness and agility within formal frameworks.

4.5 Organisational observations

From an organisational perspective, the skills and capabilities that offshore MNEs are retaining onshore include account management skill and technical skills. The MARS project delivery director describes these skills as:

“…the bits which … require customer intimacy and intimacy with the business users. Those are the bits that, you know, people are almost presuming that they cannot be moved offshore.”

On EUROPA, the account manager identifies the elements of her proposal that were most successful:

“We provided them with a solution that gave them the ability to talk about those additional bits of functionality to a set of people who understood their business pain.”

Her delivery director agrees, and notes that there are some technology skills that will also be needed onshore:

“I think there’ll always be early adopter technologies where people who are familiar with them will be of value locally. I think it’s likely that strategic consultancy, IT strategic consultancy skills…project management skills and business analysis type skills…”

This research shows that a hybrid development approach – a characteristic of the heterarchical development model – is preferred. On project MARS, for example, the bank stipulated that offshore resources be brought onshore to the bank’s premises for the duration of the project, as described by the MARS programme manager:
“They felt that it wasn’t an option to do any of it offshore. It would have been a preference for Capgemini to do components offshore, but they (the bank) weren’t prepared to consider that because they felt that the timescales were too quick… and the risks involved in doing that would be too great. And they felt they didn’t have the maturity as an organisation to do that. So they were absolutely clear they didn’t want anything built offshore.”

5 CONCLUSIONS

5.1 The emergence of the modern heterarchy

While the evidence that emerged from the case data is mostly consistent with previous studies, the use of an international business lens to interpret the se sheds new light on these findings and creates new insights from this research. From observing the offshore MNE on the projects, it is clear that a new organisational form is emerging that exhibits the essential qualities of the heterarchy described by Hedlund (1986).

The construct of the heterarchical organisation describes a networked organisational model. A key strategic difference with traditional organisational paradigms is that the heterarchical company seeks to exploit competitive advantage from any part of the global organisation, and not just from the ‘home’ market. The structural differences are more complex, and posit that the heterarchical company has many centres; that subsidiaries and their management are equally capable of contributing strategic thinking and value; that organisation is collaborative in nature rather than coercive, and generally that each part of the organisation is a reflection of the whole. This latter point implies that every member of a heterarchical organisation is aware of all aspects of the firm’s operation (ibid).

Hedlund presented his model as ‘radical’ and saw it more as a ‘loosely-defined’ or theoretical construct than an actual manifestation of reality. He predicted that such organisations might emerge in the future, possibly in newly developing countries. Writing in 1986, Hedlund used words like ‘novelty’ and ‘radical’, and his goal was to generate debate. He coined the term ‘hypermodern MNC’ to suggest that existing ‘modern’ theories and notions used in international business thinking were inadequate, and used ‘heterarchy’ as an antithesis to hierarchy. (Hedlund’s multi-national corporation (MNC) is synonymous with the multi-national enterprise (MNE)).

Predicting where such companies emerge, Hedlund identifies industries characterised by:

“…the use of many different technologies, high but not maximum global homogeneity of demand, fast rate of technical and market change, non-trivial scale economies (but not necessarily in manufacturing), and absence of strong local barriers to entry”. (Hedlund, 1986)

and notes that IT and biotechnology are obvious (if boring!) candidates. More importantly, he suggests that:

“In terms of geographical and corporate origins, heterarchical MNCs are more likely to evolve from less than gigantic firms, and from contexts with a history of rather autonomous and entrepreneurial subsidiaries. This may give European firms an advantage over US ones. In a larger picture, MNCs from newly modernising nations may stand an even better chance.” (ibid)

The reason the term ‘modern’ is used as a qualifier is because Hedlund’s construct does not describe IS offshoring MNE perfectly. He was writing in 1986 and even in the space of 23 years, much has changed. The pace of globalisation has accelerated, and its nature and profile greatly debated. IS offshoring in 1986 was at an early stage of development, and bears little resemblance to the nature of the phenomenon today. Although instinctively grasping the statelessness of the heterarchical MNE, Hedlund nonetheless defines the strategy of the firm in terms of ‘home’ markets, an irrelevant concept for the modern heterarchy:
“The heterarchical MNC differs from the standard geocentric one both in terms of strategy and in terms of structure. Strategically, the main dividing line is between exploiting competitive advantages derived from a home country base on the one hand, and actively seeking advantages originating in the global spread of the firm on the other.” (Hedlund, 1986)

Further, his notion of heterarchy implies differentiation, similar to the ‘differentiated network’ described by Rugman and Verbeke (2003). The modern heterarchical firm is decidedly undifferentiated, deploying its resources in a manner dictated not exclusively by location (for example, from a ‘centre of excellence’) but by a mix of factors including cost, availability, location, proximity to the client and strategic intent (for example, by the desire to expand a presence in a particular country). A good example of this was provided by Capgemini’s use of Accelerated Development Centres on project EUROPA: resources from France, Holland, India and the UK were deployed to optimise cost and expertise.

Similarly, the rebalancing of project risk on project MARS is giving rise to a new commercial model to accommodate heterarchical operation. It acknowledges that the traditional ‘brokerage’ business models of the western IS providers are changing to a more equitable distributed business model. The flexible approach towards development toolsets and methodologies on both EUROPA and MARS projects typifies modern development techniques. It is moreover entirely consistent with the heterarchical construct to the extent that the development infrastructure (telecommunications, tools, methodologies) can be defined as heterarchical. The Internet is stateless, networked and (mostly) immediate, and the collaborative toolsets that comprise Web 2.0 technologies are collaborative, peer-to-peer and instant. From an organisational perspective, the deployment of varied skills across distributed locations is consistent with the concept of a heterarchical enterprise, which recognises that low-cost offshore development on its own does not necessarily meet client demands; nor does aggressive labour arbitrage on its own represent a wise competitive stance (Hedlund, 1986).

All of these outcomes provide further evidence of the emergence of a heterarchical enterprise, and a move towards a more strategic deployment of offshoring analogous to that illustrated in Carmel and Agarwal’s stage model of offshoring (Carmel and Agarwal, 2002).

5.2 Evolution and maturity of the modern heterarchy

The research indicates that IS organisations will not necessarily find the evolution to becoming a modern heterarchy easy, particularly those organisations that are at an early stage of development and only now coming to understand the implications of a truly global market for IS service provision. This is a difficult transition for most onshore organisations, and there is little information available to guide them:

“The newly integrating nature of this global labor market has strategic and tactical implications for companies and countries alike. Information and insight about it are sparse, however, and executives and policy makers have little of either for making the decisions they face.” (Farrell et al, 2005)

Moreover, there is no definitive model: the modern IS heterarchy is not entirely uniform. The world is not flat, as Friedman (2005) has described it: it is bumpy and uneven, containing all sorts of inequalities, inconsistencies and irregularities, and one size does not fit all. For the MARS and EUROPA projects, for example, the recruitment and resourcing process was novel and problematic, as described by the Capgemini UK account manager:

“...it was difficult because it was a new process. So it was difficult identifying the right skills and getting the handshake between the UK and Mumbai working effectively. ... so we had somebody managing this, more or less full time, for about two weeks, two or three weeks, setting up the process, setting up the documentation around it, so there was clarity around who’d been interviewed...”

This led to delays in the project start date for both projects, something that was complicated by the fact that the public processes to facilitate offshoring were not optimised, and required significant client as well
as multi-shore organisation involvement. For example, the UK’s Home Office was not geared up to accommodate large scale offshoring in the UK, and the Capgemini project manager for MARS had to spend a good deal of time resolving these issues:

“Yeah, there were (difficulties bringing developers to the UK from India) and we had to write letters to the Home Office explaining what the contract was. We had to give them copies of the contract. … Initially, we got the visas for too short a period, and so we had to have people who went offshore. We sent them back to Mumbai, so that they could get visa extensions. So that was quite complicated and costly and disruptive.”

5.3 Theoretical and practical contributions of this research

This research uses powerful explanatory constructs from the related but substantially different discipline of international business and successfully applies it in the field of IS. Specifically, Hedlund’s construct of the heterarchy is adopted - together with related organisational taxonomies from Perlmutter, Porter and Bartlett and Ghoshal - to provide insight on offshoring.

This has the effect not only of validating the imported constructs, but also of illuminating the topic being researched. In this instance, the contribution validates the construct of the heterarchy and shows its continued ability to explain complex aspects of IS offshoring such as how IS offshore MNEs are organising. Such cross-disciplinary borrowing is endorsed by one of the leading researchers in this area:

“In its successful era, international business researchers not only imported concepts and paradigms, they also exported them to neighbouring areas. This does not seem to be occurring at the moment.” (Buckley, 2002)

While the explanatory power of the heterarchy is significant, it does not completely describe the new offshore IS organisations. This research accordingly extends the construct to take account of the elements of offshore MNEs that are new and different to previously researched MNEs. The extended construct – the modern heterarchy - offers a richer view of these new IS organisations, and therefore provides a significant contribution to the wider field of IS studies. It is potentially of interest also in the field of international business studies where the constructs originated.

This study has direct relevance for IS organisations engaged in IS offshoring. If it is accepted that the modern heterarchy is the form of organisation to which many IS MNEs will evolve, then there is much in this and in antecedent research that can help inform this evolutionary continuum. For ‘end user’ organisations, there is enough detail regarding the phenomenon to provide guidance in the deployment of offshoring with both offshore and onshore systems integrators.

5.4 Future research directions

While this research did not evaluate pure play MNEs to the same level of detail, secondary evidence from respondents suggests that they also are becoming modern heterarchies. This is because the pure play strategy is now focused on building solid customer relationships in local markets, while retaining the efficiencies and disciplines that come from centralised control. The onshore IS firms are also changing strategy: to compete against the structured, centrally-driven offshore organisations, these firms are developing development ‘factories’ in offshore and onshore locations that are modelled on the offshore organisations’ ‘global’ strategy. In effect, onshore and offshore IS companies are now indistinguishable in strategic intent, and each has co-opted elements of the other’s strategy.

Further planned research will look to validate this conclusion. It will also extend this reasoning to other dimensions identified in the analytic framework to assess, for example, the impact of global IS offshoring on IS practitioners ‘onshore’ and ‘offshore’.
6 REFERENCES


INTERNAL MARKETS AS A SOURCING OPTION FOR THE DELIVERY OF IS SERVICES: IMPROVING OUTSOURCING AND INSOURCING

Bulchand-Gidumal, Jacques, University of Las Palmas de Gran Canaria, Facultad de CC.EE. y Empresariales, Campus Universitario de Tafira, 35018 Las Palmas, Spain, jbulchand@dede.ulpgc.es

Abstract

The delivery of IS services is very resource consuming in all organizations. In an era in which management optimization is one of the bases for competitive advantage, sourcing options must be judged carefully. Traditionally, the dilemma has been considered in terms of make (insourcing) or buy (outsourcing). But both of these options have shown a certain amount of problems. Although outsourcing has allowed firms to concentrate on their core competencies, it has also shown several issues that prove it is not a panacea. Insourcing, on the other hand, does not allow a great deal of flexibility and we also find that it is every time more frequent that firms do not want their amount of staff to grow. Between both options, the internal mechanism has arisen. Its use entitles the creation of an intermediate society between the organization and the market. This article shows two cases of organizations from different environments in which the sourcing process involved the creation of an intermediate organization to act as the described buffering mechanism. Both processes were related to the IS/ICT function, one in Spain and another one in Italy. The study of the two cases has helped us identify and confirm empirically a number of features achieved by this internal mechanism use.

Keywords: Internal markets, outsourcing, insourcing, IT function.
1 INTRODUCTION

In the last few years, we have seen the importance of information systems and information and communication technologies (hereafter IS/ICT) in organizations steadily increasing. This increase has entitled that organizations have had to deal with the way to deliver these services. Traditionally the alternatives to do this have been two poles of what is now considered to be a continuum but which for some time was considered to be just the two possible existing options: outsourcing (buy) and insourcing (make).

Outsourcing is the process by which external agents perform organization’s activities while insourcing is the process in which the organization performs the functions in-house. We also refer to insourcing when a firm undertakes an outsourcing process and later decides that the services should be performed again by its own personnel and even in the case in which an organization examines the costs and possibilities of outsourcing but finally decides not to use this sourcing option (Hirscheim and Lacity, 2000).

As time has passed, both alternatives have been examined in detail in the IS/ICT sector by different authors, practical cases have been made available, and a certain amount of problems have been found in both of them.

On the one hand, outsourcing has been blamed for making organizations lose control over IS/ICT assets (King and Malhotra, 2000; Lee et al. 2003; Gupta and Gupta, 1992), lose flexibility (Lee et al. 2003; Gupta and Gupta, 1992), have to go through the possibility of threats of opportunism from the supplier (King and Malhotra, 2000), lose IT expertise and corporate memory (King and Malhotra, 2000), lose of qualified personnel (Lee et al. 2003; Gupta and Gupta, 1992) and lose of competitive advantage in information management (Lee et al., 2003) as well as in the innovation capacity (Earl, 1996).

Moreover there are also problems derived from the complexity of breaking a contract in case of dissatisfaction (McFarlan and Nolan, 1995; Lacity et al. 1996), the use of not very qualified personnel by the contractor (Earl, 1996), the possibility of a misalignment between the outsourcing organization and the contractor (Lacity et al. 1996) and even an increment in costs when many organizations outsource as a way to cut down costs (Lacity et al. 1996).

On the other hand, insourcing means not achieving the main reasons that make organizations outsource. This is not been able to control costs, not being able to concentrate on the nuclear capacities of the organization and having a very low level of flexibility over workers.

As we can see, both outsourcing and insourcing of IS/ICT have major drawbacks. However, we must consider that these drawbacks will have a greater weight depending on how important IS/ICT are for the organization. This is, for organizations in which IS/ICT are just a commodity which must be dealt with, choosing the best combination possible of outsourcing and insourcing will probably be enough. Normally we will find that very small organizations completely outsource IS/ICT, and that as they grow they tend to start creating their own IS/ICT services; once they grow over a certain amount of employees (or the IS/ICT needs grow above a certain limit) they try to increase the proportion of outsourcing.

But, what happens when the firm considers IS/ICT to be a source of competitive advantage, a nuclear activity? Due to the problems with insourcing and outsourcing, these organizations have had to find other alternatives to overcome these problems. Between these alternatives we can mention internal markets, strategic alliances, business process outsourcing and selective sourcing. We now define these four alternatives.
Strategic alliances. Two or more organizations jointly develop functions in the search of competitive advantages. For a strategic alliance to function and be successful over time, all the participating organizations must consider that they are obtaining value from the alliance; in other words, that a “win-win” situation occurs (King, 2001). There are three types of strategic alliances (Barney, 2002): non-equity alliances, equity alliances and joint ventures. Non-equity alliances are based on contractual relationships, so firms are all stakeholders, but not necessarily shareholders. In equity alliances and joint ventures an independent company is created.

Internal markets. A company creates its own organization to undertake certain tasks or it allows one or more of its units to act autonomously and to transact with other units of the company. This is, strategic alliances and internal markets are similar mechanisms, just in the first it is two or more firms getting together to create an independent company, while in internal markets it is just one enterprise the one creating the new company.

Business Process Outsourcing (hereafter BPO). BPO consists on handing over control of an IS/ICT based process (for example, human resources management or accounting) to another organization (Rouse and Corbitt, 2004). The firm thus eliminates the need for certain IS/ICT areas since the company who supplies the service will be responsible for the IS/ICT required to provide it.

Selective sourcing. Selective sourcing consists in subcontracting just certain parts of a given organizational function while retaining others in-house (Lacity et al. 1996), this is, to find a mix between what is done outside the organization and what is done inside (King, 2000). It is specially used in the IS/ICT area, due to the complexity of the function and the number of tasks involved.

As we can now see, the two alternatives formerly mentioned in the article are just two points in a continuum (see Figure 1), that starts in insourcing, passes through outsourcing and arrives to BPO. In this continuum, as we move from the starting point to the ending point we find less effort from the company in the management of IS/ICT but also, subsequently, less potential in the form of competitive advantages.

Figure 1. The source continuum

Sourcing alternatives have been widely explored in the literature, quite often in relation to IS/ICT. Probably the theoretical framework most widely used to explore sourcing alternatives, this is, the make-or-buy decisions has been transaction cost economics, derived from the original works of Coase (1937) and Williamson (1979). According to transaction costs, markets would always be used as the organizational form, but every time a market failure appears, a firm (i.e. a bureaucracy) is created as an answer to this failure (Ouchi, 1981). These failures could be due to issues such as bounded rationality, uncertainty, complexity, opportunism or small numbers.

Bounded rationality means that people try to behave rationally but, due to limited knowledge and imperfect cognitive capabilities, this rationality is limited (Greiner and Goodhue, 2005). This bounded rationality causes uncertainty since one side can never be sure how the other side is going to behave in a negotiation. Complexity is to be expected in any relation, but in the IS/ICT area it is increased due to the object of the relation, that are normally not easily to measure (i.e. it is very difficult to estimate.
precisely how long installing a new operating system and migrating all email accounts can take). Small numbers appear when in a relation, and due to the typical evolution of tasks, the process of specialization leads to a point in which it is difficult to find an alternative supplier for a product or service. Opportunism is based on the idea of self-interest and increases even more the uncertainty.

But sometimes, in the presence of any of the above failures, moving from markets to bureaucracies is not the solution, and some other organizational forms are needed (Ouchi, 1981). This happens when tasks are “[...] highly unique, completely integrated, or ambiguous for other reasons [...]” (Ouchi, 1981: 134). As we can see, these attributes are frequently applied to IS/ICT tasks, which gives us a hint that probably in this area, the decision between make (bureaucracy) or buy (market) has a much wider spectrum of possibilities that should be explored. Ouchi himself suggests the clan form: a group based in common values and beliefs, thus allowing minimization of goal incongruence and tolerating high levels of ambiguity in performance evaluation. Greiner and Goodhue (2005) explore open source communities as an intermediate option, these open source communities having some characteristics similar to the clans found by Ouchi. However, clans are not a mechanism easily found, since there are based more in perceptions and values of the individuals than on general parameters usually found in different environments.

Strategic alliances have also been explored frequently, in the IS/ICT sector and outside it. But there seems to be a certain lack in the literature of articles exploring the buffer mechanism that internal markets allow. In this article we are going to focus on this alternative, the internal market strategy, trying to answer the following research questions: what are the advantages and disadvantages of the internal market strategy? When and how should it be used? What issues have to be dealt with in the process?

In order to answer these questions, we have based our research on two practical cases developed in Spain and in Italy. Both these cases were related with the IS/ICT function. The one in Spain involved the use of this solution in a Public University as a way to outsource certain IS/ICT functions while retaining a good level of control over workers and over management of these functions. The one in Italy was that of a SME which decided to create a buffer company to manage the implementation project of an ERP System.

This article is organized as follows. First, we explore the internal market strategy. We then present the two cases that have been studied, followed by a section in which we show the research methodology that was used in each of the two cases. We also show other cases that we know about and that we expect to study in further detail in order to confirm better the results found. Then some common findings to both cases are shown. We believe many of these findings can be generalized to all internal market strategy cases. Last we discuss on the results, we highlight some limitations to the work done up to this point and we end with some conclusions which we hope can be used by managers facing sourcing alternatives.

2 INTERNAL MARKETS

As has already been defined, internal markets is a mechanism in which a company creates its own subsidiary firm to undertake certain tasks; it is also a term used when a unit from the mother company is separated from it and starts operating as an autonomous unit. Usually this new organization or the autonomous unit will provide services not only to the parent company but also to other companies in the market, thus guaranteeing a competitive pricing structure and an appropriate quality of service (King and Malhotra, 2000) since the subsidiary should be almost the same as any other external supplier (Dearden, 1987). In the following we are going to concentrate in the case in which there is a new company and it is not a unit that has been allowed to become independent.

This created company acts as a buffer organization between the mother organization and the market, thus allowing for “[...] increased responsiveness of internal suppliers, better quality with lower cost of internally-supplied services and products, elimination of fluff, de-bureaucratization, de-
monopolization, uniform measures for comparing the performance of various units, and greater opportunity for development of management skills [...]” (King and Malhotra, 2000: 327). This is due to the created company being very close to the mother organization but, at the same time, being independent.

3 THE TWO CASES

3.1 Case 1 – Spanish Public University

The first case was that of a Spanish Public University that decided to create the buffer society as a way to outsource certain IS/ICT functions while retaining a good level of control over workers and over management of these functions. The university has roughly 20,000 students and a total budget for the management of IS/ICT of around 8 M€. It has 45 staff dedicated to IS/ICT management.

As the university faced the need to enlarge the number of staff dedicated to IS/ICT following the natural increase of IS/ICT services needed that has taken place in all organizations, it found that growing the internal structure was very difficult due to tight restrictions in Spanish Public Administrations\(^1\). Outsourcing was also considered but the geographical position of the university (the Canary Islands, more than 2,000 km. from Mainland Spain) and the specificity of tasks to outsource meant there was an important shortage of providers available.

Due to this, the decision to create a buffer organization was taken. This organization was created during the year 2005 and it started providing services in 2006. Its budget for 2007 was about 700,000 Euros with a number of staff of about 25. It provides help desk services as well as some application development services. The objective of the enterprise is to break-even at the end of each financial year.

Rough estimates showed that if the same tasks had been developed in-house the total cost would have been around 1 M€, due to higher salaries that public employees are entitled to. On the other hand, outsourcing would have also meant a figure close to the same 1 M€, since, as was stated previously, enterprise from Mainland Spain willing to provide the services would have charged that extra quantity in accountancy for their benefits and for the extra costs of having to operate a unit in a remote place. This shows that, very roughly, the internal market option has allowed for savings of about 30%. This is not absolutely true since certain overhead costs (creation of the enterprise, costs associated to the heads of IS/ICT in the university who are part of the Board of Directors of the newly created company, etc.) have not been considered.

3.2 Case 2 – Italian SME

The second case was an Italian SME (we will subsequently call it MainFirm) in the iron and steel sector that decided to create a buffer company (that will be called BufferFirm) in order to manage the implementation project of an ERP System. The company was created in 1996 and started providing consulting services immediately. The following year, the implementation project took place and the new firm started providing consulting and technical services as programming, customization and training. The turnover was 200,000 Euros the first year, 700,000 the second year, 1.5 million the third year and 2 million the fourth year. The firm ended its activity in 2001, once the rollout of the ERP was completed and it was stable.

In this case, the bases to create this firm came from the high level of outsourcing that was already being used and the reduced dimension of the IS/ICT department in the main firm. Due to this, the top

\(^1\) In Spain, Public Universities have to follow the same rules that are applied to Public Administration in general.
management decided to create a new consultancy company which was owned by the firm (45% of the shares) and by four experts in business process reengineering, programming and project management (55% of the shares).

The new firm started analyzing the main business processes, in order to identify the critical aspects that had to be taken into account in the ERP selection phase. As result of the analysis, the adoption of a national ERP called Diapason was suggested. This ERP had been developed by a software vendor (which will be named SoftwareFirm) and the recommendation was made due to economic reasons and because Diapason was considered to be the reference in the industry in Italy.

The BufferFirm management decided that the implementation would be done through a gradual rollout per module and per site. This is, rather than buying an ERP implementation project (software and services), partnership was signed with the SoftwareFirm. According to the contract, BufferFirm would become a business partner of SoftwareFirm, which meant that BufferFirm’s personnel were involved in some ERP implementation projects as observers. This would allow a group of consultants in BufferFirm to be able to understand how to run an ERP implementation project from both the technical and from the managerial points of view.

After six months BufferFirm became able to run implementation projects on SoftwareVendor's software. From that moment BufferFirm’s personnel were involved in the MainFirm’s project and in some other projects of SoftwareVendor in which “body rental” agreements were conducted.

3.3 Other cases

Apart from the two mentioned cases which have been studied in detail, the authors also know, and are in the process of studying, a number of cases in different environments which could help confirming the main results that have been found in both the above cases:

- Car dealer in Spain. This firm distributes the cars of one of the biggest worldwide carmakers. Initially they concentrated in just one region in Spain but over time they have developed in South America and Africa. Some time ago, they found the IS/ICT applications provided by the automaker were not enough for their management, so they started developing their own. As time passed, they started selling these applications not only to their own subsidiaries but also to other franchises of the automaker around different part of the world. In order to better deal with the whole process they decided to create an IS/ICT company, that provided IS/ICT services to the mother firm and to other enterprises. Now it is one of the firm’s main sources of competitive advantage.

- Banking sector. There are several cases in different European countries in which Savings Banks have created organizations specifically dedicated to the management of their IS/ICT. Our idea is to explore some of these cases, of which we already have proof in Spain and in Italy.

4 METHODOLOGY

In both cases studied, the key to the research methodology was that one of the authors of this work was an agent in the process. This, obviously, introduces a certain bias into the investigation. This bias has been dealt with by having each of the possible conclusions drawn from each of the cases carefully analyzed by the other author, in order to avoid the mentioned bias. However, even with this is not absolutely possible to state that there is not a certain bias in the whole research, which is going to be further avoided by analyzing other cases.

Four types of techniques or data collections methods where used: participant observation, document retrieval, in-depth interviews and focus groups. Evidence obtained was qualitative as well as quantitative. The time frame considered starts at the creation of the enterprise and ends when the later
had been working for one year. This accounts for a total of 20 months for the case of the Spanish University and 36 months for the Italian SME. In both cases, the analysis took about three months and it started after the enterprise's first year of operation, except for some data that was collected as the process was happening, as explained below.

Participant observation and document retrieval was possible since, as has been stated previously, the authors of this work were part of the process. More than 50 hours were spent over the one year study period directly observing the help desk area in the case of the public University. In the case of the Italian SME, one of the authors was involved in the ERP implementation project and he spent three days per week taking part in the design, modelling and training sessions. Data was transcribed, structured and double checked by the authors to ensure accuracy.

Observation was supplemented with several in-depth interviews that were carried out with participants in the two processes. In total, more than 30 interviews were done. These interviews lasted typically around 75 minutes and used a semistructured questionnaire in which several issues were considered: satisfaction with the process and with the relations between the mother and the created firm, problems found at the beginning, what was expected to happen in the near future, etc.

In case 1, four focus groups were carried out. Two of them took place with the staff from the application development and user support areas of the university while the participants in the other two were their counterparts from the created enterprise. The objective of these focus groups was to analyze with the main actors in the process what their impressions of the process were, what their evaluation of the first year of work and suggestions on improvements for the following years.

Also in case 1, measures of the performance of the services delivered by the firm were obtained through a survey carried out by the evaluation department of the university. In this survey, every service delivered by the university was evaluated: student registration process, IS/ICT services, campus services, transportation, etc.

5 FINDINGS

The study of the two cases has helped us identify and confirm empirically a number of features achieved by this internal mechanism use. First, the internal market allows a very good balance between the best of insourcing and the best of outsourcing. On the one hand, there is a good control over human resource practices and over technical decisions (i.e. which products to specialize in). On the other, a good level of control over the function being sourced is achieved, since it is no more a black box where the managers do not know what is going on, becoming a function that can be analyzed.

Second, in the process of subcontracting to the buffer organization, the mother organization learns how to organize contracts, which means to dimension properly prices, times and requirements as well as even the process of choosing which things to outsource and which not. This learning process can be very useful for a later process of outsourcing to markets directly. This is quite important, since many of the problems found with the outsourcing processes that take place in organizations could be solved by acquiring this experience. So the newly created organization really acts as a buffer between the mother firm and the market, becoming a learning experiment. Outsourcing is complex and, as was seen at the beginning of this paper, there are many possible flaws in the process, so going through and intermediate phase can help organizations to better outsource.

Analyzing the costs involved in the process, on the one hand the cost control and cost reduction that is usually followed by outsourcing processes is also achieved, as well as allowing for a good level of alignment between the objectives of both organizations. In this case, we can concrete that in the Spanish University case a savings of roughly 30% was achieved. But on the other hand it is true that, as will be seen later, the process of creating and launching a new organization entitles a series of costs that must not be undervalued.
Internal markets also ease the resistance to change in the workers of the mother organization, since the process is not a full outsourcing, which is always seen as more fearful, but a subcontracting to an organization which is quite close to the mother organization, and over which this later retains a certain level of control.

Finally, this mechanism allows for the implementation of human resource practices similar to those in the mother organization to take place, reducing resistance to change as well as problems derived from frictions between workers in similar job posts with very different contracts.

But the experience in the two mentioned cases has also shown that this is a very complex process and that there are several issues that have to be carefully dealt with. First of all, and since this kind of mechanism is not used very often, the process has to be carefully explained to all the involved workers. Even doing this, a certain resistance to change and certain fears will probably be found in workers, since they may not fully understand what is going to happen to them, specially in the long run. Basically, they can think the final objective is to pass their job posts to the buffer enterprise. Related to this, workers of the mother enterprise may not fully understand the buffer organization is an independent organization, with its own managers that make their own decisions. This means that influencing human resource practices can be done only at a very high level, not at the base level.

Last, but not least, the buffer organization works best when it also provides service to third party companies (King and Malhotra, 2000), forcing it to compete with market rules to avoid falling in the small number (Ouchi, 1980) trap which would lead towards a bureaucracy again. The dilemma is that it is quite tricky to achieve a good balance between the quantity of effort dedicated to each kind of client, specially since the mother firm will control part of the board of directors and have shares in the created company, thus being able to influence its behaviour.

6 DISCUSSION

When and why should the internal markets mechanism be used? Basically, when the organization (the bureaucracy) and the market fail to deliver a consistent solution to organizations. This is, under certain conditions the traditional make-or-buy decision can provide what organizations need.

If the organization and its stakeholders are comfortable with an increase with the number of human resources that it dedicates to a function (being IS/ICT or any other), insourcing can be used. If the organization considers the task not nuclear, and has a number of service providers in the market around it that can deliver those services, outsourcing should be considered. The latter being done taking into consideration the dangers of outsourcing that have been warned by several authors and that have been explored previously in this article.

Also small enterprises will tend to prefer more classical solutions (i.e. outsourcing, insourcing and selective sourcing) before getting into the management complexity associated with internal markets and with strategic alliances.

But when none of the above happens, when organizations cannot or do not want to increase the number of internal human resources and when outsourcing is not advisable or desirable, new sourcing strategies have to be explored, since the make-or-buy poles do not provide a good enough solution for many organizations. These sourcing alternatives include strategic alliances and internal markets. When the company can find other firms that do not strategically compete with it and that have similar interests, the strategic alliance can be explored. If it wants, or it has, to follow this path on its own, then they will be facing an internal market strategy.

Our experience with two very different cases (different geographical regions, different sectors) show that, although the process of building a new enterprise has a high initial cost, the benefits that are achieved are quite rewarding: good control over the function, organizational learning and cost savings.
7 LIMITATIONS

As was stated previously, the authors of this paper have been part of the processes being studied, which introduces a certain bias into the overall process and conclusions obtained. Although every statement has been double crossed by the authors and, when possible, all documentation has also been cross examined, this does not fully ensure that this bias has been eliminated. The authors expect that further research into other cases will allow removing this bias and better guarantee that conclusions can be generalized to a greater extent.

On the other hand, there has to be a consideration about how the selection of the cases that has leaded us to this paper was done. The selection was done, initially, just on the base of being two similar cases found by authors which came to write this paper together precisely because of having worked in similar cases in the past and having written about them. Due to this the paper is based on two very different cases. One is a public Spanish university, the other one an Italian SME. Two sectors involved, education and industry, more precisely iron and steel sector. The selection is not fully representative and was not done because of any other reason than those exposed above. This introduces some limitations in the extent to which conclusions can be extended.

8 CONCLUSION

We consider that these two cases prove that the use of the internal market mechanism trough buffer organizations allows enterprises to find a very interesting sourcing alternative, especially for their IS/ICT function, by providing a method that brings together the best of insourcing and outsourcing. This is, we have proven in practice that the internal market approach is superior to outsourcing and insourcing, as was theoretically suggested by King and Malhotra (2000).

Can we guarantee that these results can be used in other sectors and in other countries? As was stated previously, in the limitations section of this paper, from the case selection it is obvious that there is no security that under different conditions and in different sectors, the conclusions can be used. But, on the other hand, it is true that common issues that have arisen in such different environments and that coincide in both cases so precisely can be reasonably expected to be found in other cases. However, further research to be developed will allow us to confirm this.

References

Sourcing and Automation Decisions in Financial Value Chains

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0584.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Process Reengineering (BPR), Decision making, Global service organisations, Offshoring / Outsourcing</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
SOURCING AND AUTOMATION DECISIONS
IN FINANCIAL VALUE CHAINS

Henneberger, Matthias, FIM Research Center Finance & Information Management, 
University of Augsburg, Universitaetsstrasse 16, 86159 Augsburg, GER, 
matthias.henneberger@wiwi.uni-augsburg.de

Katzmarzik, Arne, FIM Research Center Finance & Information Management, University of 
Augsburg, Universitaetsstrasse 16, 86159 Augsburg, GER, 
arne.katzmarzik@wiwi.uni-augsburg.de

Müller, Stephan, FIM Research Center Finance & Information Management, University of 
Augsburg, Universitaetsstrasse 16, 86159 Augsburg, GER, mueller.stephan@mytum.de

Pleie, Frans-Matthis, FIM Research Center Finance & Information Management, University 
of Augsburg, Universitaetsstrasse 16, 86159 Augsburg, GER, pleie@in.tum.de

Abstract

As information-based processes are usually independent of the location or even the processor, they 
can be oftentimes either automated or relocated to foreign sites to profit from differences in wages. 
Both strategies bear enormous micro-economic potential in terms of cost savings. However, with the 
main focus on cost reduction, risk due to the uncertain development of effective labor costs or future 
transaction volumes are oftentimes either inadequately considered or neglected. This systematically 
leads to false decisions, in particular since the two strategies – relocation and automation – result in 
different risk profiles. In this paper, we analyze the conditions for automating or relocating parts of 
business processes and propose a decision model that suggests a risk/return efficient allocation to the 
alternatives. In particular, we consider how uncertainties of effective labor costs and transaction 
volumes influence the decision. As shifting tasks to other locations has effects on the workload at the 
original location, we also take into account costs for social effects. The practicability of our approach 
is demonstrated with an example that is based on real data of a major financial services provider.

Keywords: business process sourcing, relocation, automation, social effects, risk, decision model

1 INTRODUCTION

Today, globally acting companies that are integrated into world-wide value chains are constantly in 
demand to make sourcing decisions. This includes relocating business processes entirely or in parts to 
foreign countries to profit from effects such as increased cost efficiency due to wage differences and 
opening up new labor pools. Such relocation is accomplished either by outsourcing the business 
processes to external services providers that are located abroad or by shifting the processes internally 
from one site to another. Relocating business processes, either kept in-house or outsourced, has 
become very attractive in the last years and is facilitated by better communication technologies 
(Agrawal et al. 2003). Consequently, this trend is predicted to continue in the next years, which is 
underscored by Gartner (2007), who predict the worldwide market volume for business process 
outsourcing to grow from $160 bn in 2007 to $235 bn in 2011 with an annual growth rate over 10%.

Another ongoing trend that is facilitated by technological progress is business process automation. As 
pressure for efficiency increases, firms have to evaluate whether and to what extent business processes
should be automated. Most often a mixture of both – automation and relocation – seems to be most reasonable. Gartner (2008a), for instance, conceives a combination of automation and outsourcing as a way to reach the fullest extent of efficiency and cost reduction and predicts process automation to rise from 10% in 2007 to 55% in 2017 of all processes outsourced (Gartner 2008b). Oftentimes, process automation has become a true alternative to relocation. For example, simple tasks in the financial services domain such as entering paper orders in an order processing system could either be accomplished cost effectively by staff members in low-wage countries or those tasks could be (at least partially) automated. Applying, for instance, optical character recognition (OCR) technologies, the paper orders could be automatically scanned and processed. Therefore, for making sourcing decisions, an integrated decision support model considering relocation and automation is crucial.

Relocation as well as automation bear risk and recent studies show that saving expectations were often not met (Lacity et al. 2000). Two effects are especially important when assessing the risk and the expected return of both alternatives. First, uncertainties concerning effective labor costs, wage developments in foreign countries, and future transaction volumes have to be considered. If, for instance, the risk of rising wage levels is not taken into account, cost savings intended by relocation cannot be realized. False sourcing decisions are the consequence, leading to higher costs than before (Rouse et al. 2004). Similarly, neglecting uncertainties may result in false technology investments for automation as discussed for instance in Benaroch (2001).

Second, social effects that may arise when tasks are relocated or automated have to be considered carefully. For example, relocating the entering tasks outlined in the example above to low-wage countries may have important negative effects (including reputation problems) if at the same time employees are laid off at the original site. While those effects may be unavoidable in bad times, an interesting question is how enterprises can support growth in good times by exploiting the benefits of relocating or automating without provoking negative social effects.

The objective of this paper is to analyze the economic effects of relocation and automation in such a setting (i.e. from an enterprise’s point of view) and to contribute a normative decision model proposing an optimal allocation of (parts of) business processes in a risk/return efficient way. Thus, we primarily contribute to sourcing theory by providing an integrated view over the alternatives: retention at the origin site, relocation to a new site and automation. We consider the different cost characteristics of the alternatives taking into account also possible negative effects of cutting jobs and analyze how these alternatives relate to each other. We examine the effects of uncertainty of effective labor costs and transactions volumes in a model via simulation. The applicability of our model is shown by studying a real world case with data from a large financial services provider (FSP).

The remainder of this text is organized as follows: In section 2, we give an overview of related literature. Additionally, this section lays the conceptual fundment for this paper (including the definition of important terms). In section 3, the decision model is presented, analyzed and illustrated by an operationalization. Finally, we provide some managerial implications and limitations of the model (section 4) and conclude in section 5 by giving an outlook to further research.

2 LITERATURE REVIEW

Sourcing decisions can be separated into the organizational and the regional dimension which are to a certain extent independent of each other. With respect to the organizational dimension, outsourcing is defined as the procurement of services from sources that are external to the organization (Lankford et al. 1999). With respect to the regional dimension, near- and offshoring refers to relocating jobs to foreign countries without distinguishing whether the provider is external or affiliated with the firm (Levy 2005). In this text, we generally speak of relocation (standing both for near- and offshoring) and focus on the regional dimension since the cost reduction potential of relocating to low-wage countries can be realized both with in-house and outsourcing engagements. Furthermore, we consider automation as an alternative action in our decision problem. Automation refers to replacing human work by computers or machines (Bainbridge 1983). In the following, we will discuss the literature on
relocation and automation in order to elaborate cost structures and risk factors, before we provide an overview of decision models proposed by other authors for sourcing or automation decisions.

In literature, numerous articles on the drivers and criteria for relocation decisions have been published (Farrell 2005, Quélin et al. 2003). Although a number of motives for relocating the execution of business processes such as access to a larger pool of human capital, improved position in global markets, concentration on core business activities or more flexibility in reacting to market changes have been mentioned, the main motive is wage arbitrage due to lower human resource (HR) costs (Quélin et al. 2003). This is especially true for countries with high wages such as the US, the UK or Germany because relocation yields enormous cost reduction potential due to large differences in HR costs in comparison to low-wage countries such as India, where wages are about 15 percent of the US HR costs (neoIT 2006). However, relying only on HR costs may lead to false decisions. Further costs as well as different productivity levels have to be taken into account, too. The so-called loaded costs per employee consist of costs for HR, benefits, space as well as overheads. In general, the loaded costs only show a difference of about 25-35% (Everest 2005). The effective labor costs additionally consider that productivity levels may differ from site to site resulting in another mitigation of the cost arbitrage (Criscuolo et al. 2005). The effective labor costs represent the cost efficiency in conducting a process for each site. Relocating business processes (or parts of them) to a remote location usually causes transaction costs such as management or communication costs.

Relocating business processes bears risk. In the worst case, underestimating that risk may result in higher costs than expected or may even make back sourcing inevitable, i.e. re-integrating outsourced or relocated tasks to the origin site (Rouse et al. 2004). Due to global sourcing, risk influence factors like cultural differences, environment, communication, financial markets, technology, intellectual property and law have to be considered for the decision (AT Kearney 2007, Beck et al. 2008, Winkler et al. 2006). As the volume of business processes, which are sourced to low-wage countries, and following the demand at the labor market in low-wage countries is increasing, estimating future wage levels is conjunct with growing uncertainty (Vestring et al. 2005), i.e. especially the effective labor costs are risky. At this point, it is necessary to clarify our perception of risk: In accordance with financial theory (e.g. Copeland and Weston 1988), we understand risk as deviation of an expected value. Integrating risk in our decision model consequently means to take into account the uncertainty of input parameters by considering those deviations (in the form of a probability distribution).

Costs for business process automation are primarily affected by process-specific soft- and hardware such as acquisition costs or license fees (Alpar 1992). Because of the complexity of some processes, it is not always possible to automate the entire process or it may result in uneconomical high expenses for the latest technologies to rebuild the most complex process steps (Nikolaidou et al. 2001). Thus, by rising degree of automation, the costs increase exponentially. Business process automation bears risk, too. Due to initial investments, particularly uncertain transaction volumes are troublesome because resulting uncertain earnings could lead to reduced revenue and upfront investments may not be amortized (Benaroch 2001). Therefore, one has to evaluate the effects of unstable future transaction volume, which in general cannot be forecasted exactly due to incomplete information.

Relocation and automation of business processes may lead to a reduction of workload at the origin location resulting in costs for reassigning employees to other tasks (including e.g. training costs) or even in layoffs including severance payments (Lee 1997). Furthermore, such situations may result in lower productivities due to decreasing employee motivation (Brockner 1988). Thus, the effects on employees should be considered. In the sequel, we consider severance payments, costs for reassignments as well as trainings and possible effects on productivities as costs for social effects.

In the literature, most articles dealing with decision support on sourcing or automation are qualitative approaches (see e.g. Levina et al. 2008, Rouse et al. 2004 for sourcing and Gebauer et al. 1999, Stohr et al. 1997 for automation). There are only few quantitative approaches in related research areas:

- Yang et al. (2007) identify influence factors for sourcing and propose a decision model using the analytic hierarchy process method. Beimborn (2007) analyzes cooperative sourcing with different
independent actors. His analytical models and simulation approaches are based on game and agent theory. Consequently his work and the work of Yang et al. (2007) differ from our approach both in methodologies and in the research questions covered.

- Vom Brocke et al. (2004) and vom Brocke (2007) apply investment accounting methods to support sourcing decisions on business processes. They distinguish three levels of evaluation: the operational level, the budgeting level and the corporate level. At the operational level, in-payments and out-payments are directly related to process design and sourcing decisions. In contrast to our model, their approach requires detailed modeling of the processes and allows decisions for individual process steps in a multi-period model. We are, however, interested in the general relationship between sourcing and automation (and the different characteristics with respect to uncertainty) and therefore apply a one-period model.

- In the classical area of location theory, quantitative decision models can be found applying decision criteria similar to ours. For instance, Hanink (1985) employs a mean-variance approach for finding factory locations considering returns and risks. Similarly in IT sourcing theory Zimmermann et al. (2008) present a method to allocate software development projects efficiently over sites using Markowitz’s portfolio theory. In both cases, research questions are different to our contribution as automation is not considered.

- Finally, in the area of business process automation, Wei et al. (1998) propose a quantification of the optimal automation degree considering task load and process complexity. Additionally, Sheridan et al. (2000) propose a method to quantify the expected value of the gain of either human execution or automation. Both approaches neither propose a risk-return-integrated model nor do they consider sourcing or social effects.

Summarizing – to the best of our knowledge – there is no publication that takes an integrative quantitative approach combining sourcing and automation. Thus, we want to fill this research gap.

## 3 A MODEL FOR SOURCING AND AUTOMATION DECISIONS

Our leading questions in the following are: Which degree of relocation and automation is optimal for a business process considering the specific cost structures elaborated in section 2? How do costs for social effects influence the decision? How does a combination of two major risk factors – namely the uncertainty of effective labor costs both at the origin site as well as on the new site and the uncertainty of the future transaction volume – affect the optimal relocation and automation degree? To answer these questions, we introduce a quantitative decision model that is developed in two steps. We introduce a basic model in subsection 3.1. A first analysis provides an in-depth understanding of the underlying optimization problem (under certainty). In subsection 3.2, uncertainty is included into the model, then. The section concludes with an operationalization including a simulation approach.

### 3.1 Basic Decision Model

**Notations and Assumptions**

In our one-period model, we consider a FSP that currently conducts a business process at its origin site (in general a high-wage country such as the US or Germany). For each transaction processed, the FSP receives a fixed income \( E \). The FSP plans to reengineer and to re-evaluate the sourcing strategy for the business process. The business process (or a part of it) can be carried out at the origin site, it can be relocated to a new low-wage site or it can be automated. More precisely, we assume that the overall work required to conduct the process can be performed with any possible combination of the substitutive alternatives (retention, relocation and automation). To model this, we introduce the decision variables \( \omega, \kappa \) and \( \lambda \) \((\omega, \lambda, \kappa \geq 0)\) where \( \omega \) represents the degree of work conducted at the origin site (retention), \( \lambda \) represents the degree of relocation and \( \kappa \) represents the degree of automation. The variables \( \omega, \kappa \) and \( \lambda \) are normalized to 1 \((\omega + \lambda + \kappa = 1)\), i.e. their sum represents the complete workload of the business process.
The cost structures of the alternatives differ significantly: “Manual” work at the origin or at the new site causes costs that are assumed to be variable and proportional to the amount of work. For the new site additionally transaction costs $T$ arise due to international coordination. Automation, in contrast, causes a fixed upfront investment depending on the specific degree of automation chosen. These costs are assumed to grow exponentially by $\gamma$ ($\gamma>1$) with the degree of automation and with a maximal amount of money $A$ for full automation.

At present, the FSP processes an amount of transactions $V_0$ at its origin site. In the following period, the FSP has to conduct a number of transactions $V_l$. As outlined in the introduction, we are especially interested in situations where the transaction volume is expected to grow, i.e. $V_l>V_0$. If – due to relocation and/or automation – the new transaction volume conducted at the origin site falls below $V_0$, costs for social effects have to be considered, which are assumed to increase proportionally depending on the reduction in volume. The estimated costs for a complete shutdown of the origin site are represented by a parameter $S$.

Model Description

To reduce writing overhead in the following, we introduce the index $n$ to denote the different sites, work can be conducted at. The origin site ($n=1$) and the new low wage site ($n=2$) are characterized by their productivity $P_n$ and their loaded costs $L_n$. To determine the effective labor costs of the business process at a site, we calculate the ratio of loaded costs and productivity as shown in equation 1:

$$L_n = \frac{L_n}{P_n}$$

The overall effective labor costs are $\omega \cdot V_l \cdot L_1$ for the origin site and $\lambda \cdot V_l \cdot (L_2 + T)$ for the new site, then$^1$. Depending on the degree of automation $\kappa$, the maximal expenses for automation $A$ and the exponent $\gamma$, automation costs are calculated by $\kappa^\gamma \cdot A$. Costs for social effects only arise, if the share of volume retained at the origin site falls below the original volume because in this case the workforce on the origin site would be larger than required. We model this applying the maximum function, i.e. $\max\left(\frac{V_0}{V_l}, \omega, 0\right) \cdot S$. Finally, substituting $\omega$ by $1-\kappa-\lambda$ and expressing the return only depending on $\kappa$ and $\lambda$ leads to the following objective function:

$$R(\kappa, \lambda) = V_l \cdot E - \left( (1-\kappa-\lambda) \cdot V_l \cdot L_1 + \lambda \cdot V_l \cdot (L_2 + T) + \kappa^\gamma \cdot A + \max\left(\frac{V_0}{V_l}, (1-\kappa-\lambda), 0\right) \cdot S \right) \to \max$$

s.t.: $\lambda \geq 0; \kappa \geq 0; \lambda + \kappa \leq 1$

The optimization problem can be solved by differentiating between the two cases: no costs for social effects occur and costs for social effects occur. Comparing both cases finally delivers the solution to the optimization problem. A general (and not very surprising) finding is that with a prohibitive high cost factor for social effects $S$ and in case costs for relocation are smaller than costs for retention, the optimal automation degree and the optimal relocation degree are chosen so that costs for social effects just do not occur. Additionally, it can be shown that for the case that no costs for social effects occur, it is economically reasonable to automate process steps until the marginal costs of automation reach the marginal costs of human conduction (relocation or retention). Interestingly, an automation degree of $\kappa=0$ is not a feasible solution, i.e. under the given assumptions it is always reasonable to automate at least a marginal share of the process. Since the costs of relocation increase proportionally with $\lambda$ and the costs of retention increase proportional with $\omega(=1-\kappa-\lambda)$, the relationship between relocation and retention is rather simple: If the sum of effective labor costs at the new site and transaction costs is higher than the effective labor costs at the origin site ($L_2+T > L_1$), then the maximal possible volume

---

$^1$ We are aware that there are differences between inhouse-relocation and outsourcing. Nevertheless, production and transaction costs come up for both types, but with a different extent. Therefore, the model is designed to be applicable for both types. The specifications of inhouse-relocation and outsourcing must be set by parameterization.

---

Proceedings ECIS 2009
Having understood the general structure of the model, we can now extend the model and analyze the effects of uncertainty due to unpredictable effects that may occur during the planning horizon.

### 3.2 Analyzing the Effects of Uncertainty on the Decision

We consider two important risk factors, uncertainty of the effective labor costs and of the transaction volume, in the model. For this integrated approach, we introduce Assumption 1:

**Assumption 1:** To model risk, the effective labor costs \( L_n \) and the transaction volume \( V_1 \) are assumed to be normally distributed \((N(\mu, \sigma))\) random variables. Risk is hereby understood as possible negative or positive deviation from the given expected value \( E(L_n) = \mu_{L_n}, E(V_1) = \mu_{V_1} \) and is quantified by the given standard deviation \( \sigma(L_n) = \sigma_{L_n}, \sigma(V_1) = \sigma_{V_1} \), respectively.

As the input parameters \( L_n \) and \( V_1 \) for the return function are uncertain now, its result \( R \) is uncertain, too. This uncertainty has to be considered by the decision maker, who needs to make the decision based on a random variable \( R \) (instead of return \( R \) as defined in equation 2). This means the specific risk/return position of each allocation (i.e., each possible combination of \( \kappa \) and \( \lambda \)) has to be evaluated. Implicitly or explicitly this is done by a utility function that expresses the decision maker’s attitude towards risk and enables calculating the utility of an uncertain result \( R \).

**Assumption 2:** There exists a utility-function \( u(R) \), which assigns a specific utility to every random variable \( R \) and which is compatible with the Bernoulli-principle. We assume a risk averse decision maker that maximizes utility.

Based on a utility function, one can derive a preference function that integrates return and risk and can be used as a decision rule. A classical \( \mu\sigma\)-rule that is compatible – under the constraints of normally distributed random variables (Assumption 1) and a risk averse decision maker – with the Bernoulli-principle (Assumption 2) is given by the following equation (cp. Freund 1956):

\[
\Phi_R(\kappa, \lambda) = \mu_R - \frac{1}{2} \cdot \alpha \cdot \sigma_R^2 \rightarrow max!
\]

This function calculates a preference \( \Phi_R \) based on the expected value of the return \( \mu_R = E(R) \), the risk in realizing the return \( \sigma_R = \sigma(R) \), quantified by the standard deviation, and the decision maker’s attitude towards risk, which is represented by the Arrow-Pratt parameter \( \alpha \). The parameter \( \alpha \) is a positive value, which indicates risk aversion (Arrow 1971). The expected return \( \mu_R \) and the variance \( \sigma_R^2 \) can be derived from equation 2 and from the distribution parameters of the input parameters \( L_n \) and \( V_1 \) \((\mu_{L_n}, \sigma_{L_n}, \mu_{V_1}, \sigma_{V_1})\), respectively). However, in contrast to equation 2, determining a closed analytical solution to the optimization problem is not possible, since more than one risk factor is considered at a time (several stochastic random variables are multiplied). Thus, we employ a simulation approach, which is introduced in the following subsection.

### 3.3 Operationalization and Simulation

Now, we demonstrate how to apply the model in a typical decision situation. The following operationalization is based on a real business case of a FSP. Names as well as all identifying details are omitted and the business case data have been anonymized for reasons of confidentiality.

The A-BANK, a large European FSP, operates all over the globe and plans to reengineer and reorganize several of its business processes. For reasons of clearness, we consider only the process for handling high-value payments and checks for business clients which consists of 23 process steps. Besides cost reduction, the reason for reorganizing the process is motivated by an expansion and the number of transactions is expected to grow from presently 1,210,000 to 2,200,000 in the next period.
The following alternatives have been identified by A-BANK for the process under consideration:

- **Site 1 – Germany:** Up to this point, A-BANK processed their checks at its origin site characterized by high wages, but also high productivity. If less transaction volume than at the beginning is handled at this site, it would result in layoffs and/or partial reassignment of staff to new tasks requiring additional training. Both may result in increased costs and resistance of staff. A complete shutdown of the process at the site is estimated to result in costs for social effects of about €2.5 mn.

- **Site 2 – India:** A-BANK runs a large shared service center in India. This offshore site has become very popular due to the large and mature talent pool. A disadvantage lies in the cultural differences to Western countries. The labor costs are significantly lower than in Germany, but with a considerable large rate of increase. In addition to labor costs, transaction costs due to international distribution arise. Thus, the total costs for each transaction consist of labor and transaction costs.

- **Process Automation:** A-BANK is able to implement several process steps with a new software system and can install OCR systems for check handling. There are different levels that differ in the scope of the automation and in the recognition quality. The latter influences the manual effort required in subsequent process steps. For instance, low recognition quality means that (manual) controlling steps are still required in the process. By rising degree of automation, the costs are expected to increase exponentially.

By means of the model introduced before an optimal allocation of the process steps to the different alternatives should be found. A major challenge was to parameterize the model. For the automation costs, the following procedure turned out to be reasonable: For each process step, automation costs were estimated. Similar process steps were clustered by complexity and costs. The analysis revealed that some process steps were easy and inexpensive to automate whereas others were extremely complex or difficult resulting in very high automation costs. Five different types could be differentiated with estimated costs for the automation as shown in Table 1 and Figure 1. Based on the cost estimations, we approximated automation costs to increase exponentially with an exponent $\gamma$ of 2.0 (cp. Figure 1). The maximal automation costs (parameter $A$) were determined by the sum of the costs for automating all steps, which were estimated to €21 mn.

<table>
<thead>
<tr>
<th>Type</th>
<th>Quantity</th>
<th>Estimated workload share concerning the process [%]</th>
<th>Estimated automation costs [€ mn]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>per step</td>
<td>per type</td>
<td>per step</td>
</tr>
<tr>
<td>I</td>
<td>9</td>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>II</td>
<td>7</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>III</td>
<td>5</td>
<td>4.8</td>
<td>24</td>
</tr>
<tr>
<td>IV</td>
<td>1</td>
<td>5.5</td>
<td>5.5</td>
</tr>
<tr>
<td>V</td>
<td>1</td>
<td>6.5</td>
<td>6.5</td>
</tr>
<tr>
<td>Σ</td>
<td>23</td>
<td>-</td>
<td>100</td>
</tr>
</tbody>
</table>

*Table 1. Breakdown of the automation costs*

In order to determine the risk aversion parameter $\alpha$, we compared the cost estimations and risk surcharges of the business case with the estimated variances of our approach. Based on the ratio of investments to estimated risk, we found a constant risk aversion parameter $\alpha$ with the value of 0.0001 to be reasonable. Figures about transaction volumes were available from the A-BANK. Further parameters such as productivity parameters, loaded costs and costs of social effects could be estimated.
based on experience and based on public (e.g. neoIT 2006) as well as A-BANK internal reports. Table 2 summarizes the input parameters for the simulation:

<table>
<thead>
<tr>
<th>Site</th>
<th>( V_0 )</th>
<th>( V_1 )</th>
<th>( E )</th>
<th>( P_a )</th>
<th>( LC_a )</th>
<th>( T )</th>
<th>( S )</th>
<th>( A )</th>
<th>( \gamma )</th>
<th>( \sigma_{V_1} )</th>
<th>( \sigma_{L_1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 1</td>
<td>1,210,000</td>
<td>2,200,000</td>
<td>€4.8</td>
<td>1.1</td>
<td>€5.0</td>
<td>-</td>
<td>€2.5</td>
<td>€21 mn</td>
<td>2</td>
<td>0-20%</td>
<td>3%</td>
</tr>
<tr>
<td>Site 2</td>
<td></td>
<td></td>
<td></td>
<td>0.7</td>
<td>€1.9</td>
<td>€0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Model parameters for the simulation

The simulation works as follows: For each possible \( \kappa-\lambda \)-combination (we iterate over the feasible interval \([0,1]\) for both \( \kappa \) and \( \lambda \) in 0.01 steps), we draw values for the normally distributed random variables \( \bar{L}_n \) and \( \bar{V}_1 \) using the Monte-Carlo method (according to the distribution parameters \( \sigma_{V_1} \) and \( \sigma_{L_1} \)). With these values, we calculate the return \( \bar{R} \) according to the return function 2. This procedure is repeated 10,000 times (denoted by \( i \)). With the resulting 10,000 values for \( \bar{R} \), we calculate the mean \( \mu_R \) and the standard deviation \( \sigma_R \):

4) \[ \mu_R = \frac{1}{i} \cdot \sum_{i=1}^{10000} \bar{R}_i \] \[ \sigma_R = \sqrt{\frac{1}{i-1} \sum_{i=1}^{10000} (\bar{R}_i - \mu_R)^2} \]

These values are inserted into the preference function (equation 3). Repeating this procedure for each possible \( \kappa-\lambda \)-combination, we get a value \( \Phi_R \) for each \( \kappa-\lambda \)-combination. Consequently, the \( \kappa-\lambda \)-combination with the highest value of \( \Phi_R \) is the optimal allocation for the current constellation of input parameters.

To analyze the influence of increasing risk to the optimal \( \kappa \) and \( \lambda \), we repeat this procedure with different values for \( \sigma_{V_1} \) and \( \sigma_{L_1} \), increasing both \( \sigma_{V_1} \) and \( \sigma_{L_1} \) from 0% to 20% by 0.5% and determine the optimal \( \kappa-\lambda \)-combination for each \( \sigma_{V_1}-\sigma_{L_1} \)-combination as described above. \( \sigma_{L_1} \) always has the value of 3% during the whole process.

Figure 2 depicts the change of the optimal automation degree (\( \kappa \)) (left plot) and relocation degree (\( \lambda \)) (right plot) for increasing both \( \sigma_{V_1} \) and \( \sigma_{L_1} \). For illustration purposes, we extracted profile cuts I, II, III and IV (cp. Figure 2) and show them in Figure 3 and Figure 4:

Figure 2. Graphical representation of the results of the simulation

---

2 We are aware that the return \( R \) as specified in equation 2 may not be normally distributed in any case (because we multiply several normally distributed random variables and a maximum function). However, an analysis of the distributions in the simulation approach has shown that \( R \) is at least approximately normally distributed, i.e. the differences cause little and acceptable deviations in the results.
In the following, we analyze the influences of the different risk factors and of costs for social effects. Growing uncertainty of the effective labor costs at the new site $\sigma_{t_2}$, e.g. caused by unpredictable wage developments or productivity deviations, in tendency fosters automation as the less risky alternative ($\kappa$ rises, $\lambda$ decreases; cp. Profile I). This effect can be observed up to the point where automation becomes too expensive due to increasing process step complexity. After exceeding this point, the automation degree stagnates (cp. Profiles I and II). This means, automation can be seen as a substitute for relocation as long as it is not surpassed in attractiveness by retention. In contrast, rising uncertainty of the transaction volume leads to a higher share of relocation and less automation. This is due to high upfront costs for automating a process and in turn probably lower earnings resulting from uncertain transaction volumes. Interestingly, when the automation degree $\kappa$ equals zero, relocation becomes less benefiting, too (cp. Profile III and IV), which means, uncertainty of the transaction volume affects both, automation and relocation.

Considering the effects of both factors together, we find that in the area of low uncertainty (low $\sigma_{t_2}$ and $\sigma_{t_1}$), the relocation degree $\lambda$ reacts very sensitive on deviations (cp. Profile III). This effect, which is especially strong for $\sigma_{t_1}$, is caused by the linear behavior of relocation costs, i.e. in the area of low uncertainty, little changes may have enormous effects on the optimal allocation resulting in completely different relocation degrees $\lambda$. Furthermore, changing uncertainty of the transaction volume $\sigma_{t_1}$ in general has a stronger impact on the optimal solution than changing uncertainty of effective labor costs $\sigma_{t_2}$ which is because the first affects both $\kappa$ and $\lambda$ (cp. Profiles III and IV). Interestingly, the results in this particular case additionally show that automation is still a viable alternative even with volume risk of 7% (cp. Profile II).
As there is also uncertainty of effective labor costs at the origin site ($\sigma_{\lambda_1} = 3\%$ here), retention is no longer a certain alternative to relocation and automation. The effects are the following: With rising $\sigma_{\lambda_2}$, the automation degree $\kappa$ is steadily rising with a small slope (cp. Profile II). The relocation degree $\lambda$ is affected stronger since retention and relocation have similar cost characteristics. Thus, rising unattractiveness of the first – in terms of risk – directly means higher attractiveness of the latter. This effect is illustrated in Profiles I and III with high values for $\lambda$ in particular in the range of low $\sigma_{\lambda_2}$.

While – even high – costs for social effects are accepted when there is no or only low uncertainty (cp. Profiles I and III), increasing uncertainty fosters the effects of these costs since cost reduction potential of relocation and automation becomes uncertain and the costs for social effects also support this trend (cp. Profiles II and IV). In general, we can state that avoiding costs for social effects, for instance by moderate automation and relocation, is a preferable option as the sum of the optimal allocation $\lambda$ and $\kappa$ is running on or nearby the social cost border (cp. Profiles I and II).

4 DISCUSSION AND LIMITATIONS

Summarizing, uncertainty significantly influences optimal relocation and automation decisions on business processes. One source of uncertainty in today's global economy are effective labor costs. In the near past, for instance, wage levels in countries such as India increased faster than expected, which turned out to be a major problem also for the A-BANK. Our model includes those risks suggesting automation as an interesting alternative, if at the same time the uncertainty of the transaction volume is not too high. In this case, the decision maker should first choose process steps that are easy and inexpensive to automate. A further result is that with increasing risk, taking costs for social effects should be avoided. To put it in other words, the decision maker should have complete information or should at least be able to estimate parameters with the required precision before relocating work to a foreign country and cutting jobs at the origin site. If effective labor costs and the transaction volume are uncertain, the automation and relocation potential should be exploited only up to the point where costs for social effects would arise. If uncertainty is exceptionally high, even staying significantly below this threshold is a favorable strategy. In the light of the current economic crisis with high uncertainty of the transaction volume, the results suggest that postponing costly automation projects and keeping the work at the high wage and less risky site can be the preferred alternative.

The introduced model is based on a set of assumptions. These limitations at the same time define extension potential. First, we assumed infinitely divisible processes. In reality processes may not always be cut as proposed by the model. Still, most processes consist of several steps and can nearly be allocated as optimized. Thus, feasible discrete allocations may lie near the theoretical optimum, also delivering a good economic solution. Our model is especially applicable, when process steps can be more or less independently sourced or automated since interdependencies are not explicitly included in the calculations. Second, we proposed a one-period model which was reasonable to understand the fundamental relationships between sourcing and automation under uncertainty. Extending our model to multiple periods (in terms of a cash-flow oriented perspective) is an interesting direction for further research. Investment accounting methods (as discussed in the literature section) may be a starting point for this extension. Third, specific characteristics of in-house relocation versus (external) offshoring could be considered. Fourth, we assumed normal distribution of the random variables – which is a common assumption in literature (e.g. Hanink (1985)). Nevertheless, our approach could be extended by including empirical estimations of probability distributions which could then be applied in an enhanced model. Last but not least, not all of the results can be generalized and further simulations are required to validate our findings.
5 CONCLUSION

Though relocating and automating business processes bear much economic potential, there are only few quantitative decision support models. In particular, there is a lack of research considering both alternatives in an integrated approach. Thus, in this paper we propose a risk/return based approach to calculate optimal degrees of relocation and automation for individual business processes considering the specific cost structures and risk profiles. In comparison to previous research, this approach comes up with several enhancements: First, we simultaneously consider relocation and automation. Second, such decisions also have effects on the original site since jobs may be abolished. Hence, we consider such social effects by quantifying negative outcomes of disestablishing jobs. Finally, we analyze the influence of uncertainty. Applying the model, the results provide a recommendation for automation or site selection for business process steps. Employing data of a business process of a major FSP, the applicability of this approach is illustrated.

The model provided evidence for several facts. First, automation becomes unattractive, if information of the future transaction volume is incomplete. But, automation is a viable alternative, if the effective labor costs cannot be predicted precisely. Second, relocating process steps is affected negatively both by uncertainty of transaction volume and uncertainty of effective labor costs. Third, taking costs for social effects is only sensible, if the decision maker has sufficient information. This is an interesting finding since it suggests that social effects should be assessed carefully. Finally, not considering uncertainties may lead to false decisions. Summarizing, the analysis of the proposed model for supporting sourcing and automation decisions not only revealed interesting insights, it can also form the foundation for further research.

References


Proceedings ECIS 2009
TOWARDS OPERATIONAL RISK-AWARE INFORMATION SYSTEMS: A CRITICAL REALIST PERSPECTIVE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0438.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Risk, Critical Realism, IS Philosophy, Decision making</td>
</tr>
</tbody>
</table>
TOWARDS OPERATIONAL RISK-AWARE INFORMATION SYSTEMS: A CRITICAL REALIST PERSPECTIVE

Kristian Rotaru, Monash University, Department of Accounting and Finance,
900 Dandenong Rd, Caulfield East, Victoria, 3145 Australia,
kristian.rotaru@buseco.monash.edu.au

Carla Wilkin, Monash University, Department of Accounting and Finance,
900 Dandenong Rd, Caulfield East, Victoria, 3145 Australia,
carla.wilkin@buseco.monash.edu.au

Andrzej Ceglowski, Monash University, Department of Accounting and Finance,
900 Dandenong Rd, Caulfield East, Victoria, 3145 Australia,
andrzej.ceglowski@buseco.monash.edu.au

Leonid Churilov, University of Melbourne, Melbourne, Victoria, Australia,
leonid.churilov@gmail.com

Abstract

Following calls to enhance risk-sensitivity of second generation Operational Risk-Aware Information Systems (ORISs), this paper aims to address the lack of ontological/epistemological grounding for the concept of Operational Risk (OR). Herein, OR is regarded both as a property of a real system and as a representational phenomenon forming part of the core of ORIS in line with Weber’s (2003) view of the core of IS. The paper explores how the ontological/epistemological position of the Critical Realist philosophy of science assists in the Requirements Definition of ORISs by providing an ontology-driven representation of the heterogeneous nature of OR. The retroductive mode of logical inference enabled by Critical Realism supports the discovery of OR causal mechanisms when the historical data about operational loss events is limited. The ontological/epistemological position suggested in the paper contributes to better understanding and representation of OR, informs OR assessment in conditions of a constantly changing socio-economical environment, and so assists in the Requirements Definition of ORISs.

Keywords: Operational Risk, Critical Realism, Causality, Operational Risk-Aware Information System
1. INTRODUCTION

Last decade saw a rise in frequency and magnitude of Operational Risk (OR) loss events in most areas of business including supply chain management (Lewis 2003) and banking and finance (Chernobai et al. 2007; Moosa 2008). In response the introduction of regulatory requirements and recommendations (i.e. Sarbanes-Oxley Act 2002; Basel Committee for Banking Supervision (BCBS) 2006) established the role of Operational Risk Management (ORM) as a strategic function that assures sustainability of organisations (Scandizzo 2007; Chernobai et al., 2007; Lewis 2003). Today, in ORM theory and practice, Information Systems (ISs) are regarded as both a source of ORs (BCBS 2006; Wolf 2005) and a tool for managing ORs (Flores et al. 2006; Chartis Research Ltd 2008). This research focuses on the second aspect, wherein ISs are seen as a tool for managing ORs, and explores their ability to represent the phenomenon of OR. Reports on risk technology (cf Chartis Research Ltd 2008) detail a forthcoming transition from first generation “compliance ‘tick-box’ approaches”, to a new generation of Operational Risk-Aware Information Systems (ORISs), although explicit support in the research for this is still scarce (Flores et al. 2006).

Following Weber (2003 p.vii), this research views “representation... [as] the essence of all information systems”, which implies the ability of “track[ing] states of and state changes in other systems” (ibid). By “other systems” Weber (2003 p. vii) means the real world systems that exhibit the patterns of events and behaviours according to certain rules. These also inform the development of “artificial information systems” (ibid). Leveraging off Weber’s work, we assume that an IS can represent OR as an intrinsic component of the real world socio-technical system. To facilitate this understanding we explore the nature of OR as an object of research inquiry (ontological positioning). Further, given research in the IS field has insistently called for “examin[ing]g and clarif[y]ing the underlying set of ontological and epistemological assumptions that underpin every research activity” (Kanellis and Papadopoulos 2009, p.1; Becker and Niehaves 2007), we also look at how the knowledge about this phenomenon is built (epistemological positioning). To date research on OR has not explicitly explored the ontological assumptions that support understanding about the heterogeneous nature of OR (Holmes 2003), nor has it provided the epistemological grounding to enable the knowledge acquired through different OR methodological approaches to be addressed. Consequently, this questions the ability of ORISs to represent, in a non-contradictory and rigorous way, the phenomenon of OR as part of the corresponding real system.

The aim of this paper, therefore, is to address the research gaps mentioned above by exploring how the ontological/epistemological position of the Critical Realist (CR) philosophy of science assists in: 1) representing the phenomenon of operational risk as part of the real system reflected by an Operational Risk-Aware Information System; and 2) more informed decisions in the Requirements Definition stage when designing an Operational Risk-Aware Information System. The novelty and original contribution of this paper concerns the non-contradictory ontological/epistemological support provided by multimethod operational risk assessment that combines the data-driven and scenario-based approaches as an integral component of Operational Risk-Aware Information Systems.

The reminder of this paper is organised as follows. The next section addresses the issues of OR representation and assessment, and suggests a list of desirable properties that aim to overcome these issues. In Section Three the suitability of the dominant philosophical paradigms adopted in both the IS and Risk Analysis literature are critically assessed by examining their ability to meet these desirable properties. This is followed by discussion of how the alternative ontological/epistemological position brought by the philosophy of science of CR leads to a better set of desirable properties. Section Four illustrates the implications of the CR ontological/epistemological support in the development of ORISs. This is followed in Section Five by an illustration of how CR can inform the OR assessment process of an ORIS model. Section Six provides conclusions and offers suggestions of future research directions.
2. REPRESENTING OR AT THE REQUIREMENTS DEFINITION STAGE OF DEVELOPING ORIS

Despite IS being predominantly treated as a source rather than a tool for ORM, the body of literature that explores ORM in IS provides the reference points that are necessary to understand the desirable properties of ORIS. Based on extensive analysis of both the ORM and IS literature, Wolf (2005, p. 104) suggests a list of criteria to assess the potential of existing approaches to manage IS risks: 1. “comprehensiveness” as sufficient coverage of all stages of the risk management process; 2. “life cycle coverage” as the ability to “address the risks of the system life cycle”; and 3. “IS risk sensitivity” as the “coverage of IS risk categories” that are reflected in the BCBS (2006, p.144) definition of OR: “the risk of loss resulting from inadequate or failed internal processes, people and systems, or from external events [emphasis added]”. In this paper we adopt this widely accepted definition of OR. This paper’s aim to explore issues of OR representation at the Requirements Definition stage of the ORIS development life cycle limits the first two criteria introduced by Wolf (2005). In this paper we are focusing on: 1) the stages of the risk management process that are directly related to the representation of OR, i.e.: risk identification, risk analysis and risk evaluation; and 2) the Requirements Definition stage of the ORIS development life cycle. The third criterion introduced by Wolf (2005) is addressed fully in this paper.

The Revised Basel II Framework (BCBS 2006) provides the major source of methodological guidelines and regulatory requirements for organizations that plan to transfer to an ORIS. It primarily provides regulatory requirements and high-level managerial guidelines for the banking sector. Hence, more than 50% of organisations that adopt an ORIS are banking institutions that have to meet Basel II regulatory requirements as part of their ORM strategy (Chartis Research Ltd 2008). This justifies exploration of the methodological implications suggested by the BCBS (2006) as part of the Requirements Definition stage of ORISs. The Revised Basel II Framework (BCBS 2006) suggests three methods for calculating operational risk capital charges. The Advanced Measurement Approach (AMA) is considered to be the most risk sensitive and sophisticated of these approaches (Flores et al. 2006). The organisations that adopt this OR measurement approach considerably minimise their capital allocation charges. A recent global market survey of 318 banks and insurance companies reveals that 52% of respondents aim to adopt the AMA approach as their ORM system by 2011 (Chartis Research Ltd 2008).

In order to be able to adopt the AMA approach the organisation is required to meet a set of “Standards” set up by the BCBS (2006, pp.142-152): “General Standards”, “Qualitative Standards” and “Quantitative Standards”. The first two sets of standards refer to the organisation of the ORM function in a bank, whereas “Quantitative Standards” refer to the methodological requirements for OR representation and assessment. Column 1 of Table 1 represents a full list of the “Detailed Criteria”, (a)-(f) (BCBS 2006, pp. 151-152) that have been interpreted in a more concise form by Scandizzo (2007, p. 56). The “Criteria” are part of the “Quantitative Standards” for the AMA introduced by the BCBS (2006). Column 2 represents the research evidence that supports the “Detailed Criteria” contained in column 1. Based on the outputs of columns 1 and 2, column 3 provides the set of desirable properties that aim to address the issues of OR representation and assessment. Finally, column 4 suggests a list of references that support the arguments and desirable properties in columns 2 and 3.

The first desirable property is related to the need to provide greater crossover between different OR assessment methods including: a) different degrees of subjectivity in relation to the object of the research; and b) different types of reasoning that underlie different methods of OR assessment (for example, subjective approaches that allow educated guesses such as scenario analysis vs. data-driven approaches based on statistical analysis of operational loss data) (Mestchian et al., 2005; Scandizzo 2007). Interestingly, both regulators (BCBS 2006) and practitioners (Breden 2007, p.193; Mestchian et al. 2005) call for the crossover between loss data and scenario-based approaches. The second desirable property is driven by the BCBS (2006) requirement to cover both expected and unexpected losses as
well as the major drivers that underpin ORs. This property calls for understanding of the complex nature of causality that underlies operational loss events (Lewis 2003; Holmes 2003).

Finally, the third desirable property relates to adopting a mode of logical inference that would aggregate the results of different OR assessment methods. This implies the ability to assess and represent the heterogeneous nature of OR by adopting, if necessary, a range OR assessment methods based on opposed modes of logical inference. This categorisation of OR assessments methods has been used by Scandizzo (2007) who distinguished between proactive (ex ante) and reactive (ex post) methods for OR assessment. The adoption of this desirable property would assure “soundness” (BCBS 2006, p.150) and rigour when weighting and aggregating the results of different OR assessment methods.

<table>
<thead>
<tr>
<th>Basel II “Detailed Criteria” for AMA adoption</th>
<th>OR research evidence</th>
<th>Desirable Properties</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>(e) key features include four fundamental elements of OR assessment: internal data, external data, scenario analysis and factors reflecting the business environment and internal control systems</td>
<td>1. Calls to overcome incommensurability of the OR assessment methods based on opposed directions of research reasoning and different value assumptions: reactive (ex post) inherently objective approaches vs. proactive (ex ante) inherently subjective approaches.</td>
<td>1. A non-contradictory multimethod OR assessment</td>
<td>Scandizzo (2005, 2007); Moosa (2008); Mestchian et al. (2005); Wolf (2005); Mestchian et al. (2005); Breden (2007); Holmes (2003); Jarrow (2008).</td>
</tr>
<tr>
<td>(a) consistency with the Basel definition and categorisation</td>
<td>2. The Basel II definition of OR is “causal-based” and breaks the concept down into four major categories of OR drivers (Chernobai et al. 2007, p.17). The requirement to include unexpected losses as part of OR assessment emphasizes the importance of understanding causal mechanisms based on the interplay of the major OR drivers.</td>
<td>2. Understanding of the complex nature of causal mechanisms that underlie OR losses, both expected and unexpected</td>
<td>Scandizzo (2005, 2007); Holmes (2003); Mestchian et al. (2005); Marais (2005); Chernobai et al. (2007); Cooke and Rohleder (2005); Lewis (2003); Wolf (2005); Diegardt (2007).</td>
</tr>
<tr>
<td>(b) regulatory capital includes expected and unexpected losses</td>
<td>3. The need to assess OR, which is heterogeneous in nature, calls for the relationship between the data-driven and scenario-based approaches to OR assessment to be understood. This assures a consistent basis for weighting and aggregating the assessment results that cover the different components of OR.</td>
<td>3. Adoption of a mode of logical inference that assures consistency and rigour when aggregating the results of multimethod OR assessment</td>
<td>Scandizzo (2005, 2007); Moosa (2008); Mestchian et al. (2005); Breden (2007); Wolf (2005); Cooke and Rohleder (2005); Chernobai et al. (2007).</td>
</tr>
<tr>
<td>(c) sufficient granularity so the major drivers of OR can be captured</td>
<td>4. The Basel II definition of OR is “causal-based” and breaks the concept down into four major categories of OR drivers (Chernobai et al. 2007, p.17). The requirement to include unexpected losses as part of OR assessment emphasizes the importance of understanding causal mechanisms based on the interplay of the major OR drivers.</td>
<td>4. Understanding of the complex nature of causal mechanisms that underlie OR losses, both expected and unexpected</td>
<td>Scandizzo (2005, 2007); Holmes (2003); Mestchian et al. (2005); Marais (2005); Chernobai et al. (2007); Cooke and Rohleder (2005); Lewis (2003); Wolf (2005); Diegardt (2007).</td>
</tr>
<tr>
<td>(d) regulatory capital which is the sum of the various risk estimates (full correlation assumption)</td>
<td>5. The need to assess OR, which is heterogeneous in nature, calls for the relationship between the data-driven and scenario-based approaches to OR assessment to be understood. This assures a consistent basis for weighting and aggregating the assessment results that cover the different components of OR.</td>
<td>5. Adoption of a mode of logical inference that assures consistency and rigour when aggregating the results of multimethod OR assessment</td>
<td>Scandizzo (2005, 2007); Moosa (2008); Mestchian et al. (2005); Breden (2007); Wolf (2005); Cooke and Rohleder (2005); Chernobai et al. (2007).</td>
</tr>
<tr>
<td>(f) a sound approach for weighing the four fundamental elements of OR assessment introduced in the ‘detailed criterion (e)’</td>
<td>6. The need to assess OR, which is heterogeneous in nature, calls for the relationship between the data-driven and scenario-based approaches to OR assessment to be understood. This assures a consistent basis for weighting and aggregating the assessment results that cover the different components of OR.</td>
<td>6. Adoption of a mode of logical inference that assures consistency and rigour when aggregating the results of multimethod OR assessment</td>
<td>Scandizzo (2005, 2007); Moosa (2008); Mestchian et al. (2005); Breden (2007); Wolf (2005); Cooke and Rohleder (2005); Chernobai et al. (2007).</td>
</tr>
</tbody>
</table>

Table 1. Desirable properties for OR representation and assessment

The next section discusses how the two ontological/epistemological positions that dominate the IS and general risk literature support these desirable properties.

### 3. THE RELEVANCE OF DOMINANT PHILOSOPHICAL PARADIGMS IN SUPPORTING OR REPRESENTATION

This section explores the extent to which the ontological/epistemological positions, which are part of the dominant philosophical paradigms adopted in the context of IS and Risk Analysis research, support the set of desirable properties for OR assessment suggested in Section 2.
The OR literature (Moosa 2008; Flores et al. 2007) raised concern about the lack of methodological assistance provided by the BCBS (2006) to meet Basel II requirements. Specifically these concerns were related to the “Detailed Criteria” for AMA adoption. This questions the effectiveness of the resulting capital allocation models (Jarrow 2008). Contributing to this concern, we note that the description of the nature of OR provided by the Revised Basel II Framework (BCBS 2006) does not go beyond specifying a set of possible “major drivers of operational risk” (ibid, p.151). The same is true when it comes to specifying the ways knowledge about OR can be built (cf Detailed Criterion (e) of Quantitative Standards). Herein, we argue that the lack of ontological (which refers to the nature of the real world phenomena) and epistemological (which refers to the nature of knowledge about the real world phenomena) foundations in representing OR questions the “soundness” of the results of OR assessment required by the BCBS (2006, p.150). Consequently, a consistent ontological and epistemological position is required to support the properties throughout the Requirements Definition stage of the ORISs development life cycle.

The ontological and epistemological status of OR has been taken for granted in OR research. Moreover, research on ontology and epistemology of OR has been almost exclusively undertaken within the research domain of Risk Analysis (cf Rosa 1998; Klinke and Renn 2002), a multidisciplinary approach to assess risks related to complex socio-technical systems. Thus, we demonstrate the lack of explicit ontological/epistemological grounding for OR in Figure 1 by using a dashed line. This dashed line represents the integrated ORIS and subsequent ontological/epistemological position, which is in line with the aim of this paper.

Unsurprisingly the dominant philosophical paradigms discussed in the Risk Analysis literature (cf Rosa 1998; Klinke and Renn 2002) correspond to the paradigms widely discussed in the context of IS research (Becker and Niehaves 2007; Lee and Baskerville 2003; Walsham 2006; Kanellis and Papadopoulos 2009): positivism and interpretivism/constructivism (herein, the elements of the interpretivism/constructivism tandem are distinguished according to Becker and Niehaves (2007)). In assessing positivist and interpretivist/constructivist philosophical paradigms this research focuses on two major aspects of Kuhn’s (1970) definition of a paradigm; namely, ontology (that refers to) and epistemology (that refers to the nature of knowledge about the real world).

![Figure 1. Ontological/epistemological support for representing the OR phenomenon as part of the Requirements Definition phase of the ORIS development life cycle](image_url)

### 3.1 Addressing Desirable Property 1: “A non-contradictory multimethod OR assessment”

The positivist ontology supports the view that reality exists separately from the human mind. Epistemologically, positivists claim that the empirical knowledge accumulated through perceptual experience can provide objective knowledge about the real world (Becker and Niehaves 2007). Grounded in empirical methods, positivism provides limited justification for OR multimethod research because the value-free observation of the objective reality is considered to be the only valid approach to acquire knowledge about the object of the research. Thus, it excludes values and normative...
judgements from the OR assessment process or at least undermine their importance. The epistemological position adopted by positivism is not able to address the role of the social values in the knowledge building process. This is a serious limitation for a realistic OR assessment as the critical analysis of social values is the essential component of such OR assessment methods such as scenario analysis (Scandizzo 2005, 2007; Marais 2005; Cooke and Rohleder 2005).

The ability of constructivist/interpretivist paradigm to support this desirable property is also limited by its ontological position, associated with “‘internal’ or ‘subjective’” realism. It represents reality as an intersubjective or a personal construction (Walsham 1995 in Smith 2006). By equalising “subjective meaning” and “objective reality” (Lee and Baskerville 2003, p. 230), interpretivists/constructivists claim that “risk assessments constitute mental constructions that can be checked at best against standards of consistency...” (Klinke and Renn 2002, p. 1073). By: a) emphasizing the perception of OR as a key component of OR assessment and rejecting; or b) strongly undermining the value of data-driven methods, the interpretivist/constructivist position run the risk of biased OR perception. And this may finally lead to a wrong OR estimation (Rosa 1998).

3.2 Addressing Desirable Property 2: “Understanding of the complex nature of causal mechanisms that underlie OR losses, both expected and unexpected”

Causation in the positivist tradition is built on a number of conditions first introduced by David Hume. According to Cook and Campbell (1979, p.10) “the most important and the most positivistic” of these conditions is the condition of ‘constant conjunction’ (ibid). The latter, also known as the “uniformity of nature assumption” (Lee and Baskerville 2003 p.225) or “perfect regularities” condition (Anderson and Vastag 2004, p. 93), implies a “regularity of constantly conjoined pairs of events (Effect = f(Cause))” (ibid). This is only possible under a closed system assumption where the relationships between components are determined and constant (Blanskar 1978, 1979). The criticism of the ‘constant conjunction’ assumption brought by the positivist tradition relates to the claim that “high correlations demonstrate or are synonymous with causation” (Cook and Campbell 1979, p.10). This “den[i][es] … all conceptual status to unobserved phenomena” and is characterised as “logically inadequate” (ibid).

Focusing on the consistent description of events and behaviours under study rather than on uncovering universal laws (Lee and Baskerville 2003) constructivists/intepretivists emphasize the multidimensionality of factors that constitute the concept of risk (Rosa 1998). At the same time, while accepting that causes of risks are triggered by the interplay of diverse unspecified factors, it reduces the ontological reality of OR causes to the realm of the socially constructed epistemology (ibid). Another point of critique is introduced by Lee and Baskerville (2003). They argue that constructivism (same as positivism) adopts the ‘uniformity of nature assumption’ when presenting research outcomes and generalizing from them. In the context of OR - which is the product of the non-uniform interrelationship between the actors and the components of the inherently open socio-technical systems (Marais 2005; Cooke and Rohleder 2005) - the “uniformity of nature assumption” is not tenable. This said, both positivism and constructivism fail to represent the complexity of the causal dimension of OR.

3.3 Addressing Desirable Property 3: “Adoption of a mode of logical inference that assures consistency and rigour when aggregating the results of multimethod OR assessment”

The above discussion of how positivism and constructivism/interpretivism address the two desirable properties for OR assessment introduced in Table 1 demonstrates that each paradigm focuses on opposite poles of the dichotomous nature of risk. Epistemologically they co-exist as competing and even antagonist perspectives of risk (Klinke and Renn 2002; Rosa 1998). Due to the incommensurability of the ontological assumptions adopted by constructivism and positivism, the conceptual grounding for integrating knowledge supported by both perspectives is still missing (cf Rosa 1998; Downward and Mearman 2007). As both paradigms fail to appropriately support the multimethod OR research and uncover the complexity of OR causality, neither provide a consistent reasoning for rigorously aggregating the results of the multimethod OR assessment.

Proceedings ECIS 2009
In summary, Klinke and Renn’s (2002, p.1076) representation of risk as both a “social construction and representation of reality” proves to be appropriate in the context of addressing the set of desirable properties for supporting OR representation introduced in Table 1. In the next section Critical Realism (CR) is introduced and its ability to address the abovementioned issue is reviewed.

4. CRITICAL REALIST SUPPORT FOR OR REPRESENTATION

Critical Realism (CR) is a philosophy of science that is most commonly associated with the work of Roy Bhaskar (1978, 1979). Introduced as a direct critique of positivism, as well as a postmodern philosophical paradigms including constructivism, CR provides a well-established philosophy of science that strongly emphasises the primacy of ontology over epistemology (Bhaskar 1979). CR comes with an idea that the “reality is both intransitive (existing independently of humans) and stratified” (Mingers 2000, p. 1261). The ‘stratified’ CR ontology divides reality into three distinct ‘domains of being’ (Bhaskar 1979) that are commonly referred to as the domains of ‘real’, ‘actual’ and ‘empirical’ (Danemark 2002; Mingers 2000, 2006). The ‘real’ domain consists of underlying structures and generative mechanisms. Their relations trigger patterns of events and behaviours that reside in the ontological domain of ‘actual’ and represent all events and behaviours produced by causal structures and generative mechanisms even though they are ontologically separated from them. The domain of ‘empirical’ is made up of events that are actually experienced or observed by humans. On the epistemological level, CR aims to provide causal explanations about the limited empirical-based knowledge of events and behaviours. This section explores how the basic ontological/epistemological premises adopted by CR can help address the desirable properties defined in Table 1 that are not fully met by the constructivist/interpretivist and positivist paradigms discussed in Section 3.

4.1 Ontology-driven support for multimethod OR assessment

Critical realists claim that the strengths as well as the limitations of research methods are driven by, but also limited by the ontological position that methods are tied to (Downward and Mearman, 2007). This is in line with Kuhn’s (1970) representation of paradigms as constructs that relate particular research methods to the corresponding ontological, epistemological and other philosophical assumptions. The CR ontological/epistemological position addresses the first desirable property introduced in Section 2 (“A non-contradictory multimethod OR assessment ”): by a) representing an ontologically stratified view of reality as well as the objects of the research inquiry that are components of this reality; and b) claiming that the way reality is constructed determines the way knowledge about reality is obtained and generated (Bhaskar 1979; Mingers 2000; Danermark et al. 2002).

After Rosa (1998) and Klinke and Renn (2002), CR denies the ontological assumption that confines the phenomenon of OR to human perceptions and social construction alone. Hence, OR can neither be reduced to the “objective outcomes [that are] monetary in nature and easily traceable to entities in the general ledger account” (Breden 2007 p 179), nor to objective outcomes in terms of injuries, fatalities etc (Klinke and Renn 2002). As an object of social reality, the phenomenon of OR comprises components that correspond to ontological domains of ‘real’ (OR causal structures and mechanisms), ‘actual’ (the population of all possible operational loss events invoked as a result of causal powers (Bhaskar 1979) of the OR generative mechanisms) and ‘empirical’ (i.e. data accumulated in operational loss event databases). In order to assure rigour and “soundness” (BCBS 2006, p.150) of OR assessment, these ontologically distinct components of OR have to be addressed by multiple research methods that are able to target these research domains (cf Downward and Mearman 2007).

For example, the analysis of internal and external operational loss event data is mostly concerned with the ‘empirical’ component of ORs; and both ‘scenario analysis’ and analysis of ‘factors reflecting business environment and internal control systems’ (BCBS 2004, p.147) aim at uncovering OR at the real ontological domain where the underlying generative mechanisms and structures reside. Understanding that these major OR assessment methodological approaches focus on different
ontological aspects of OR provides additional motivation for combining these approaches, but also allows these approaches to be combined in a non-competitive way by distinguishing the ontological aspects of OR that they target.

4.2 Overcoming reductionism in representing the OR causality

In this subsection the second desirable property, “Understanding of the complex nature of causal mechanisms that underlie OR losses, both expected and unexpected”, is addressed based on the CR understanding of causal mechanisms. These ontologically separated objects generate patterns of events and behaviours that are both perceived/experienced and not.

As mentioned in Section 3, the ontological/epistemological position acquired by the positivist paradigm reduces the representation of causality to ‘perfect regularities’ that was first introduced by David Hume. It does not account though for “imperfect regularities” that may represent an empirical trace for a “genuine causal relation” (Anderson and Vastag 2004, p.93). According to CR the root-cause of the biased representation of causality suggested by both positivism and constructivism relates to the blurring of the ontological and epistemological dimensions. CR represents the object of research as an emerging property of an open system. Thus, CR denies the artificial closure of socio-technical systems that do not correspond to the nature of reality that is emergent and complex. The socio-technical reality being part of the real world is also represented as an open system. Hence, according to CR, the causality underlying the concept of OR is not necessarily a reflection of the succession of distinct causes and effects but rather an emergent property of the complex interplay between the agency and technical components of an open socio-technical system. Taken from this perspective, “[c]ause thus has an ontological depth” (Downward and Mearman 2007, p. 88). Causality underlying the concept of OR is thus the product of an interplay of the complex interrelations between agent’s values and the components of the operational environment that are grounded in the ontological domain of ‘real’.

The ontology-driven position of CR on the nature of causality is in line with the discussion in OR research. Thus, Holmes (2003, p. 84) when providing a “reality check” of “measuring operational risk”, concluded that major operational risk losses cannot be predicted by operational loss data-driven approaches to OR assessment. These are typically characterised as “descriptive and backward-looking, with limited intuition about how key features could create a risk event” (ibid, p. 85). The reason such losses cannot be predicted is because the causal structures that underlie major operational loss events are complex (Diegardt 2007; Marais 2005) and cannot be reduced to the individual OR “drivers” specified by the BCBS (2006, p.150). Rather they represent the complex relationships of these drivers “all acting in combination” (Holmes 2003, p. 84; Wolf 2005). Thus, the ability to understand the nature of complexity underlying OR causes is the key to proactive and more effective assessment and management of ORs (Scandizzo 2005).

4.3 Enabling the exploration of OR causality based on multimethod OR assessment

In this subsection the third desirable property, “Adoption of a mode of logical inference that assures consistency and rigour when aggregating the results of multimethod OR assessment”, is addressed by the retroductive mode of logical inference espoused by CR. This allows knowledge about the OR generative mechanisms and structures to be acquired, based on the limited empirical data.

The retroductive mode of logical inference is an ampliative (synthetic) mode of “inference from effects to explanatory structures” (Hartwig 2007 p. 257). It was first introduced by Charles Peirce as abduction however at the later stage of his research career he used both terms, abduction and retroduction, interchangeably (Lizska 1996). Retroduction aims to obtain knowledge about “what properties are required for a phenomenon to exist” (Danermark et al. 2002, p.206) and as such has been incorporated as part of the CR epistemological apparatus (Downward and Mearman 2007).

In the context of OR assessment, the retroductive process of elicitation of real causes of ORs (that account for both perfect and imperfect regularities) starts by classifying the empirical
events/phenomena (Resolution) and representing them in a theoretically significant way (Redescription) (Mingers 2000). Based on these empirically viable results, a 'creative model' of the possible OR generative mechanism(s) is built. Herein the reasoning moves from the acquired operational loss event data to the postulation of the OR underlying generative mechanism(s), which if it (they) existed would causally enact the operational risk events the given empirical data (Retroduction) (Mingers 2000; Downward and Mearman 2007; Danermark et al. 2002). Here the concept of emergence, which is “the methodologically significant consequence of the stratification of reality” (Danermark 2002, p. 61) plays a role. It allows to “isolate causal mechanisms downwards through strata” to understand the new non-reducible properties and mechanisms that are added to each stratum (ibid). Next, the existence of OR generative mechanism(s) is demonstrated by: firstly, isolating the hypothetical generative mechanism; and secondly, eliminating alternative hypotheses (Elimination). This is followed by identifying a correct OR causal structure in the model under review (Identification).

The adoption of the retroductive mode of logical inference enables the integration of the OR assessment approaches into a unified inquiry about the underlying causes of ORs based on the accumulated empirical data on operational loss events. In contrast to inductive inferential logic that is used by data-driven OR approaches, retroduction discovers new causal connections between the components of the inherently open socio-technical system. Importantly, the outcomes of the retroductive causal inference are not necessarily built on initial premises. This makes retroduction the only mode of logical inference that is able to produce new knowledge (Danermark et al. 2002; Downward and Mearman 2007). Given the incompleteness of the operational loss event databases in organisations (Scandizzo 2005) and the tendency of operational loss data to lose its relevance with the changes of socio-technical environment (Holmes 2004) - without this type of causal inference the understanding of the generative mechanisms underlying ORs would be less structured and more failure-prone.

In this section the basic CR ontological and epistemological assumptions proved to address the aggregated set of desirable properties built in Table 1. Hence, CR proved to be a viable alternative to positivist and constructivist paradigms in supporting the representation of the OR phenomenon at the Requirements Definition phase of the ORIS development life cycle. In the next section we illustrate these findings by extending a conceptual model of an ORIS with a view that supports evaluation of the regulatory capital required under the revised Basel II Accord (BCBS 2006).

5. ILLUSTRATION AND DISCUSSION

Today more than 50% of ORISs developed globally are integrated in the banking sector and thus are driven by Basel II (BCBS 2006) regulatory requirements (Chartis Research Ltd 2008). As the practice of wrong estimation of regulatory capital for OR coverage under the BCBS is still a norm rather than an exception (Jarrow 2008; Moosa 2008), the realistic estimation of OR losses and the application of these estimates to calculate subsequent economic capital is one of the ORIS priorities (Flores et al. 2006). At the same time, based on OR research, a better understanding of the inherent causes of operational failures provides a proactive and more reliable approach for evaluating ORs, as well as for building possible OR scenarios (Scandizzo 2005, 2007; Marais 2005). The identification of inherent causes of empirically gathered operational losses means it is possible to prevent certain major operational failures (Cooke and Rohlender (2005) and/or allocate an appropriate level of economic resources to assure continuity of business processes and sustainability of organizations.

Through exploration of the issues related to designing ORISs in the banking sector, Flores et al. (2006, p. 385) developed a conceptual model that depicted ORIS “calibration” role in OR regulatory capital estimation under AMA (the core of their model is depicted on the right hand side of Fig. 2). The model introduced by Flores et al. (2006) demonstrates that an ORIS potentially faces the danger of under- and overvaluation of the OR regulatory capital requirements due to the inability to appropriately represent and assess “Real Operational Risk” as opposed to “Estimated Operational Risk”. 

Proceedings ECIS 2009
Risk” (ibid p. 385). Thus, the undervaluation of OR by an ORIS is characterized in Flores et al.’s (2006, p. 385) conceptual model as: “the “entity systematically fails to cover risk fully. High possibility of system disruption”. In the case of overvaluation the “entity systematically covers risks with unnecessarily high capital charge. This prevents valuable financial leverage and damages competitive position”.

Thereby, we argue that the ramifications of adopting the CR ontology/epistemology presented in this paper can contribute to solving the issue of discrepancy between “Real Operational Risk” and “Estimated Operational Risk” brought by Flores et al. (2006) (cf Fig. 2). By doing so it contributes to a more effective capital coverage of ORs.

Figure 2  Informing OR assessment and OR regulatory capital estimation results through the retroductive mode of logical inference enabled by CR ontological stratification

As discussed in Section 4, the CR stratified ontology provides a basis for multimethod OR assessment. Further, the retroductive mode of inference espoused by CR provides the necessary epistemological tools to effectively hypothesise the nature of the causal mechanisms that underlie operational loss event data. It has been demonstrated that the empirical operational loss event data can provide access to the intransitive domain of ‘real’ through the ontological domain of ‘empirical’, which is a subset of the ‘actual’ ontological domain. This is represented as step 1 in the process of informing OR assessment and OR regulatory capital estimation results through the retroductive mode of logical inference in Figure 2. The enquiry about the real causes of ORs based on the empirical data was made possible by the adoption of the retroductive inference mode, which is espoused by CR as part of its epistemological apparatus (Downward Mearman 2007) (step 2). Processing operational loss event data with the view of understanding underlying causal mechanisms of ORs is represented as a separate conceptual block that is enabled by Critical Realist retroductive inference (step 3). It should be noted, however, that like Flores et al. (2006) we treat the hypothetical ORIS as a black box (step 4). Therefore, the question of how an ORIS should be designed in order to process the OR scenarios and their causal underpinning enabled by CR retroductive reasoning are outside the scope of this paper. This point is demonstrated by the question mark between the outcome of the CR retroduction and ORIS. Next, the OR assessment results are informed: a) by the empirical operational loss event data that is accumulated in the ORIS (Step 5); and b) by results of the retroductive process that provides better understanding of causal structures that underlie the data about historical OR losses (Step 6). And finally, the amount of OR regulatory capital is calculated (step 7).

Therefore, the ontological/epistemological support provided by Critical Realism helps address the “dilemma” between the full coverage of OR exposures and unnecessary OR capital charge (Flores et al. 2006, p. 385) through a better understanding of OR causes underlying the available historical data on operational losses. From the systems development perspective, it has been demonstrated how
critical realist ontological and epistemological position supports rigorous and consistent representation of the OR phenomenon at the Requirements Definition stage of the ORISs development life cycle.

In this section we extended the Operational Risk-Aware Information System (ORIS) conceptual model of Flores et al. (2006) and illustrated how a Critical Realist ontological/epistemological position supports an ORIS by reporting on the outcomes of OR regulatory capital evaluation in accordance with the requirements of the Revised Basel II Accord (BCBS 2006).

6. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper we set out to address two main research gaps that hinder risk-sensitive approaches to Operational Risk (OR) by Operational Risk-Aware Information Systems (ORISs). The first concerned the lack of explicit ontological/epistemological positioning for representing OR. The second was the lack of understanding about how OR should be explored in a non-contradictory manner as a representational phenomenon forming part of the core of an ORIS.

We have addressed these research gaps by demonstrating how the ontological/epistemological position of the Critical Realist philosophy of science assists: 1) in representing the phenomenon of OR as part of the real system tracked by an ORIS; and 2) at the Requirement Definition stage of the ORIS development life cycle by taking into consideration more informed decisions. In doing so the paper demonstrates how the retroductive mode of logical inference brought by the Critical Realist epistemology informs ORIS in the discovery of OR causal mechanisms under the conditions of limited operational loss event data.

The research is limited by focusing only on the Requirements Definition phase of the systems development life cycle and on the risk assessment stage of the risk management process without covering risk treatment, monitoring and communication (cf AS/NZS 4360:2004). For simplicity the research took a black box approach in representing ORIS, focusing on how its representational properties may be supported by the Critical Realist ontological/epistemological position, and so left the question of how ORIS components should be designed to further research. Future research may include addressing these limitations as well as empirical evaluation of the model suggested in Figure 2.

References


WHY RISK MANAGEMENT MATTERS IN IT OUTSOURCING – A SYSTEMATIC LITERATURE REVIEW AND ELEMENTS OF A RESEARCH AGENDA

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0652.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Offshoring / Outsourcing, Risk, Literature review, IT/IS management</td>
</tr>
</tbody>
</table>
WHY RISK MANAGEMENT MATTERS IN IT OUTSOURCING –
A SYSTEMATIC LITERATURE REVIEW AND ELEMENTS OF A
RESEARCH AGENDA

Martens, Benedikt, University of Osnabru?ck, Katharinenstr. 1, 49069 Osnabru?ck, Germany,
benedikt.martens@uni-osnabru?ck.de
Teuteberg, Frank, University of Osnabru?ck, Katharinenstr. 1, 49069 Osnabru?ck, Germany,
frank.teuteberg@uni-osnabru?ck.de

Abstract

In this paper, we present a systematic literature review in the field of IT Outsourcing with a focus on risk management. The source material of the review consists of 97 high quality journal articles published in 18 journals between 2001 and September 2008. Besides an analysis of related work, this review provides an overview of applied research methods and theories in the field of IT Outsourcing. The articles are then analyzed from a risk management point of view to highlight key risk factors and their specific impact on IT Outsourcing. Identified risk factors are further analyzed in order to assign each risk factor to the phases of a typical IT Outsourcing process (life-cycle). The results of the review show that empirical research is the most applied method and that action research and reference modelling have not been used at all so far. Furthermore, elements of a research agenda are discussed in order to determine further steps to the construction of a reference model for risk management in IT Outsourcing. This paper mainly aims at an audience of experienced researchers in the field of IT Outsourcing who are looking for research ideas and at junior scientists (e.g. PhD students) entering this emerging field of research.

Keywords: IT Outsourcing, Risk Management, Systematic Literature Review, Research Agenda.
1 INTRODUCTION

IT Outsourcing has been one of the most discussed phenomena of the recent years (King & Torkzadeh 2008, p. 205). Considering the high growth rates that are being projected for this industry, the relevance of IT Outsourcing is strongly increasing. A cross-European study conducted by Ernst & Young in 2008 demonstrates the general significance of IT Outsourcing. According to this survey, 68% of all European companies have outsourced their IT at least partially. Seeing that, each company should implement a risk management system for IT Outsourcing to reach the business objectives. This can only be achieved by considering main risks and required safeguards (Bahli & Rivard 2003). A generally growing interest in this topic is also mirrored by its presence in major scientific research journals. For example, MIS Quarterly released a special issue dedicated to the topic of offshore outsourcing in June 2008. The field of risk management has been investigated by scientists from 1988 to 2000 (cf. Dibbern et al. 2004, pp. 34-38, 54). Dibbern et al. identified articles with a focus on topics of risk management in IT Outsourcing, ranging from 1994 to 1999. In our review we continue this work and analyze articles ranging from 2001 to 2008. The objective of this article is to provide an overview of the current research status on risk management in IT Outsourcing as it presents itself in international high quality journals. By means of a systematic literature review, an analysis of past, present and future research on IT Outsourcing and risk management in IT Outsourcing is presented.

Generally, the term ‘outsourcing’ applies to the transfer of business units/ functions to external service providers (Dibbern et al. 2004, pp. 9-10, Lee et al. 2000). In IT Outsourcing, this transfer process is limited to the realm of IT, where it can encompass a wide range of measures – from hiring an external software developer to delegating the management of a company’s complete data processing center to external service providers. The reasons for IT Outsourcing can be put down to a general shift in business strategies. The former diversification strategy aiming at risk distribution that used to be predominant until quite recently has now been replaced by a focus on the company’s core competencies. IT is usually not regarded as a core competency and is therefore frequently outsourced by companies. In the course of the IT Outsourcing transaction, economies of scale and external expert knowledge contribute to a more efficient realization and administration of IT. A second reason for IT Outsourcing lies in the current discussion about the incalculable value contribution of IT (Levina & Ross 2003, p. 332). Executive managers often regard the costs of IT as overhead costs which should be reduced to a minimum. In summary, IT Outsourcing can be regarded as a make-or-buy-decision in the context of IT management.

The paper is structured as follows: section 2 starts with a discussion of the applied research method (systematic literature review), the underlying review process and related work. The results of the systematic literature review are presented in section 3, which starts by presenting different classifications (e.g. research methods applied, theories applied and research topics) of the analyzed articles and ends with a systematic analysis of risk factors. Section 4 takes a critical look at the overall results and provides a summary of findings and recommendations for further research activities in the field of risk management in IT Outsourcing. An outlook on future research trends is also given.

2 RESEARCH METHODOLOGY AND FRAMEWORK OF ANALYSIS

2.1 Systematic Literature Review and Analyzed Journals

Considering the increasing number of books, journals, conferences and workshops, the application of a systematic literature review has become indispensable (Fettke 2006, p. 257). In a systematic literature review, relevant work and current findings are analysed with regard to a particular research question. The objective is to describe, summarize, assess, appraise, resolve or to integrate selected primary research results. The focus of the analysis can be set on methodology, theory, content or other aspects. Webster and Watson (2002) show a more content-oriented approach to writing a review article. In
their opinion, the main motivation for a research article should always lie in the additional benefit it has for the research community. Every figure and evaluation should increase the research community's knowledge. The thematic scope of the article should also be clearly defined. In addition, a methodology or model should be developed that further research could be based on. Finally, a review should imply conclusions relevant for other researchers and managers alike. To improve the quality of the analyses, both authors of this paper were involved in reviewing and coding the analyzed articles. The inter-rater reliability was good (inter-rater percentage agreement: > 90% in all analyses). The limitations of a systematic literature review lie in the paper selection and categorization process, which requires some judgment calls. However, we tried to minimize this risk by following a proven course of action for the creation of a literature review (Swanson & Ramiller 1993). The restriction of the source material to high quality articles leads to reliable results about the state of the art of IT Outsourcing.

The analyzed journals and the number of published articles regarding IT Outsourcing are shown in Figure 1. The selection of journals is relied on the journal ranking list published by WKWI (2008). This list has a wide acceptance among researchers. It declares 23 information systems journals as high quality journals. Figure 1 shows that the journals “Communications of the ACM”, “Information & Management” and “MIS Quarterly” make up a large part with 14, 10 and 14 published articles on IT Outsourcing. The analyzed time period spans 8 years, from 2001 to September 2008, because a review article by Dibbern et al. 2004 analyzed articles dating from 1988 to the year 2000. Amongst others, they analyzed papers on risk management in IT Outsourcing (Dibbern et al. 2004, pp. 34-38, 54). The work of Gonzales et al. 2005 spans a time period from 1988 to 2005 but does not analyze the topic of risk management in detail. The search for relevant articles was conducted as follows: As a first step, a matching list of pre-defined key words was applied to the search engine of each journal or to publisher independent journal data bases like EBSCO (Business Source Complete, EconLit (full text)) or Science Direct.

![Figure 1. Total number of articles per journal.](image)

To achieve extensive search results the following key words were used: Cloud Computing, Offshoring, Nearshoring, Outsourcing, IT Outsourcing, Sourcing, Application Outsourcing, Contracting Out, On-demand Computing, Application Service Provider, Software as a Service, Software on demand, Netsourcing, Internet as a platform, Utility Computing, Service Industry, Service Engineering, Service Outsourcing, IT Service Provider, Business Service Provider, Commerce Service Provider, Storage Service Provider, Vertical Service Provider, Full Service Provider and Managed Service Provider. The inclusion of many synonyms and/or semantically very similar expressions led to more exhaustive search results. The overall results were first saved in a list. In a second step, each article was checked for its relevance to the topic of IT Outsourcing by reading the respective abstract. The selected 97 articles were then analyzed by using the method of a systematic literature review.
Following Swanson & Ramiller 1993 we studied each paper’s abstract, introduction, discussion section and conclusion to classify the paper according to its research method, theoretical lens and research topic.

2.2 Related Work

In order to determine the research status it is necessary to evaluate former literature reviews about IT Outsourcing. In total, the authors found 5 different review articles dealing with IT Outsourcing (cf. Table 1).

| Title, Journal, Year, Authors | Information Technology Outsourcing in Australia – a literature review, Information Management & Computer Security, 2001, Christina Costa (Costa 2001) |
| Period, Publications reviewed | Period not explicitly named; Articles, books, surveys |
| Research objective(s) | analysis of economic and technical considerations regarding IT Outsourcing: e.g. success factors of IT Outsourcing arrangements |
| Review method | not explicitly described |
| Results | an overview of: cost reduction, technical considerations and relevance of core activities |

| Research objective(s) | framework for cataloging literature about IT Outsourcing: research foci, theoretical perspective, definition of the methodologies utilized to conduct the analysis, areas of consensus and suggestions for future research |
| Review method | search of journals by using “outsourc$” as search phrase; reading of abstract, introduction, discussion section and conclusion of articles to determine their research focus, theoretical foundation and methodology |
| Results | IT Outsourcing research is dynamic and vibrant: growing research synergy, healthy exchange between researchers; early research analyses: why, what function and how to outsource IT systems or services; recent research analyses: relationships and psychological aspects |

| Title, Journal, Year, Authors | Information systems outsourcing: A literature analysis, Information & Management, 2006, Reyes Gonzalez, Jose Gascoa, Juan Llopisa (Gonzalez, Gascoa & Llopisa 2006) |
| Period, Publications reviewed | 1988-2005, 131 articles ("prestigious” journals [10] as well as management and business journals [8]) |
| Research objective(s) | identification of main topics, most frequently applied methodologies, survey of authors and countries contributing to IT Outsourcing, suggestions to improve research on IT Outsourcing |
| Review method | search for several terms in the ABI database; and manual search in indices of the journals |
| Results | progressive growth of the topic, empirical research is often conducted (especially field studies), examples of new topics: computer staff and job threat |

| Title, Journal, Year, Authors | Why 'Nearshore' means that Distance matters, Communications of the ACM, 2007, Erran Carmel, Pamela Abbott (Carmel & Abbott 2007) |
| Period, Publications reviewed | 1998-2006, 150 textual sources: journal/magazine (45%), promotional (23%), web-based texts (22%), consulting articles (7%), and academic texts (3%) |
| Research objective(s) | categorization of publications into nearshoring location, dimensions of the nearshoring construct and evidence for the assertion of difference between near- and offshore |
| Review method | critical, systematic, and qualitative content analytical method; inductively derived coding categories which can be used as a basis for the analysis and can then be applied deductively to the textual sources |
| Results | “nearshore” has become a convenient label; locational and geographical differences will continue to play a role; distance is viewed as multi-dimensional (e.g. physical meters and time zones) |

| Period, Publications reviewed | June 2008, 43 articles submitted to MISQ for a special issue on IS Offshore Outsourcing (solely empirical research) |
| Research objective(s) | presentation of a representative extract from research that is being conducted on the topic of IS offshore outsourcing |
| Review method | extraction of research question(s), definition of offshoring (if given), context of the study, theoretical perspective(s), vendor countries, sample frame, general research method, analysis method, findings |
| Results | research regarding offshoring is still in nascent phase; research is still mainly qualitative and/or exploratory |

Table 1: Related Work.
Table 1 categorizes all reviews in chronological order and summarizes their research objectives, review methods and results. The specific difference of our article from the 5 indentified review articles (cf. table1) is:

- Only papers in high quality journals are analyzed. Specific analyses of the current research status and the articles’ theoretical perspective are conducted.
- In particular, risk management articles on IT Outsourcing are categorized and in-depth analyzed.

3 ANALYSIS OF RESULTS

3.1 Distribution of Articles over time

Figure 2 illustrates the development of practical, empirical and other (i.e. non-empirical and non-practical) articles over the last eight years. An article was counted as practitioner article if at least one of the contributing authors had a distinctly practical orientation that was outlined in the author information. In the years 2003 and 2008 an especially high number of articles on IT Outsourcing was published. Generally, the number of articles has grown during the time period discussed, reaching its peak in 2008. The number of published scientific articles follows the same pattern. The number of articles with empirical approaches, which amount to 31% of all papers published in the covered time span of 8 years, also shows an upward trend. However, the application frequency of the empirical research method has declined from 50% in the year 2001, to 20% in 2004 and 26% in 2008. The application of empirical methods can generally be regarded as a sign of scientific progress, because this research approach validates previously developed models, methods and concepts, sharpens the research area and contributes to its long-term establishment (Webster & Watson 2002, p. xiii). Finally, the fact that IT Outsourcing is a concept developed in practice (Dibbern et al. 2004, p. 9) explains the existence of a (small) number of practitioners who contribute occasionally to scientific research articles.

![Figure 2. Total Number of Articles per Year during the Period from 2001 to September 2008.](image)

3.2 Research Methods

The research methods that are commonly applied in IT Outsourcing are listed in Figure 3. Due to the distribution of articles to more than one research method the overall number is higher than 97. This also applies generally to the following analyses. A spectrum of IS research methods that we have applied is shown in Wilde and Hess 2007. The analysis shows that empirical research methods dominate the area of IT Outsourcing. The high number of conducted case studies underlines the strong practical component of IT Outsourcing. Formal models as well as trend and development analyses were included in the category of formal argumentative-deductive research. The authors realized that in most cases the hypotheses of empirical analyses were not based on pre-developed models, concepts or methods, but rather on general aspects of IT Outsourcing to keep track of changing demands.
Furthermore, one should note that in spite of numerous new models being developed, their practical implementation by means of empirical methods or case studies (including an iterative improvement process) is mostly either incomplete or does not happen at all. The three review articles identified in our systematic literature review are complemented by two more articles in Table 1. These were identified by means of an unsystematic review of literature conducted with search engines like EBSCO, Science Direct, etc. Furthermore, compared to explorative research, other methods are clearly underrepresented in this research area. But also previously unused research methods like action research or reference modeling should be applied in order to gain new insights. Action research aims to solve current practical problems while expanding scientific knowledge by using an iterative research process that both researchers and research subjects benefit from (Baskerville & Myers 2004). A generic reference model represents a class of domains; e.g. the development or improvement of processes, systems and organizational demands in certain fields (Fettke & Loos 2003). For the reference modeling approach BITKOM recommended the development of a reference model that supports the realization and implementation of IT Outsourcing.

![Figure 3. Research Methods Applied in IT Outsourcing.](image)

### 3.3 Theories Applied

Figure 4 illustrates the different theories applied in the 97 analyzed articles. The data shows that IT Outsourcing researchers strongly tend to base their approach on the transaction costs and agency theory. Other theories like incomplete contract theory or psychological contract theory are not very frequently utilized. The transaction cost theory is applicable because it deals with asset specificity, overall cost advantage, the threat of opportunistic vendor behaviour, and the complexity of the transaction (Bahli & Rivard 2003).

![Figure 4. Theories applied in IT Outsourcing research.](image)

Agency theory deals with principal-agent conflicts, information asymmetry and goal incongruence between an agent (the external vendor) and a principal (the client). Obviously, researchers in this area are still mainly relying on theories which have already been applied before, but they have also started to test other theories for their applicability to IT Outsourcing. This shows that the research area of IT...
Outsourcing is vivid, because both established and new theories play a role in it. As a result of a future literature review, a list of new theories to be established would be desirable. Furthermore, the application of multi-theoretical approaches could be especially identified.

3.4 Risk Management in IT Outsourcing

Main phases of risk management are context analysis (1), risk identification (2), risk analysis (3), risk evaluation (4), risk treatment (5), monitoring and review (6) and communication and continuous improvement of risk strategy (7). Since at least two companies are involved in an IT Outsourcing process, the range of risk factors expands and new risks could appear which are yet unknown to client companies. Furthermore, due to an increase in national and international regulations and quasi-standards (e.g. SOX, EuroSOX, Basel I, ITIL, Cobit, etc.) the risks that companies are facing grow and risk management for IT Outsourcing is explicitly demanded by suppliers, clients and states. To identify the topics of the articles we first chose an inductive approach and generated a topic list by reading each article. In a second step we aligned our list with the categorization by Gonzales et al. 2006. Figure 5 presents the results and illustrates the topics that IT Outsourcing articles are dealing with. It shows that risk management presently takes a relatively weak position. Apart from the chosen topic of risk management other articles address domains like quality issues or the project management of IT Outsourcing, but these issues are secondary if the IT Outsourcing arrangement is not successful because main risks and required safeguards are disregarded (Bahli & Rivard 2003). In total, 11 articles which deal with risk management in IT Outsourcing have been identified and examined in detail.

Figure 5. Distribution of Articles by Topics.

Figure 6 illustrates the distribution of research methods applied in the articles. Besides empirical research, formal, argumentative deductive analysis is often used to validate theoretical concepts and to formulate formal models for risk management functions.

Figure 6. Scope of Analysis for Risk Management in IT Outsourcing.

For prototyping and reviewing, Figure 6 indicates a research gap. The method of prototyping could be applied in the development and implementation of software for IT Outsourcing decisions and for facilitating IT Outsourcing (e.g. Service-oriented Architectures).
<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.7 High Asset Specificity (3): overspending due to high transaction costs and a small number of providers on market</td>
<td>Bahli &amp; Rivard 2003, Oh, Gallivan &amp; Kim 2006, Gellings &amp; Wüllenweber 2006</td>
</tr>
<tr>
<td>1.8 Low Financial Stability (3): the provider’s financial stability is important</td>
<td>Günther et al. 2001, Aubert, Patry &amp; Rivard 2006, Ngwenyama &amp; Sullivan 2006</td>
</tr>
<tr>
<td>2.5 High Performance Oscillation (3): provided performance after contract conclusion has high oscillations</td>
<td>Iacovou &amp; Nakatsu 2008, Saeed &amp; Leitch 2003, Tafati 2005</td>
</tr>
<tr>
<td>2.6 Excessive Dependence on Provider (2): customer has a limited scope of action</td>
<td>Gonzales, Gasco &amp; Llopis 2004, Kern, Kreijger &amp; Willcocks 2002</td>
</tr>
<tr>
<td>2.7 High Task Complexity (1): the service or task complexity influences the achievement of objectives.</td>
<td>Aubert, Patry &amp; Rivard 2006</td>
</tr>
<tr>
<td>3.3 Lack of Customer Expertise with Law (4): gained provider experience regarding IT Outsourcing contracts</td>
<td>Bahli &amp; Rivard 2003, Tafati 2005</td>
</tr>
<tr>
<td>3.4 Irreversibility of outsourcing decision (1): back-sourcing is usually not economical</td>
<td>Gonzales, Gasco &amp; Llopis 2004</td>
</tr>
<tr>
<td>5.2 Poor User Integration (2): IT users have insufficient influence on the IT Outsourcing project/services</td>
<td>Gonzales, Gasco &amp; Llopis 2004, Tafati 2005</td>
</tr>
</tbody>
</table>

Table 3: Risk Factors in IT Outsourcing.

The risk factors listed in Table 3 were extracted from the analyzed 97 articles and were assigned to one of the five categories (economical, organizational, legal, technical or psychological). The three most frequently mentioned and therefore most discussed risk factors are Quality Inferior to...
Anticipation and Lack of Privacy/Data Security. A total of 23 different risk factors have been identified.

The risk factors listed in table 3 have specific effects and macro effects on the IT Outsourcing process. Figure 7 illustrates this causal relationship. The effects range from budget overrun (wrong cost expectations) to bad financial/economic performance. Every mentioned risk factor can trigger any of the possible effects, which in return may lead to macro effects. Macro effects are classified as one of four IT Outsourcing failure levels: IT Outsourcing process failure (the process/project is not completed within the time and budget); expectation failure (IT services and new processes do not match user and company expectations); interaction failure (users’ attitudes towards outsourced IT services are negative); and correspondence failure (there is a discrepancy between IT Outsourcing process, performed service and the planned objectives).

Figure 7. Risk Factors, Effects and Macro Effects (cf. Aloini et al. 2007, p. 553).

Figure 8 exemplifies a typical IT Outsourcing process (Georgius & Heinzl 2005) with risk factors assigned to the individual process steps. The process is initiated by the customer and ends with the expiration or renegotiation of the contract. The risk factors are taken from table 3. The Information Security Forum reported in 2008 that many companies only accounted for their potential IT risks after the outsourcing process had been completed. This makes it all the more necessary to develop, implement and establish risk methods and concepts for IT Outsourcing. As a guideline, the distribution of risks along a typical outsourcing process can be useful for risk prevention in IT Outsourcing.

Figure 8. Risk Factors in a Typical Outsourcing Process (cf. Georgius & Heinzl 2005, p. 5).
4 OUTLOOK ON FORTHCOMING RESEARCH

4.1 Current Trends in IT Outsourcing Research

A trend clearly indicated by the analysis is the special focus that researchers placed on offshore outsourcing in the year 2008. It is also apparent that IT Outsourcing is not location dependent any more, which creates new tasks and challenges. Two factors for this development are certainly the globalization of IT and the improvement of ICT. The total number of papers on IT offshore outsourcing has continuously increased, ranging from 0 in the year 2001 to 1 in the years 2002 to 2004 and 10 in the year 2008. Indirect requests for research by several ‘Calls for Papers’ have influenced this development as well (e.g. MIS Quarterly, June 2008). Other current developments show that IT Outsourcing is strongly influenced by trends like cloud computing or the software as a service model. The notion of “cloud computing” has been especially dominant in journals aimed at readers with a practical background. Cloud computing could cause major changes in IT business in the near future; several providers like Amazon, Salesforce and Google are already offering IT services via the internet which are processed by the “cloud” (Hayes 2008, p. 10). Along with the increasing spread of these concepts and technologies, new fields of activity entailing new risk factors emerge and require a new design of risk management in IT Outsourcing (e.g. Buyya et al. 2008, p. 5).

4.2 Elements of a Research Agenda

Looking at the research results in Table 3 we noticed that a stronger focus on performance measurement seems to be needed to support the progress of ongoing research (Rustagi et al. 2008, pp.127-128). Another gap which was discovered while assigning the risks to the steps of the outsourcing process in Figure 8 is the small number of risks in the last process step. However, more risk factors that need to be considered may be identified through more detailed examination. Also, it became apparent that the risks and benefits of multi vendor sourcing are not well known yet (e.g. Sia et al. 2008). Furthermore, none of the analyzed articles applied the method of reference modeling. To the authors’ knowledge, no reference model for risk management in IT Outsourcing has been published to date. Risk management is not popular compared to other research topics in IT Outsourcing, but it is necessary for companies. Therefore, a proposed research agenda illustrated in Figure 9 includes the development of a reference model for risk management in IT Outsourcing which is based on best practices. Cooperating business partners could then implement such a reference model, which could be continuously improved by means of iterative loops. Figure 9 illustrates one way of constructing such a reference model for risk management in IT Outsourcing by applying the presented research methods. Already existing models, as e.g. the „eSourcing Capability Model for Client Organizations” (Hefley et al. 2006) or the COBIT framework could support these specific steps. In a next step, a comparison of the presented results with methods and theories of general IT risk management could strengthen the research agenda and is therefore projected by the authors.

Figure 9. Elements of a Research Agenda.
Acknowledgement: The authors are indebted to Ms Anja Grube and several anonymous reviewers for fruitful discussions and substantive comments relating to this article.

References
A complete list of the 97 analyzed articles in our literature review is available for download at: www.uwi.uos.de/itoutsourcingreview.pdf.


Proceedings ECIS 2009
EXAMINING THE RELATIONSHIP BETWEEN TRUST AND CONTROL IN IT OUTSOURCING RELATIONSHIPS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0672.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Trust, Offshoring / Outsourcing, Dynamic Capabilities / relationships / perspective, IT/IS management</td>
</tr>
</tbody>
</table>
EXAMINING THE RELATIONSHIP BETWEEN TRUST AND CONTROL IN IT OUTSOURCING RELATIONSHIPS

Beimborn, Daniel, Department of Information Systems and Services, University of Bamberg, Feldkirchenstr. 21, 96045 Bamberg, Germany, daniel.beimborn@uni-bamberg.de

Schlosser, Frank, Department of Information Systems and Services, University of Bamberg, Feldkirchenstr. 21, 96045 Bamberg, Germany, frank.schlosser@uni-bamberg.de

Weitzel, Tim, Department of Information Systems and Services, University of Bamberg, Feldkirchenstr. 21, 96045 Bamberg, Germany, tim.weitzel@uni-bamberg.de

Abstract

What is the role of control in maintaining trust in outsourcing relationships? Although the literature is quite rich on conceptualizing the relationship between control and trust in inter-organizational relationships, there exist quite sparse quantitative works which help to evaluate the models developed. In this paper, we analyze data from 156 IT outsourcing relationships of German banks in order to get insights into the actual relationships between different modes of control and the level of relational trust in these relationships. Additionally, we examine the role of service quality (measured in terms of reliability and responsiveness) in this context. The results show that trust is positively related with most modes of control, and that there is indicative evidence that control supports trust in high-service quality situations while it leads to a reduction of trust in relationships suffering from bad service quality (cycle of trust vs. distrust).

Keywords: Trust, Control, IT Outsourcing, Relationship Management, PLS, Survey
1 MOTIVATION

Although there exists a quite mature strand of research on how to manage an IT outsourcing relationship, only a few works have actually empirically studied the field of relational governance in this context. Among them are both quite early and fundamental works like (Grover et al. 1996; Lee and Kim 1999), which evaluate the role of relationship determinants on outsourcing success, and younger works, such as (Goles and Chin 2005) which develops a measurement model for relationship quality in IT outsourcing. Nevertheless, works have seldom trickled down the outsourcing governance determinants to a level on which direct and usable implications for the management of outsourcing relationships can be identified. For example, what are the control mechanisms and items that actually impact relationship quality dimensions like commitment and trust? By focusing on a particular aspect of this research strand, this paper wants to contribute both to a deeper and more managerial understanding of the relationship between control and trust in outsourcing relationships. There is a quite rich literature on conceptualizing the relationships between control and trust in inter-organizational relationships (IOR); nevertheless, there is not much quantitative research found that empirically evaluates these concepts and theories in IOR, in general, and in IT outsourcing relationships, in particular. One of these works is (Langfield-Smith and Smith 2003) which investigates the impact of control mechanisms from a transaction cost economics perspective by conducting a single case study. Further, the existing literature that has investigated the interrelationship between control and trust has derived contradicting findings, up to now. While some researchers found control to be positively related with trust, others suggest negative interrelations (Das and Teng 2001; Langfield-Smith and Smith 2003).

In order to contribute to this important strand of research, we want to answer the following research question: What is the relationship between trust and different managerial control mechanisms in outsourcing relationships? How does the level of service quality, in terms of reliability and responsiveness, affect this relationship?

The remainder of this paper is structured as follows: section 2 introduces related work and develops our research hypotheses. Section 3 outlines the chosen approach and gives an overview about our sample while section 4 empirically evaluates the proposed hypotheses. Section 5 discusses the results and derives implications for research and management before section 6 concludes the paper.

2 RESEARCH MODEL

Previous literature has highlighted the shortcomings of solely considering contractual issues for assessing the quality of IT outsourcing arrangements, as those contracts concluded at the beginning of an outsourcing relationship do not comprise all future eventualities and thus are inherently incomplete (Hart 1988; Macneil 1980). Consequently, researchers have shed light on the importance of a good cooperative relationship being another relevant facet of outsourcing management (Klepper 1995; Lacity and Willcocks 1995; McFarlan and Nolan 1995). In a first empirical investigation, Grover et al. (1996) have addressed this extended view of an outsourcing relationship by factors like trust and commitment which reduce threats from opportunistic behavior, help to better integrate partners and to reduce formal contracting. Other authors suggested incorporating additional aspects or dimensions aiming at a better and more comprehensive understanding of outsourcing relationship quality. For a good overview and summary on this as well as an empirical validation, see (Goles and Chin 2005).

Although it is difficult to capture and to measure, inter-firm trust has shown to be a fundamental dimension of relationship quality (Kanter 1994; Ring and Van de Ven 1994). Trust has been defined as “the firm’s belief that another company will perform actions that will result in positive outcomes for the firm, as well as not take unexpected actions that would result in negative outcomes for the firm” (Anderson and Narus 1990). Since the early 1990s, there have been numerous works which defined trust as a multi-dimensional concept (Das and Teng 2001). One of the most popular conceptualizations distinguishes between competence trust and relational trust (Anderson and Narus 1990; Das and Teng...
2001) – or “concern a partner’s ability to perform according to agreements (competence trust), or his intentions to do so (goodwill trust)” (Nooteboom 1996, 990, p. 990). Others have proposed a distinction between the need for trust and the actual level of trust (Gallivan and Depledge 2003). In this work we focus on the latter one.

The outsourcing literature moreover has shown the importance of trust for maintaining and managing an ongoing outsourcing relationship and tested its contribution to outsourcing success in numerous qualitative and quantitative works (Goo and Nam 2007; Grover et al. 1996; Kern 1997; Langfield-Smith and Smith 2003; Saherwal 1999; Willcocks and Kern 1998). In this study, we investigate the relationship, and thus trust, between client and vendor from the client’s perspective.

One of the antecedents of low outsourcing risk, high service quality, and success in IT outsourcing is monitoring the provider (“control”). In general, control is related to comparing as-is and to-be in order to reach the goals set by the firm through corrective actions, where deviations appear. However, since esp. in the area of IS, outcomes and corrective actions are not always clear ex ante, a broader view on control has been discussed (Kirsch 1997; Lee et al. 2008). Osterloh and Weibel (2006) differ between the control object, the point in time, and the control direction. The control object itself can then be examined in several ways. First, process control or behavioral control involves controlling process execution and also monitoring the employees’ behavior (Ouchi and Maguire 1975). Second, result control or outcome control traditionally is the comparison of as-is and to-be (Das and Teng 2001). Both behavioral and outcome control represent formal modes of control and thus can be characterized as a performance evaluation strategy (Eisenhardt 1985). Third, social control adheres to the compliance of codes and values (Ouchi 1979), is known as an informal mode of control, and can e.g. be assessed at different levels like group (clan control) and individual (self control) level (Kirsch 1997). Control can be conducted in an ongoing manner during an arrangement (feed forward controls) or at its end (feed-back control). In both cases, directed controls implemented to check for determined behaviors and performance standards, and undirected controls like e.g. early warning systems are possible. In general, the mode of control that is appropriate and therefore should be implemented is dependent on the context and the available information, e.g. it makes sense to implement outcome control when outcomes are measurable (Kirsch 1997).

In short term, it might be attractive from an outsourcing firm’s point of view to achieve a high level of control. This helps to identify variations and to initiate retaliatory action as early as possible, thus avoiding escalations (Zaheer and Venkatraman 1995). When viable controls are agreed upon in the outsourcing contract, this can have positive effects in a twofold way. While the service receiver can monitor the quantity and quality of the service, the service provider can substantiate that the delivered services meet the standards as determined in the contract (Langfield-Smith and Smith 2003). Thus, control can also be viewed positively from a vendor’s perspective.

However, accounting for elements like trust as described above when intending to build a long-term relationship, too much of control can result in undesired effects within a work relationship and destroy trust (Lorange and Roos 1992). The occurrence of such effects does not reflect the original goals of the controlling party and reasons can be found in an employee’s strive for relative independence. This can lead to a vicious circle as described by Argyris (1952) in a way that more control causes a decrease of satisfaction, causes less willingness to perform, causes less efficiency, causes more control, and so on. This is confirmed by Das and Teng (2001) who argue that formal control will undermine trust since it takes autonomy and independence away from the controlled partner firm. Considering this, one central issue for managers of outsourcing relationships would be to choose the right or optimal level of control, both from an organizational point of view (control mechanisms and systems) and a social point of view. In doing so, the benefits (reduced incentives for opportunistic behavior) and the costs (e.g. reduction of trust level) of control have to be balanced carefully.

Nevertheless, as already outlined, control can also have positive impacts in building and sustaining trust because specifying performance goals and controlling that they are achieved helps to motivate people who do their job well (Sitkin 1995) and because control reduces relational risks and thus facili-
tates trust and collaboration (Coletti et al. 2005). More control and certainty can lead to greater satisfaction, better conflict avoiding, cost reduction, and the development of trust in buyer/seller relationships (Marcolin and McLellan 1998). Moreover, Das and Teng propose a positive relationship by reversing the causality: “goodwill trust and competence trust will enhance the effectiveness of all control modes (behaviour, output, and social) in an alliance” (Das and Teng 2001, 265, p. 265). In their conceptual work on strategic B2B alliances, they argue that result controls will have a negative effect on trust while social controls lead to a higher level of trust. This is in line with findings that people who aim at building trustful relationships tend to avoid formal controls (McEvily et al. 2003).

Summarizing the different lines of argumentation, we hypothesize the different modes of control to be either positively, negatively, or concavely (reverse-U shaped) associated with the level of trust.

**H1a:** Control will be negatively related with the client’s trust in the vendor.

**H1b:** Control will be positively related with the client’s trust in the vendor.

**H1c:** There will be a concave relationship between control and the client’s trust in the vendor.

Furthermore, literature has discussed the existence of cycles of trust and distrust (Ring 1996) or spiraling effects (Marcolin 2002) when considering positive and negative impacts of control. A cycle of trust does appear if control is used to credit the provider for high performance. Good results from controlling the provider lead to increasing trust. By contrast, if the provider does not deliver high service quality, the reverse effect will appear: trust is destroyed by the cycle of distrust, since the client firm increases the control level and may find more deficiencies. Correspondingly, Sabherwal (1999) showed in a case study that good performance in an outsourcing relationship arises from a balance between trust and structural control, while the opposite will occur when an inappropriate structure comes together with a specific type of trust (over or under control).

This high interdependence between trust, control, and service quality as contingency variable leads to our second hypothesis.

**H2:** High service quality will lead to a cycle of the client’s trust in the vendor, while low service quality will lead to a cycle of distrust.

3  METHODOLOGY

3.1  Unit of Analysis and Approach

In our study, we surveyed the relationship of the 1,000 largest German banks to their IT service provider who is responsible for providing the primary loans system (i.e. operating and maintaining the IS used to grant and manage private construction loans and mortgages). The data were collected between April and October 2008.

In Germany, all banks belong to one of three sectors: while commercial banks are mostly independent when selecting an IT service provider, public savings banks and cooperatives both have joint IT service firm subsidiaries within their sectors. Moreover, particularly banks that belong to one of the latter two sectors show very ancient IT outsourcing relationships (cf. Table 1).

<table>
<thead>
<tr>
<th>Bank type</th>
<th>Population</th>
<th>Sample</th>
<th>Duration of relationship to IT service provider</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>0-5 years</td>
</tr>
<tr>
<td>Commercial banks</td>
<td>136 (13.6%)</td>
<td>23 (13.6%)</td>
<td>3</td>
</tr>
<tr>
<td>Savings banks</td>
<td>422 (42.2%)</td>
<td>55 (32.5%)</td>
<td>6</td>
</tr>
<tr>
<td>Cooperatives</td>
<td>442 (44.2%)</td>
<td>91 (53.9%)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>1,000 (100.0%)</td>
<td>169 (100.0%)</td>
<td>12</td>
</tr>
</tbody>
</table>

**Table 1.** Distribution of bank types in the overall population and in the sample
All banks have been initially contacted via phone in order to identify the appropriate person (Chief Information Officer or other IT manager, responsible for managing the relationship to the IT provider) and to ensure that the relevant IT system is not provided by an in-house unit. If it turned out that the bank was not engaged in outsourcing, it was dropped from the list of addressees and been replaced by the next largest one. Only 52 banks out of the initially largest 1,000 stated to have the operations of the loans system provided in-house. All other managers received the questionnaire (via mail, fax, or e-mail, depending on their preferences). After two reminders (paper-based and call-based) we eventually received 169 usable questionnaires (response rate = 16.9%).

For testing the basic relationship between trust and control (H1) we will provide both a regression analysis and a group comparison which allows us for a very simple but robust identification of a change in the relationship (such as concave relation which turns from an increasing relationship to an decreasing relationship at a certain degree of control). The second hypothesis is tested by hierarchical regression-based moderator tests and group comparisons.

3.2 Measurement

All of our constructs were measured by multi-item scales in order to increase reliability. The particularities are explained in the following and the indicators are listed in Table 7 in the appendix.

Control: To better understand the impact of control, we distinguish two of the three dimensions as discussed in the literature review above but more strongly dedicate them to the application domain in order to test for control action items: First, performance reports delivered from the IT service provider to the bank (CONTROL_REP) represent formal result control mechanisms (output control). Second, actions conducted by the provider firm itself (CONTROL_PROV) are examined, such as employee satisfaction surveys on the client side (output control). Third, detailed activity monitoring done by the client itself (CONTROL_BANK) represent the third dimension (behavioral control).

Trust (TRUST): We focus on the concept of relational trust. This construct was measured by reflective indicators derived from (Goles and Chin 2005; Kettinger and Lee 1995; Parasuraman et al. 1988).

Service quality: For capturing service quality, SERVQUAL (Parasuraman et al. 1988) has become one of the most commonly used measurement instruments both in Marketing Sciences but also in Information Systems (e.g., IS-SERVQUAL (Kettinger and Lee 1994)). From this instrument, we take reliability (i.e., ability to perform service dependably and accurately; SQ_reliability) (Grover et al. 1996; Lee and Kim 1999) and responsiveness (flexible reaction to problems and changing demands from client side; SQ_responsiveness). Especially, responsiveness often shows to be a critical facet of service quality when asking vendor managers.

To be able to conduct group comparisons and hierarchical regression analyses, we applied confirmatory factor analysis (CFA; principal component analysis, in particular) to achieve construct scores from the multi-item measures.

4 RESULTS

In this section, the proposed hypotheses are empirically tested, as described above. However, first of all we evaluate the measurement model in order to ensure reliability and validity of the scales used.

4.1 Validation of the Measurement Models

All of our tests are based on reflectively measured constructs. We deleted all cases which showed a missing value in any of the items used. This resulted in a reduced sample of 156 data sets.

Measurement instruments have to be analyzed regarding content validity, indicator reliability, and construct validity. Content validity examines the degree to which the supposed meaning of a construct
is reflected by its measures (Boudreau et al. 2001). Content validity was ensured by developing questions for indicators from preceding research as well as by performing pre-tests to check for ambiguities. The findings from the pre-tests were incorporated into the questionnaire after adaptation or elimination of single questions. For ensuring *indicator reliability*, loadings should be significant and above 0.7 (Hulland 1999) which is fulfilled by all our indicators (cf. Table 8 in the Appendix).

For ensuring *convergent validity* as an aspect of *construct validity*, composite reliability has to be above .7 (Nunnally 1978) and the Average Variance Extracted (AVE) should be above .5 (Chin 1998). Table 9 in the Appendix shows that all of our constructs fulfill this requirement. Finally, *discriminant validity* represents the extent to which the items of a latent variable differ from items of other latent variables in the same model (Hulland 1999). As Table 10 in the Appendix shows all inter-correlations between the latent variables are lower than the square root of the AVE (shown in the shaded cells).

The presence of *common method bias* was tested both by Harman’s single factor test and by integrating a common method factor into a PLS model consisting of all constructs only being related to their items and to the common method factor, as suggested by (Podsakoff et al. 2003) and e.g. executed by (Liang et al. 2007). The largest component of a CFA, capturing all indicators, explained 30.7% of the overall variance. Since we have only three basic constructs, this is acceptably low. Moreover, the common method factor did not count for substantial explanations of the theoretical constructs (most paths in the PLS model even showed to be completely insignificant). Finally, we added two theoretically unrelated marker variables which showed no significant correlation with any indicator used.

### 4.2 Examining the Relationship between Trust and Control (H1)

We tested hypotheses 1a-c, representing the interrelationship between control and trust within IT outsourcing arrangements, both by a multivariate regression analysis (TRUST = β₁ * CONTROL_REP + β₂ * CONTROL_BANK + β₃ * CONTROL_PROV) and a group comparison approach.

<table>
<thead>
<tr>
<th></th>
<th>CONTROL_REP</th>
<th>CONTROL_BANK</th>
<th>CONTROL_PROV</th>
<th>R² (R² adjusted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>standardized β</td>
<td>.502*</td>
<td>-.127</td>
<td>.207*</td>
<td>.280 (.265)</td>
</tr>
</tbody>
</table>

*Table 2. Regression results (* = significant at p<.01)*

While the linear regression test showed a significantly positive relationship between output control (by reports and provider-initiated control) and trust, and no significant (linear) relationship between behavioral control and trust (Table 2), we chose a group comparison approach in order to test for a concave relationship (H1c) or at least a tendency towards it. For each control measure, we ordered the banks by the respective level of the control measure and split the sample into three groups (with the split being at the terciles, leading to sub-groups consisting of 52 data sets each). For each of the three control measures, we compared the level of trust between the three sub-groups specified on the level of control (Mann-Whitney test). A comparison of the mean values of trust based on the control values separated into three groups may help to uncover a concave relationship (i.e., moderate controls lead to a high trust level, while less or (too) much control leads to a lower trust level).

<table>
<thead>
<tr>
<th></th>
<th>Low control group (n=52)</th>
<th>Medium control group (n=52)</th>
<th>High control group (n=52)</th>
<th>Level of sig. low control group – medium control group</th>
<th>Level of sig. medium control group - high control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>control = CONTROL_REP</td>
<td>2.01</td>
<td>2.44</td>
<td>2.89</td>
<td>.026</td>
<td>.002</td>
</tr>
<tr>
<td>control = CONTROL_BANK</td>
<td>2.21</td>
<td>2.51</td>
<td>2.61</td>
<td>.168</td>
<td>.484</td>
</tr>
<tr>
<td>control = CONTROL_PROV</td>
<td>2.29</td>
<td>2.28</td>
<td>2.77</td>
<td>.902</td>
<td>.026</td>
</tr>
</tbody>
</table>

*Table 3. Mean values of control dimensions related to level of trust and significance of difference (based on Mann-Whitney test)*
As the data in Table 3 show, the mean values in all dimensions and for the overall control construct (combining all seven control indicators) increase with the level of control. Although not all differences between the control subgroups in each control dimension are significant, a concave relationship can be screened out. A more detailed analysis within the control subgroups by again dividing the high-control group into three equally large sub-groups (again based on the level of control) did not reveal any additional insights.

Thus, we can partially accept H1b and will discard H1a+c. Moreover, detailed results show that (1) the more detailed and customer-comprehensible the performance reports generated by the IT service provider, the higher is the level of trust by the bank; (2) the provider’s activities in conducting employee satisfaction surveys at the client as well as in providing evidence for IT trainings of its own employees at least partly increase the level of trust reported by the bank; (3) behavioral control by the bank is slightly positively but insignificantly related with the level of trust.

4.3 Testing the Role of Service Quality as a Differentiator (H2)

Hypothesis 2 proposes the role of service quality to affect the relationship between control and trust. The higher the service quality, the more complementary rather than substitutive would the relationship between control and trust be. Since the described cycle of trust or distrust represents a temporal and furthermore reciprocal phenomenon, we cannot sufficiently validate it by cross-sectional data. Nevertheless, we tried different approaches in order to get at least indicative evidence for its existence.

First, we did a simple group comparison between outsourcing relationships showing high vs. low service quality. Since we measured service quality along two dimensions, we formed different groups based on both measures, as well. In order to achieve discriminant validity, we defined the two groups based on the lowest and highest third, omitting the mid-level third. Following our hypothesis, we would expect a negative relationship between trust and control in the low-SQ third while a positive relationship should appear in the high-SQ third. The following table shows the regression results for trust = β₀ + β₁ * CONTROL_REP + β₂ * CONTROL_PROV + β₃ * CONTROL_BANK + residual.

<table>
<thead>
<tr>
<th>Standardized regression coeff.</th>
<th>Low level</th>
<th>High level</th>
<th>Low level</th>
<th>High level</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTROL_REP (β₁)</td>
<td>.616***</td>
<td>.141</td>
<td>.442***</td>
<td>.429***</td>
</tr>
<tr>
<td>CONTROL_PROV (β₂)</td>
<td>.103</td>
<td>.173</td>
<td>.170</td>
<td>.251*</td>
</tr>
<tr>
<td>CONTROL_BANK (β₃)</td>
<td>-.255*</td>
<td>.019</td>
<td>-.157</td>
<td>-.312**</td>
</tr>
</tbody>
</table>

Table 4. Relationship between control and trust in low vs. high SQ groups (levels of significance: ***: p<.01, **: p<.05, *: p<.1)

The proposed situation appears only in one case (CONTROL_BANK, SQ-reliability). In two other cases (CONTROL_PROV), at least the impact shows an insignificant trend in the expected direction, thus being higher in the high-SQ group than in the low-SQ group. However, in half of the combinations, a reverse result is detected.

Second, we apply a moderator test based on hierarchical regression consisting of the three CONTROL variables, the two SQ variables, and the six resulting interaction terms. Confirming the hypothesis would require positive interaction effects. The results are presented in Table 5.

Again, the analysis uncovers some contra-intuitive results: some of the interaction effects are significantly or insignificantly negative instead of being positive. Moreover, graphical data analyses showed that at high levels of SQ, the relationship between control and trust becomes actually negative (not displayed).
Standardized regression coefficients

<table>
<thead>
<tr>
<th>Standardized regression coefficients</th>
<th>Model 1 (only CONTROL)</th>
<th>Model 2 (only SQ)</th>
<th>Model 3 (all main effects)</th>
<th>Model 4 (including interaction terms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTROL_REP</td>
<td>.502***</td>
<td>.337***</td>
<td>.366***</td>
<td></td>
</tr>
<tr>
<td>CONTROL_BANK</td>
<td>-.127</td>
<td>-.166**</td>
<td>-.214***</td>
<td></td>
</tr>
<tr>
<td>CONTROL_PROV</td>
<td>.206***</td>
<td>.190***</td>
<td>.162**</td>
<td></td>
</tr>
<tr>
<td>SQ_reliability</td>
<td></td>
<td>.278***</td>
<td>.189**</td>
<td>.191**</td>
</tr>
<tr>
<td>SQ_responsiveness</td>
<td></td>
<td>.353***</td>
<td>.296***</td>
<td>.230***</td>
</tr>
<tr>
<td>CONTROL_REP * SQ_reliability</td>
<td></td>
<td></td>
<td></td>
<td>-.470***</td>
</tr>
<tr>
<td>CONTROL_BANK * SQ_reliability</td>
<td></td>
<td></td>
<td></td>
<td>.158*</td>
</tr>
<tr>
<td>CONTROL_PROV * SQ_reliability</td>
<td></td>
<td></td>
<td></td>
<td>.210**</td>
</tr>
<tr>
<td>CONTROL_REP * SQ_responsiveness</td>
<td></td>
<td></td>
<td></td>
<td>.090</td>
</tr>
<tr>
<td>CONTROL_BANK * SQ_responsiveness</td>
<td></td>
<td></td>
<td></td>
<td>-.032</td>
</tr>
<tr>
<td>CONTROL_PROV * SQ_responsiveness</td>
<td></td>
<td></td>
<td></td>
<td>-.136</td>
</tr>
<tr>
<td>R² (R² adjusted)</td>
<td>279 (.265)</td>
<td>318 (.309)</td>
<td>433 (.414)</td>
<td>.513 (.475)</td>
</tr>
<tr>
<td>F</td>
<td>19.628***</td>
<td>35.696***</td>
<td>22.924***</td>
<td>.13.769***</td>
</tr>
</tbody>
</table>

Table 5. Moderator tests for service quality moderating the impact of control on trust (levels of significance: ***: p<.01, **: p<.05, *: p<.1)

Finally, we combined the first and the second approach. We did a similar moderator test with a reversed regression function, now using SQ as dependent variable and the interaction of control and trust as determinant. These regression tests were conducted within the low and the high SQ groups instead of on the overall sample. Confirming H2 would be reflected by a complementary relationship (i.e. a positive interaction effect) in the high-SQ sub-sample, and a substitutive relationship (negative interaction effect) in the low-SQ sub-sample.

Because we found multicollinearity problems during this analysis (VIFs up to 50), we followed the regression approach proposed by (Lance 1988) and successfully applied in other studies (e.g. (Prabhu et al. 2005; Tiwana et al. 2007)): we tested the following regression function: \( SQ = \beta_0 + \beta_1 * CONTROL + \beta_2 * TRUST + \beta_3 * dXX + residual \)

This approach prevents inflation of the regression coefficients resulting from multicollinearity. We did this procedure for all combinations of \( CONTROL = \{rep, prov, bank\} \) and \( SQ = \{reliability, responsiveness\} \) within the low and high service quality sub groups. The following table shows the results.

<table>
<thead>
<tr>
<th>Standardized regression coeff. and sig. level of interaction term</th>
<th>SQ-reliability</th>
<th>SQ-responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTROL_prov * TRUST</td>
<td>-.007,.946</td>
<td>.189,.089</td>
</tr>
<tr>
<td>CONTROL_rep * TRUST</td>
<td>-.242,.019</td>
<td>-.143,.190</td>
</tr>
<tr>
<td>CONTROL_bank * TRUST</td>
<td>-.202,.052</td>
<td>.082,.734</td>
</tr>
<tr>
<td>N (sample size)</td>
<td>76</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 6. Interaction tests for control and trust showing different effects on SQ at different levels of SQ

---

1 In order to increase reliability during the regression analyses, we did not choose the terciles with low and with high service quality as group separators but used the lower and the higher half in order to have larger sub-samples for the regression calculations. Nevertheless, using the terciles does not lead to structural differences in the results displayed by Table 6; path coefficient differences show to be even larger, but the smaller samples do not enable successful tests of significance.
Interestingly, now a major proportion of the results supports our proposition. Although the path coefficients are mainly insignificant, due to the small data samples, the differences between both groups are significant in all cases (based on a bootstrapping with 500 samples). Comparing the groups, there is usually a negative relationship in the group with low SQ, while it is positive or at least less negative in the high SQ group.

5 DISCUSSION AND LIMITATIONS

While many of the conceptual and few of the empirical works (Das and Teng 1998; Das and Teng 2001; Grundei 2006; Kale et al. 2000; Langfield-Smith and Smith 2003) argue for a negative relationship between control and trust, our work found mainly a positive relationship. This can at least partially be ascribed to the characteristics of our control items. Particularly, the provider-based control items, such as employee surveys or proving IT training certificates, are more or less directly reasoned in promoting a good relationship. Thus, this dimension is very likely to lead to a positive relationship.

Further, the control reports construct was measured in a more qualitative way, which also will contribute to a positive relationship. By contrast, the bank-driven control showed no correlation with the level of trust in the relationship and rather turned into a negative path when service quality was introduced as a moderator. A reason that can be directly extracted from the different service quality dimensions is that bank-driven control does not necessarily drive reliability (that would be the most obvious relation) but rather the responsiveness by the provider, which usually indicates commitment and in turn significantly drives trust in the relationship. The negative residual direct relationship then can be explained by reversing the link. Above, we discussed the reciprocal relation between trust and control, which can be uncovered here. The negative residual path simply indicates the reverse argument that a less trustful relationship will lead to more control actions by the bank itself. Nevertheless, most of our results quite strongly and impressively showed that there are no significantly negative relations between control and trust. This indicates that there is no over-control in our sample and that the finance industry and their IT providers are used to high levels of control being a natural and “healthy” part of their business.

Interestingly, there is no correlation between bank-driven control and service reliability, and moreover, provider-driven control in both models has a significant positive effect on trust, but is not related to higher service quality. This again underlines that the primary reasoning may lie in maintaining the relationship by these issues, rather than actually increasing operational service quality.

Contra-intuitively, the basic tests of service quality moderating the relationship between control and trust showed to be negative instead of being positive, as proposed. This represents a substitutive relationship between service quality and control regarding the achievement of a trustful relationship. In case of high service quality, the need and impact of control is less than otherwise. Nevertheless, in case of high service quality, control, at least as measured by the control action items in this paper, can even better interact with the creation of trust in order to achieve a successful outsourcing partnership (cycle of trust, as indicatively confirmed at the end of section 4.3).

This leads to our implications for management: our results clearly showed the positive impact of several control action items, both the provider and the client firm can apply. Providers that set up mechanisms that fulfill social control (such as surveying satisfaction of client employees or proving training certificates for own personnel) will increase and maintain trust. Nevertheless, the client should be aware that these measures will eventually turn into an increase in service quality and not only in relationship quality (dazzling effect). Further, the client has to be aware of the cycle of distrust. Good relationship quality which is oriented towards a long-term partnership requires sensitivity and sometimes forbearance in order to support the provider during difficult situations. Otherwise, control will just worsen and maybe destroy the outsourcing relationship with high costs on both sides.

Of course, there are a number of limitations that have to be considered when interpreting the results: (1) we investigated IT outsourcing relationships restricted to the German Banking Industry; (2) our findings are based on data from only one point in time, and collected from one person in each bank,
thus inhibiting the assessment of time effects (and, the cycle of trust/distrust, in particular) and incorporating subjectivity; (3) we are aware that parts of our measurement model may be subject to being too tightly bound to the application domain. Although this might be a weakness for rigorously testing theory on the relationship of trust and control, it will improve the trade-off between the results providing valid findings while being directly applicable by managers who maintain outsourcing relationships. Also, through adopting indicators from prior research and in-depth case studies, we can assume our measurement model to be appropriate; (4) we have not tested for any contingency factors, yet. Besides e.g. firm size or strategy, in particular the duration of the relationship could have an impact. Although we suppose the relationship age not to have an effect in our data since, as shown above, most arrangements have lasted more than ten years, further investigations will include such tests. Also, we concentrated on one specific information system to avoid noteworthy contingency effects, which could e.g. result when different outsourcing scenarios (e.g. characterized by various degrees of complexity) are examined and thus trust is more or less important with respect to distinct scenarios; (5) trust has only been measured as relational trust, neglecting other dimensions such as competence trust, in particular.

6 CONCLUSION

Based on data of IT outsourcing relationships between 156 German banks and their IT service providers, and on the literature that has dealt with the coherence of control and trust in inter-organizational collaboration, this paper investigated the relationship between control and trust, and also included service quality as a moderator. We could show that different types of control do differently affect the level of trust in IT outsourcing arrangements. While control reports and provider-driven controls proved to have a positive effect, bank-driven controls did not correlate with trust and furthermore showed a negative effect when considering the level of service quality. Quite surprisingly, the interaction effects of service quality and control turned out to be negatively related with trust. However, comparing high-service quality situations with those showing low service quality, we could indicatively validate the hypothesis that there may exist cycles of trust/distrust. Finally, by splitting control into different dimensions, we did a first step to generating more concrete implications for provider managers that can help determining the right level of control.

In a next step, we will incorporate different contingency variables and a richer concept of trust, as noted in the limitations, in order to take the true complexity of outsourcing relationships better into account. Moreover, we intend to collect periodical data from selected firms in order to get longitudinal data for more adequately testing the proposed inter-temporal effects. Ultimately, we hope to contribute both to a better scientific understanding of the complexities between control and trust and to a statistically validated managerial set of control action items which helps outsourcing relationship managers to achieve and maintain sustainable and successful IT outsourcing partnerships.

References


Proceedings ECIS 2009


## Appendix

<table>
<thead>
<tr>
<th>ID</th>
<th>Item (measured by either a five-point or a seven-point Likert Scale ranging from “fully disagree” to “fully agree”)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTROL_REP1</td>
<td>The granularity of the service reports meets our demands.</td>
<td>Development of own indicators, derived from case studies (Beimborn et al. 2008) and chosen to better meet the practitioner requirements.</td>
</tr>
<tr>
<td>CONTROL_REP2</td>
<td>The provider’s service reports offer an accurate picture of the provider’s services.</td>
<td></td>
</tr>
<tr>
<td>CONTROL_BANK1</td>
<td>We can directly access service evaluation measurements to avoid sugarcoated reports.</td>
<td></td>
</tr>
<tr>
<td>CONTROL_BANK2</td>
<td>We use our own system to obtain detailed information on service quality.</td>
<td></td>
</tr>
<tr>
<td>CONTROL_BANK3</td>
<td>We regularly conduct service satisfaction surveys among users.</td>
<td></td>
</tr>
<tr>
<td>CONTROL_PROV1</td>
<td>The service provider regularly conducts service satisfaction surveys among our users.</td>
<td></td>
</tr>
<tr>
<td>CONTROL_PROV2</td>
<td>The service provider regularly shows IT training results of their employees.</td>
<td></td>
</tr>
<tr>
<td>SQ_reliability1</td>
<td>Problems are resolved reliably.</td>
<td>(Grover et al. 1996; Kettinger and Lee 1995; Lee and Kim 1999; Parasuraman et al. 1988)</td>
</tr>
<tr>
<td>SQ_reliability2</td>
<td>Applications and services are provided as promised.</td>
<td></td>
</tr>
<tr>
<td>SQ_reliability3</td>
<td>There are never any critical system failures.</td>
<td></td>
</tr>
<tr>
<td>SQ_responsiveness1</td>
<td>The service provider reacts quickly if there are problems.</td>
<td>(Kettinger and Lee 1995; Parasuraman et al. 1988)</td>
</tr>
<tr>
<td>SQ_responsiveness2</td>
<td>The service provider shows adequate readiness to respond to our requests.</td>
<td></td>
</tr>
<tr>
<td>SQ_responsiveness3</td>
<td>Provider staff has a service-oriented attitude.</td>
<td></td>
</tr>
<tr>
<td>TRUST1</td>
<td>Both parties in the relationship can be trusted to do business fairly.</td>
<td>(Goles and Chin 2005; Kettinger and Lee 1995; Parasuraman et al. 1988)</td>
</tr>
<tr>
<td>TRUST2</td>
<td>We trust that the reports of the service provider are correct.</td>
<td></td>
</tr>
<tr>
<td>TRUST3</td>
<td>We trust that the cost estimates of the service provider are correct.</td>
<td></td>
</tr>
<tr>
<td>TRUST4</td>
<td>Our service provider acts in our best interests.</td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Used indicators

<table>
<thead>
<tr>
<th>Construct</th>
<th>Indicator</th>
<th>Loading from PLS</th>
<th>Loading from CFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTROL_REP</td>
<td>CONTROL_REP1</td>
<td>0.915</td>
<td>0.916</td>
</tr>
<tr>
<td></td>
<td>CONTROL_REP2</td>
<td>0.921</td>
<td>0.916</td>
</tr>
<tr>
<td>CONTROL_BANK</td>
<td>CONTROL_BANK1</td>
<td>0.829</td>
<td>0.790</td>
</tr>
<tr>
<td></td>
<td>CONTROL_BANK2</td>
<td>0.641</td>
<td>0.807</td>
</tr>
<tr>
<td></td>
<td>CONTROL_BANK3</td>
<td>0.812</td>
<td>0.749</td>
</tr>
<tr>
<td>CONTROL_PROV</td>
<td>CONTROL_PROV1</td>
<td>0.636</td>
<td>0.775</td>
</tr>
<tr>
<td></td>
<td>CONTROL_PROV2</td>
<td>0.885</td>
<td>0.775</td>
</tr>
<tr>
<td>SQ_reliability</td>
<td>SQ_reliability1</td>
<td>0.817</td>
<td>0.807</td>
</tr>
<tr>
<td></td>
<td>SQ_reliability2</td>
<td>0.763</td>
<td>0.784</td>
</tr>
<tr>
<td></td>
<td>SQ_reliability3</td>
<td>0.774</td>
<td>0.781</td>
</tr>
<tr>
<td>SQ_responsiveness</td>
<td>SQ_responsiveness1</td>
<td>0.818</td>
<td>0.823</td>
</tr>
<tr>
<td></td>
<td>SQ_responsiveness2</td>
<td>0.782</td>
<td>0.802</td>
</tr>
<tr>
<td></td>
<td>SQ_responsiveness3</td>
<td>0.800</td>
<td>0.774</td>
</tr>
<tr>
<td>TRUST</td>
<td>TRUST1</td>
<td>0.837</td>
<td>0.841</td>
</tr>
<tr>
<td></td>
<td>TRUST2</td>
<td>0.814</td>
<td>0.820</td>
</tr>
<tr>
<td></td>
<td>TRUST3</td>
<td>0.816</td>
<td>0.828</td>
</tr>
<tr>
<td></td>
<td>TRUST4</td>
<td>0.781</td>
<td>0.765</td>
</tr>
</tbody>
</table>

Table 8. Indicator loadings from PLS and Confirmatory Factor Analysis (note: all loadings were significant at a .001 level)
<table>
<thead>
<tr>
<th>Construct</th>
<th>Composite Reliability</th>
<th>AVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTROL_REP</td>
<td>0.915</td>
<td>0.843</td>
</tr>
<tr>
<td>CONTROL_BANK</td>
<td>0.807</td>
<td>0.856</td>
</tr>
<tr>
<td>CONTROL_PROV</td>
<td>0.740</td>
<td>0.594</td>
</tr>
<tr>
<td>SQ_reliability</td>
<td>0.828</td>
<td>0.617</td>
</tr>
<tr>
<td>SQ_responsiveness</td>
<td>0.842</td>
<td>0.640</td>
</tr>
<tr>
<td>TRUST</td>
<td>0.886</td>
<td>0.660</td>
</tr>
</tbody>
</table>

Table 9: Quality measures for latent variables

<table>
<thead>
<tr>
<th>Construct</th>
<th>CTL_REP</th>
<th>CTL_BANK</th>
<th>CTL_PROV</th>
<th>SQ_REL</th>
<th>SQ_RES</th>
<th>TRUST</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTROL_REP1</td>
<td>0.915</td>
<td></td>
<td></td>
<td>0.357</td>
<td>0.336</td>
<td>0.441</td>
</tr>
<tr>
<td>CONTROL_REP2</td>
<td>0.921</td>
<td>0.416</td>
<td>0.170</td>
<td>0.346</td>
<td>0.359</td>
<td>0.465</td>
</tr>
<tr>
<td>CONTROL_BANK1</td>
<td>0.465</td>
<td>0.829</td>
<td>0.275</td>
<td>0.181</td>
<td>0.261</td>
<td>0.163</td>
</tr>
<tr>
<td>CONTROL_BANK2</td>
<td>0.294</td>
<td>0.641</td>
<td>0.284</td>
<td>0.044</td>
<td>0.100</td>
<td>0.030</td>
</tr>
<tr>
<td>CONTROL_BANK3</td>
<td>0.231</td>
<td>0.812</td>
<td>0.344</td>
<td>0.199</td>
<td>0.228</td>
<td>0.173</td>
</tr>
<tr>
<td>CONTROL_PROV1</td>
<td>0.085</td>
<td>0.194</td>
<td>0.636</td>
<td>0.096</td>
<td>0.052</td>
<td>0.153</td>
</tr>
<tr>
<td>CONTROL_PROV2</td>
<td>0.220</td>
<td>0.369</td>
<td>0.885</td>
<td>0.113</td>
<td>0.144</td>
<td>0.246</td>
</tr>
<tr>
<td>SQ_reliability1</td>
<td>0.258</td>
<td>0.120</td>
<td>0.142</td>
<td>0.817</td>
<td>0.621</td>
<td>0.465</td>
</tr>
<tr>
<td>SQ_reliability2</td>
<td>0.313</td>
<td>0.127</td>
<td>0.078</td>
<td>0.763</td>
<td>0.418</td>
<td>0.335</td>
</tr>
<tr>
<td>SQ_reliability3</td>
<td>0.337</td>
<td>0.258</td>
<td>0.093</td>
<td>0.774</td>
<td>0.346</td>
<td>0.345</td>
</tr>
<tr>
<td>SQ_responsiveness1</td>
<td>0.327</td>
<td>0.197</td>
<td>0.070</td>
<td>0.517</td>
<td>0.818</td>
<td>0.444</td>
</tr>
<tr>
<td>SQ_responsiveness2</td>
<td>0.311</td>
<td>0.273</td>
<td>0.147</td>
<td>0.402</td>
<td>0.782</td>
<td>0.341</td>
</tr>
<tr>
<td>SQ_responsiveness3</td>
<td>0.274</td>
<td>0.216</td>
<td>0.120</td>
<td>0.500</td>
<td>0.800</td>
<td>0.464</td>
</tr>
<tr>
<td>TRUST1</td>
<td>0.347</td>
<td>0.131</td>
<td>0.197</td>
<td>0.358</td>
<td>0.469</td>
<td>0.837</td>
</tr>
<tr>
<td>TRUST2</td>
<td>0.406</td>
<td>0.054</td>
<td>0.185</td>
<td>0.451</td>
<td>0.395</td>
<td>0.814</td>
</tr>
<tr>
<td>TRUST3</td>
<td>0.418</td>
<td>0.156</td>
<td>0.222</td>
<td>0.376</td>
<td>0.326</td>
<td>0.816</td>
</tr>
<tr>
<td>TRUST4</td>
<td>0.428</td>
<td>0.264</td>
<td>0.260</td>
<td>0.399</td>
<td>0.499</td>
<td>0.781</td>
</tr>
</tbody>
</table>

Table 10: Square root of AVE (shaded cells) and correlations of latent variable scores

Table 11: Correlations of indicator scores
THE ICT CONVERGENCE DISCOURSE IN THE INFORMATION SYSTEMS LITERATURE – A SECOND-ORDER OBSERVATION

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0655.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Convergence, Literature review, Discourse analysis, Ubiquitous systems</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
THE ICT CONVERGENCE DISCOURSE IN THE INFORMATION SYSTEMS LITERATURE – A SECOND-ORDER OBSERVATION

Herzhoff, Jan, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK, j.d.herzhoff@lse.ac.uk

Abstract

The idea of ICT Convergence is used by many practitioners and observers - such as economists, politicians, journalists, and academics - as an important descriptor for technological change. However, a review of previous work in this field suggests that, despite more than 30 years of research on ICT Convergence, the theoretical basis of the concept of convergence is still under-researched. In particular in the IS literature, the concept has been either relegated to the sidelines or taken for granted without further reflection. Therefore, a systematic analysis of the idea of ICT Convergence from an IS perspective is needed.

This paper aims to explore how the discourse of convergence is being shaped in the IS literature. In order to address this question, 317 articles published in ten leading IS journals from 1998 to 2008 have been examined. This study has been built around a Grounded Theory approach informed by Niklas Luhmann’s Theory of Distinction.

The findings show that convergence cannot be viewed as a single concept. Five archetypes of “convergence communication” are identified, and a conceptualization of ICT Convergence as a double process between alignment and interoperability is suggested. The main limitation of this paper is the focus on leading IS journals.

Keywords: Convergence, Luhmann, Theory of Distinction, Literature Review, Discourse Analysis
1 INTRODUCTION

"Convergence may be the most expensive word in history. It has cost people billions." —David Geffen, co-founder of DreamWorks, 2002

The idea of the convergence of information and communication technologies (ICT Convergence) has been accorded tremendous importance by practitioners and academics alike when describing the process of technological change through digitalisation. Many organizations have used the concept to justify large investments (Lind, 2004) or to solicit funds for entrepreneurial endeavours (Knox, 2003a). But convergence is also seen as one of the driving forces for the development of new information infrastructures and services, and is therefore important to understand from a systems design perspective (Lyytinen & Yoo, 2002).

ICT Convergence is not a new concept. It has been intensively discussed in the academic literature for more than 30 years and has established itself in all major disciplines adjacent to information systems, such as computer science, management, new media, or economics. However, despite the apparent maturity of this research area (Lind, 2004; Farber and Baran, 1977), there is no generally accepted definition of ICT Convergence (Nyström, 2007); instead Appelgren (2004:246) argues that "there seem to be as many definitions of convergence as there are authors discussing the topic".

Furthermore, many authors agree that the idea of convergence lacks a systematic theoretical analysis (see Hacklin, 2007; Nyström, 2007). The diffusion of new converged services like IPTV (the convergence of TV and the Internet) or mobile Internet (the convergence of mobile telephony networks and the Internet) is still very slow in many countries. It is, therefore, crucial to provide a more systematic analysis and more robust theoretical underpinning of this important phenomenon in order to inform future design and regulatory recommendations. This paper argues that because of its close link to technological change, it is highly relevant for the information systems community to engage in this multidisciplinary debate and to offer an own perspective.

Schütz (1962) suggests distinguishing between first-level and second-level constructs. A theory can be built explaining whether an observed phenomenon is a “war dance, a bartender trade, or the reception of a friendly ambassador” (Schütz, 1962 p. 54); or it can be investigated how the observed persons understand the dance. In Luhmann’s terms (2002), focusing less on the constructs than on the process of observation, either the dancers themselves (first-order observation) or how the dancers themselves understand the dance (second-order observation) can be observed. As this study is interested in how IS scholars describe convergence, the primary focus is therefore not the phenomenon of convergence itself; but on the perspective of a second-order observer (Luhmann, 2002). Therefore, at this point a definition of ICT Convergence will not be committed to upfront. This definition should instead emerge from second-order observation.

The purpose of this paper is, therefore, to provide a taxonomy based on a systematic second-order analysis of how the idea of convergence has been constructed and used in the IS literature. Based on the taxonomy, this study aims to analyse in particular what constitutes a particular form of convergence communication, namely "ICT Convergence", from an IS perspective.

The organization of this paper is as follows: The literature review provides an overview about the existing literature on ICT Convergence from both first and second-order perspectives. Next, the paper presents its analytical approach based on the work by Niklas Luhmann and gives a brief account on the applied methodology. Finally, five archetypes of convergence communication in the IS literature will be presented and analyzed based on Luhmann's form and differentiation analysis.

The research scope of this paper is limited to the conceptualisation of convergence communication, in particular ICT Convergence communication, in ten leading IS journals and therefore disregards lower-ranked journals, conference papers and books. However, an initial review of other related IS journals and conference papers has not indicated any significant differences to the results presented in this paper.
2 LITERATURE REVIEW

The following literature review draws the distinction between first-order and second-order research on ICT Convergence. The first-order research deals directly with the underlying phenomenon of convergence, its drivers and consequences. The second-order research, on the other hand, deals with how the idea of convergence has been constructed by the observers of the phenomenon.

2.1 First-order analysis of ICT Convergence

ICT Convergence has been studied from many different perspectives and may be seen as a truly multidisciplinary research topic. This review focuses on the key works in the fields of management, computer science, new media, and information systems.

Computer Science Perspective: The meaning of ICT Convergence, in the context of digitalization, has already been envisioned by H. E. Vaughan of Bell Labs in 1959 as integrated communications that can provide flexibility for new services. There are manifold debates on convergence from the perspective of computer science. Messerschmidt (1996) identifies nine key debates related to convergence in the Computer Science literature: (1) best effort versus Quality of Service (QoS), (2) scalability, (3) terminal and network coordination, (4) connection versus wireless, (5) control architecture, (6) interconnection versus interoperability, (7) embedded computing versus general-purpose computing, (8) heterogeneity, and (9) architecture and complexity management. In particular, the design issues pertaining to increased conflicts or “tussles” through convergence and how to better control networks are of primary concern (see, e.g., Clark et al., 2005).

Management Perspective: The term “technical convergence” was coined four years later in 1963 by Rosenberg, who studied the American machine tool industry of the 18th century. According to Hacklin (2007), the key debates revolve around how an organization should respond to convergence, which may be viewed in the larger context of how organizations respond to innovations or market disruptions not originated by them. The debates have proceeded on two levels of analysis: at the firm level and at the industry-level. The management literature, drawing upon Rosenberg, began to incorporate the concept of convergence into the study of strategic management, building analytical tools and offering recommendations on how companies should react to technical convergence (Pennings & Puranam, 2001). Researchers also debate how to conceptualise convergence itself. Is it an endpoint or a process? If it is a process, how can this process be described? However, only a few scholars have attempted to describe convergence and craft a theory. One of the first serious attempts to do so from a management perspective was undertaken by Greenstein and Khanna (1997)—who distinguish between convergence of substitutes, which entails one domain competing with another, and complementary convergence, which entails two different fields coming together. Stieglitz (2002) refines the model by introducing a second dimension, that of product orientation versus technology orientation. However, each of these conceptualisations focuses only on industry-convergence, taking technological aspects for granted. The latest major work on convergence from a management perspective was conducted by Hacklin (2007), who interprets convergence as a species of technological change. He suggests analysing it as a process that originates in convergences of knowledge, technology, and applications, leading eventually to industrial convergence.

New Media Perspective: Convergence is one of the many terms proposed to describe technological change through digitalisation in the media literature. Other competing terms offered in the 1970s include "comunications" (A. Oettinger) and "telematique" (Nora & Minc). However, "convergence" eventually triumphed as the dominant label for this form of technological change, and became the term to be deployed in both the management literature and the popular media. It was further popularised by Nicolas Negroponte’s famous figure of three overlapping circles, a highly static conceptualization that assumes that convergence is an end-stage. Ithiel de Sola Pool's book Technologies of Freedom (1983) uses the term to describe the convergence of different modes of communication. Henry Jenkins (2001) built upon de Sola Pool’s findings and systematises the concept of convergence by splitting it into five...
different processes: namely, technical and economic convergence, which together lead to global, cultural, and organic convergence. Jenkins (2006) argues in particular for a wider understanding of convergence and coined the term “convergence culture”.

**IS Perspective:** Only a few articles in the core IS literature deal with the concept of convergence explicitly—all, without exception, on mobile computing and information infrastructures. Lyytinen and Yoo (2002) briefly address convergence in their discussion of the drivers of a nomadic information environment. They see convergence, mobility, and mass scale as driving development for both information infrastructures and services. Convergence is defined as the “digital processing of all forms of data […] across different carriers […] with multiple devices….” (p. 379). Finally, they note that open standards are essential to convergence. Nielsen (2004) sees convergence as a “process bringing together different and heterogeneous actors as well as markets and technologies, a process not only bringing synergies but also challenges.” He argues that in particular conflicting interests might emerge from these convergence processes.

Jansen and Nielsen (2005) investigate the convergence of UMTS and WiFi infrastructures in Norway. They suggest conceptualising convergence as a form of co-evolution, pointing out that the convergence of these two infrastructures is by no means inevitable. Tilson (2008) addresses several convergence instances in the mobile and television industry. He uses actor-network-theory to ”explain convergence, the explosion in the number of interfaces requiring standardization, and other industry and standardization changes observed in the case studies” (p. 17).

### 2.2 Second-order analysis

Several studies over the last five years discuss convergence from a second-order perspective, i.e. analyzing or reflecting on the idea of convergence instead of describing the underlying phenomenon. Knox (2003b) analyses the idea of convergence from an anthropological perspective, considering how it has been mobilised in the development of new media in Manchester. She points out that convergence is used by both observers of new media (economists, academics, politicians, civil servants) and its practitioners (Knox, 2003a). Further, Knox argues that the articulation of the term often manifests a ”calculated performance" in which ”the lack of experience by these companies is self-replicated in as much as they are required to seek out novelty and new ways of working” (p. 47). Furthermore, she suggests seeing convergence not only as a singular description of a process but rather as a descriptor of change “which has gained its predominance from the fact that it cannot be pinned down to a single process, a single model, from the fact therefore of its own reproduction” (p. 120).

In a study of the usage of the term convergence in published newspaper articles between 1990 and 2004, Lind (2004) finds that the idea of convergence is often used to justify mergers and acquisitions, and also to flag impending change early in the redefinition of a market. In a study of how business and IT managers in the Finnish telecommunications and media sector perceive convergence, Nyström (2007) stresses the need for a better description of what convergence is and what it is not. Hacklin (2007) argues for the importance of a better understanding of how the term is used in the literature.

The main shortcoming of the existing body of second-order literature on ICT Convergence is, with the exception of the work by Knox (2003b) and Nyström (2008), that it has not moved further below the surface to enrich the understanding of the idea of convergence. Since the most prominent meaning of convergence has emerged in information and communication technologies (Hacklin, 2007), it is fruitful to engage in this discussion, particularly in relation to information systems.

### 2.3 Key findings

The review shows that most first-order observations in various academic disciplines take the concept of ICT Convergence for granted. Furthermore, the IS literature lacks a systematic second-order analysis of ICT Convergence. Finally, there is evidence that the concept has been diluted which makes an analysis of the distinctions of what constitutes convergence and what does not necessary.
Therefore, the expected theoretical contributions of this paper are threefold. First, this research aims to provide an initial second-order taxonomy on convergence from an IS perspective. The paper aims to show that convergence has been used in information systems in several different contexts, but only very superficially in the context of technological change. Second, this paper expects to improve the theoretical understanding of ICT Convergence from an information systems perspective. Third, it aims to contribute second-order observations to the debate on the conceptualization of ICT Convergence and also to push the analysis deeper than previous studies.

3 ANALYTICAL STRATEGIES – THEORIES OF DISTINCTION

One interesting aspect of all previous attempts to investigate the idea of convergence from a second-order perspective (with the exception of Hannah Knox’s approach using Callon’s analysis of framing and externality) is the lack of an analytical strategy as a guide for data collection and analysis. Andersen (1999) uses the notion of analytical strategies to describe the aspects involved in making an analysis from a constructivist point of departure. By doing so, he intends to emphasize that a second-order observation is not a method to be deployed in order to get closer to the truth about an object. Instead, the social perceptions of objects are to be analyzed.

Niklas Luhmann’s General Theory of Social Systems is one of the grand theories in the social sciences (Lee, 2000). Although there are many different ways to access Luhmann, this study has chosen the Spencer-Brownian route to the Theory of Distinction (Andersen, 2003, p. 64). To draw on Spencer Brown’s Law of Form (1969), the basis of form analysis is observation, which consists of two components, distinction and indication. Whenever something is observed, a distinction is made between the inner side and the outer side by choosing or “marking” the inner side as the unit of analysis. There is always something left to be “unsaid,” which is the residual category (Demetis & Angell, 2007). According to Luhmann, every researcher has to decide how he or she will observe the object of study (Luhmann, 2002). Any choice of distinction is contingent and hence open for criticism. Luhmann (2002) suggests looking instead at how the object draws the distinction between itself and its environment. Thus, this research is interested in how the authors of the texts make their own distinctions on convergence.

The approach applied by this paper uses two types of discursive analytical strategies in the data analysis, both based on the work by Niklas Luhmann: form analysis and a sub-form of systems analysis known as differentiation analysis (Andersen, 2003). Form analysis provides the foundation of the data analysis and focuses on the following question: Which distinction allows the observer to see the environment in terms of convergence, or, more specifically, which is the unmarked side of the difference when communication indicates convergence? Form analysis is not an end in itself, but it leads to the question of how social systems cope with—or, in Luhmann’s language, “de-paradoxify”—the paradoxes upon which their communication structure is built (Andersen, 2003: 101). In the analysis of how IS researchers apply the concept of convergence, this paper also applies differentiation analysis to identify distinct forms of “convergence communication”. The guiding distinction for this analysis is similarity/difference. This paper uses differentiation analysis to observe the functional differentiation between different types of convergence to build a taxonomy of convergence.

4 METHODOLOGY

The methodology is primarily influenced by the Grounded Theory approach, which has been used successfully in previous studies (Jones, 2004; Orlikowski & Iacono, 2001). This approach to analyzing IS journal papers uses a kind of content analysis in which categories are developed solely on the basis of the findings from the data and not imposed from the outside (Agar, 1980). However, it differs from the traditional content analysis since the categories were not pre-defined but have emerged during the data analysis.
This study follows the Corbin and Strauss (1990) version of Grounded Theory, but differs from their strict approach in two distinct ways. First, the data selection is based on corpus construction (Bauer and Aarts, 2000) instead of theoretical sampling. Second, the focus is not only on the core category but also on the core distinction using Luhmann’s Theory of Distinction as additional analytical strategy in the analysis of the findings.

Corpus construction (Bauer & Aarts, 2000) has the advantage of offering a vocabulary that is independent of sampling logic and that overcomes the shortcomings of theoretical sampling, as suggested by Strauss and Corbin (1990), such as multiplication of sampling methods. The goal is to select “incidents” of a phenomenon, not to sample a population (Bauer & Aarts, 2000). Barthes (1967) suggests selecting a data corpus based on relevance, homogeneity, and synchronicity. By keeping the focus on information systems relevance was ensured whereas homogeneity of the corpus was achieved by taking only journal articles into consideration. Finally, synchronicity has been maintained by focusing on journal articles that were published between 1998 and 2008.

This study focuses only on the leading academic IS journals. Ten IS journals (MIS Quarterly, Information Systems Research, Journal of MIS, Journal of the AIS, Information & Organization, European Journal of Information Systems, Journal of Strategic Information Systems, Information Systems Journal, Information & Management, Decision Support Systems) have been drawn upon based on the ranking from Louisiana State University, since it distinguishes among management, practitioners’, and “pure” IS journals. The selected journals also appear frequently in the top ten rankings in recent studies (see e.g. Rainer & Miller, 2005 or Peffers & Tang, 2003). The decision to focus on this body of literature and to exclude other journals, conference papers, etc., is driven by relevance and resource constraints. The primary aim here is to see how the highly relevant concept of convergence is treated and used in the mainstream IS journals.

A full text search has been conducted on the term "convergence" using Business Source Premier, Sweetwise, and ScienceDirect. The search has not been limited to "ICT Convergence" to achieve a broad understanding of convergence communication and to increase the variety of the findings. The only exception for the analysis period between 1998 and 2008 is the Journal of the AIS which has only been published in 1999. This timeframe has been selected to cover an extended period, but at the same time benefits from the easy data access to conduct a full-text search. For most of the journals, no electronic versions are available before 1998. Based on the search results, a corpus of 341 journal articles has been constructed and imported into the software package Atlas.ti. 24 articles which used the term “convergence” only in the bibliography have been excluded and the final corpus comprised of 317 articles.

The approach of informing Grounded Theory with the Theory of Distinction is new, especially in the IS field, but has already proven valuable in other domains of social sciences (Gibson et al., 2005). Theory of Distinction has been used as an analytical strategy to sharpen the perception of differences in the data analysis (Andersen, 2003). It is consistent with Grounded Theory in several ways. Firstly, both are interested in the emergence of meaning and focus on what has been communicated and how it has been organized. However, they differ in one distinct aspect: traditional Grounded Theory searches for the core category, whereas the Theory of Distinction is concerned with the “guiding distinction” (Gibson et al., 2005).

5. FINDINGS

Each article has been searched for the term “convergence,” and the relevant paragraphs have been coded by repeatedly asking what it means in this instance, in what context it is used, and what distinctions are made by the authors of these articles. Based on the context codes, six categories of contexts have been identified in which the authors used convergence. Three contexts are closely related to the research process itself, and three were related to the phenomenon under study.

The findings suggest that the IS community uses the concept of convergence to describe (I) research streams and theoretical concepts coming together; (II) quality criteria in methodology sections; and
The processing of quantitative and qualitative data analysis. The topical contexts can be broadly separated into (IV) decision-making; (V) technological change; and (VI) other contexts.

Table 2 illustrates the different sub-themes and the number of occurrences in the data corpus. One interesting aspect of this data is that convergence in the context of technological change has only 48 occurrences, which account for less than 15% of the articles (total = 317). In the next step, some initial concepts have been identified with their properties and their dimensions, and the data accordingly coded. These insights and the method of constant comparisons have been used to construct a first set of categories. Based on the analysis, the five identified categories or conceptualizations of convergence are: alignment, recombination, optimization, interoperability and correspondence. The category labels are rooted in the data. In the following paragraphs, all five views on convergence identified in the course of the analysis of the data corpus are presented.

**Convergence as Alignment:** In the IS literature, the concept of alignment is primarily used in the context of decision-making. In the data corpus two sub-forms are identified: one based on building up shared models between social systems and the other one on finalizing the decision-making process. In both cases, a change in both converging elements is anticipated in order to reach some sort of consensus. IT has more of a support role in the form of group or decision-support systems. Both sub-forms have an iterative process understanding of convergence. Convergence as decision-making is intended to increase focus and efficiency whereas convergence as building up shared mental models aims towards incremental change in individual accuracy.

**Convergence as Correspondence:** Convergence as correspondence focuses on similarities among concepts and highlights correlations or equality between them. It is used in the context of aligning research findings with existing research or in the process of triangulation. The converging elements are not anticipated to change, but are conceptualised in a stable state. Therefore, convergence is not seen as a process but as an end stage. The process is folded into one dimension and is not iterative:

### Table 1: Convergence Contexts

<table>
<thead>
<tr>
<th>Context</th>
<th>Freq.</th>
<th>%</th>
<th>Freq.</th>
<th>%</th>
<th>Freq.</th>
<th>%</th>
<th>Freq.</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Research focus</td>
<td>36</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interdisciplinary</td>
<td>8</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Streams</td>
<td>12</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theories</td>
<td>10</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Findings</td>
<td>6</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II. Quality criteria</td>
<td>37</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triangulation</td>
<td>12</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validity</td>
<td>25</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III. Data analysis</td>
<td>65</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturation</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neural Networks</td>
<td>20</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genetic Algorithms</td>
<td>16</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Algorithms</td>
<td>27</td>
<td>42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV. Decision-making</td>
<td>119</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groupwork</td>
<td>52</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alignment</td>
<td>22</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decision-support</td>
<td>8</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agent systems</td>
<td>37</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V. Techn. Change</td>
<td>46</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td>9</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network</td>
<td>21</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisation</td>
<td>10</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>2</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web service</td>
<td>6</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VI. Other contexts</td>
<td>14</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Globalisation</td>
<td>11</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activities</td>
<td>2</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telemedicine</td>
<td>1</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Freq. / %</td>
<td>317</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2: Convergence Archetypes in Context**

<table>
<thead>
<tr>
<th>Context</th>
<th>Interoperability</th>
<th>Recombination</th>
<th>Optimization</th>
<th>Alignment</th>
<th>Correspondence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freq.</td>
<td>%</td>
<td>Freq.</td>
<td>%</td>
<td>Freq.</td>
<td>%</td>
</tr>
<tr>
<td>Research focus</td>
<td>-</td>
<td>-</td>
<td>11</td>
<td>65</td>
<td>-</td>
</tr>
<tr>
<td>Quality criteria</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>25</td>
</tr>
<tr>
<td>Data analysis</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>65</td>
</tr>
<tr>
<td>Decision-making</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>45</td>
</tr>
<tr>
<td>Techn. Change</td>
<td>27</td>
<td>100</td>
<td>6</td>
<td>35</td>
<td>-</td>
</tr>
<tr>
<td>Other contexts</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Total Freq. / %</td>
<td>27</td>
<td>9</td>
<td>17</td>
<td>5</td>
<td>137</td>
</tr>
</tbody>
</table>
Convergence as Recombination: Convergence as recombination deals with the mixing of elements, often resulting in innovation. It is used in the context of bringing together different concepts, for example, in the form of research streams, different functionalities, or media in the context of technological change. This view focuses on the mix and the outcome and does not account for an iterative process. It assumes no change in the converging elements themselves, as observed in the case of alignment, but instead results in the creation of a new element:

The industry has experienced the introduction of nearly twenty competing products [...] convergence of functionality of hand-held devices, palm devices, small phones, and car communication systems within a short time span of about 2 years. (Ramesh & Tiwana, 1999).

Convergence as Optimization: Convergence as optimization has been found mainly in the data analysis sections of the articles. An optimization problem is at hand and is analysed in different ways (e.g., genetic algorithms) to achieve a convergence to the optimal solution. This special form of convergence assumes that there is only one element that changes to a (predefined) ideal stage. It has a strong process view, and number of iterations and rate of convergence are important properties.

Convergence as Interoperability: The interoperability view on convergence is mainly found in the context of technological change, mainly in relationship to system integration and the network organization. Both deal with the detailed technical links between two or more elements moving together. While the integration form sees convergence more as a driver for efficiency, the network organizational form points out that it is set up through standards which are again a result of a negotiation or alignment process among players.

6 DATA ANALYSIS AND DISCUSSION

From the findings so far, three points can be highlighted. Firstly, the concept of convergence as technological change is relegated to the sidelines in the leading IS journals. It has a very small number of occurrences and is primarily picked up within the themes of mobility and network organisation. Secondly, convergence can be conceptualized based on the IS literature in five archetypes: convergence as alignment, interoperability, optimization, recombination, and correspondence. Thirdly, convergence communication in the context of technological change has been primarily described through convergence as interoperability and alignment and, on some occasions, through recombination. In the following differentiation analysis, the differences and the linkages among these five categories are investigated, specifically focusing on the idea of ICT Convergence. In the final step, a form analysis is conducted to identify the overall guiding distinction drawn by the authors in the context of technological change which shapes ICT Convergence communication in the IS field.

6.1 Differentiation Analysis of ICT Convergence

If the differences among the five categories are analyzed, it is observed that the categories mainly differ in the relations of the converging elements (see figure 1). The key distinction between alignment and recombination is that alignment is not about “mixing” media or functionalities. Instead, it deals with streamlining existing ideas, interests and opinions, i.e. agreement on standards. At the same time, the main distinction between alignment and interoperability is that interoperability builds detailed technical bridges or gateways among the converging elements, i.e. web services. Correspondence is a special case, assuming that the converging elements are the same, while optimization differs from the
other four conceptualizations because it assumes that there is only one element which moves towards an ideal state or optimum.

In the light of these findings, ICT Convergence is seen as a socio-technical complex (Bauer, 2002) that is observed by a very heterogeneous field of observers. The ICT Convergence complex develops parallel to other established systems that constitute its environment. Any particular observing system might be in focus, depending on the observer and his or her research question, letting the other system move into the background. Based on the observations made so far, the observers in the context of technological change observe convergence as a socio-technical process where both elements mutually constitute each other. A good example of this can be found in a definition quoted by Pawlowski & Robey (2004) from Susan Leigh Star et al. (1997, p. 4):

*Star et al. defined convergence as 'the double process by which information artifacts and social worlds are fitted to each other and come together...a process of mutual constitution.'*

While the focus of ICT Convergence in IS seems to revolve around this double process of alignment and interoperability, other forms of ICT Convergence communication cannot be ruled out. In fact, from a design perspective it might be very useful to closely examine the other three forms and, in particular, their distinctions.

### 6.2 Form Analysis of ICT Convergence

What is the guiding distinction that indicates convergence? According to Luhmann (1991, pp. 15-16), there are three ways to make distinctions. Firstly, a distinction can be made without specifying the other side of the distinction (e.g., convergence/no convergence). Secondly, a distinction can be made to restrict the other side of the distinction (e.g., convergence/divergence). Luhmann refers to the first category as objects and to the second category as concepts. Finally, there is a special kind of concept in which a distinction is made by copying it to the inside or outside of the concept itself. Luhmann calls these concepts, which can re-enter themselves, as second-order concepts (he gives an example of government and opposition, where government can itself have a deciding fraction and an opposition).

In the following section the forms of all five types of convergence are analyzed to identify the prevailing guiding distinction.

**Convergence as alignment:** Many authors make the distinction between convergence as a concept and divergence as its counter-concept. While the other side of the distinction of convergence through building shared models is the revelation of biases and conflict, divergence is seen in decision-making as part of the brainstorming phase which seeks creativity and opens up the option space. In the case of the double process of alignment and interoperability, convergence is observed as alignment becoming a second-order concept, which re-enters itself in convergence as interoperability.

**Convergence as interoperability:** Most authors do not make any explicit distinction except in the context of network organization where the other side of the distinction indicates the traditional form of closed systems.

**Convergence as recombination:** Similar to interoperability, most authors use it as an object without any clear distinction. Interestingly, recombination has a strong relation to innovation and therefore seems to be closer to the counter-concept of alignment. On the other hand, it leads to new forms which may question the existing beliefs, bringing in diversity and may result in divergence. The other side of the distinction is, in this case, a form of separation, i.e., a concentration on a specific concept, functionality, or medium (see for e.g. the original Blackberry or iPod).

**Convergence as optimization:** Here, many authors see any divergence from the optimum or ideal stage as main difference. Although this type of convergence communication has not been used explicitly in the context of technological change in the data corpus, some convergence rhetoric is based upon the belief of convergence as an ideal (e.g. discussion on ubiquitous computing). However, from a design perspective it may become problematic if the other side of the distinction is forgotten, namely that there are other alternatives as well. A similar point has been raised by Jansen and Nielsen (2005).
Convergence as correspondence: The other side of the distinction is difference. This view on convergence blends out differences and constructs an artificial sameness between two different elements. It might be helpful from a design perspective to remain sensitive to these differences.

6.3 Discussion

In most cases, the IS researchers in the data corpus use convergence as an object without any clear distinction from its environment. Some other researchers set it explicitly against a counter-concept (this was primarily divergence), and therefore fulfill the criterion of a concept according to Luhmann. Finally, some authors used convergence as alignment as a second-order concept. The first type of distinction is not very helpful; it is a distinction between convergence and everything else. It therefore offers a form, but not a conceptualization of convergence.

So what does making the distinction between convergence and divergence tell us? The unity of the distinction between convergence and divergence could be described as "mutual dependencies between elements." Jansen and Nielsen’s (2005) theory of convergence is based on a similar distinction. They call the unity "co-evolution." This indicates that convergence itself is not inevitable and that there might be other trajectories to follow. The first finding from the form analysis is that convergence itself is taken for granted and that the possibility of divergence is mostly ignored or seen as undesirable. Therefore, the relationship between convergence and divergence seems to be asymmetrical.

The second finding is that convergence is not absolute but relative to the observer. For example, ICT Convergence between mobile telephony networks and the internet may be increasing interoperability, but at the same time decreasing alignment between the actors (e.g. increasing conflicting interests). If the distinction between convergence and divergence is more closely analyzed, it can be observed that the same process can be convergence for some observers and divergence for others. The interoperability between different information infrastructures may be convergence from an infrastructure provider but divergence – in terms of increasing option space or choice – for the user. These contradictions in the guiding distinctions between convergence/divergence reveal a paradox.

Convergence becomes divergence, divergence becomes convergence – it all depends upon the observer.

According to Luhmann, two possible ways exist to deal with this paradox (Luhmann, 1991, p. 118). The first one is to replace the paradox with a new distinction; the second one is to observe it from a second-order observation and to question why more and more communication in society deals with convergence. One advantage of keeping the convergence/divergence distinction is that it puts emphasis on referentiality; it depends on the observer to determine whether something is convergence or is divergence. If this distinction is kept it is now possible to observe how the paradox unfolds. How does society deal with it? What mechanisms does it establish? Answers to these questions need to be sought to understand better the implications for design and regulatory challenges.

7 CONCLUSION, LIMITATIONS AND FUTURE RESEARCH

The idea of ICT Convergence has existed for more than 30 years and has gained tremendous importance in both practice and academia over these past decades. However, as this paper has shown, convergence has been relegated to the sidelines in IS and has, at best, been taken for granted. The purpose of this paper has been to offer a systematic analysis of how IS researchers see the general idea of convergence and, more specifically, the convergence of ICT.

The analysis of 317 articles from ten of the top IS journals is based on a grounded theory approach informed by Luhmann's Theory of Distinction. From this analysis five types of convergence communications in IS are identified. Besides the contribution to the multidisciplinary literature of second-order analysis of ICT Convergence and to the IS literature in general, this paper aims to contribute to the emerging literature which applies Luhmann’s concepts to empirical problems in particular in Information Systems (see, e.g., Kallinikos, 2006; Demetis & Angell, 2007).
Furthermore, this study offers two methodological contributions. Firstly, to the author’s knowledge this is the first study in IS, which links Grounded Theory with Luhmann’s Theory of Distinction. This approach is very helpful for the analysis because it enables the author to look not only at similarities, but also at differences. Secondly, critics of discourse analysis point out the problem that studies applying this method give only imprecise and implicit suggestions regarding how to carry out discourse analysis (Flick, 2002). Therefore, the author suggests using Luhmann’s discursive analytical strategies (Andersen, 2003) as possible guidance for discourse analysis.

Finally, the contribution to practice is to provide for regulatory and design decisions on convergence an alternative path for understanding convergence, which might encourage a less superficial and more thoughtful discussion, changing taken-for-granted assumptions on convergence itself.

The main limitation of this paper is the explicit focus on the main-stream IS journals. The primary limitation of this approach is that it does not take books, conference proceedings, working papers, or articles in other IS-related journals into account. This decision was made to keep the focus primarily on the main body of IS research and because of time constraints. Future research might consider a wider analysis of IS research, including other IS journals, conference papers, or books. Furthermore, an empirical study would be valuable to observe how practitioners attempt to deal with the convergence paradox. Finally, it might be interesting to find out why the concept of convergence has been relegated to the sidelines in the mainstream IS literature.

References


RESEARCH 2.0: IMPROVING PARTICIPATION IN ONLINE RESEARCH COMMUNITIES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0708.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Action research, Virtual community, Web 2.0, User participation</td>
</tr>
</tbody>
</table>
RESEARCH 2.0: IMPROVING PARTICIPATION IN ONLINE RESEARCH COMMUNITIES

Ferneley, Elaine, Salford Business School, Maxwell Building, The Crescent, Salford, M5 4WT, UK, e.ferneley@salford.ac.uk

Heinze, Aleksej, Salford Business School, Maxwell Building, The Crescent, Salford, M5 4WT, UK, a.heinze@salford.ac.uk

Child, Paul, Virtual Surveys Ltd, Faulkner House, Faulkner Street, M1 4DY, UK, paul.child@virtualsurveys.com

Abstract

Web 2.0 thinking and technologies create a number of new opportunities to conduct research broadly labeled as Research 2.0. Research 2.0 is a growing area of academic and commercial interest, which includes research undertaken in online research communities. This research in progress paper explores the practice of online research communities using a case study example operated by the commercial market research company Virtual Surveys Limited (VSL) in the UK on behalf of their client United Biscuits UK Ltd.

The preliminary findings are based on VSL and academics working together to improve the online research community participants’ response rate and the quality of contributions. Data collected for this study is based on meetings, participant observation, and a pilot survey of United Biscuits online research community (snackrs.com) members.

Using the responses of 112 snackrs.com community members, a preliminary typology of motivational factors is proposed. This can be used to refine the recruitment and development of activities in an online research community. Also, a model for supporting online research communities to ensure longitudinal engagement based on an adaptation of Salmon’s (2004) 5 Stage Model for e-moderation is proposed, extending the 5 stages to 7 – adding the stages of selection and disengagement

Keywords: Research 2.0, Online Research Communities, online communities
1 INTRODUCTION

In recent years there has been an exponential growth of user generated content of the web, facilitated by the emergence of the phenomenon of Web 2.0 (NetRating., 2006). The market research industry has embraced Web 2.0 tools as mechanisms for supporting their data collection activities, indeed the industry has recently started using the phrase ‘Research 2.0’ to cover a range of research methods utilising Web 2.0 tools and environments (Oxley, 2006). To reflect the market research profession’s interest in the area, the Market Research Society conference, held in December 2007, was dedicated to Research 2.0. Three particular drivers for Research 2.0 are the observation of a) declining response rates to both online and offline surveys and polls, b) users increasingly checking and posting online reviews of products, companies and services that they interact with and c) that online research offers potential costs savings of about 40% compared to traditional survey research (Stafford and Gonier, 2007). The market research environment is changing, with users not only supplying answers but also increasingly posing questions and taking a more proactive role in shaping the areas of research (Comley, 2008). The interactivity that is demanded by users and offered by Web 2.0 tools such as discussion boards, wikis and blogs (O’Reilly, 2005) can be utilised in Research 2.0 environments to provide market researchers with an opportunity to gain richer insights and a tighter relationship with their interviewees and hence provide their clients with potentially richer data. Market research companies have used online polls for several years to collect quantitative data, but the development of bespoke online panels using community tools such as discussion forums, blogs or social networks to collect qualitative data are a more recent but rapidly expanding phenomena with in excess of 1000 bespoke research panels currently being run online (Harmon, 2005). The lead companies in this area are internationals such as the GfK Group and Communispace in the US and Virtual Surveys in the UK (Comley, 2008). Online research communities (ORCs) are typically closed communities where interaction is based on an agenda prompted by the researcher or moderator. These bespoke communities may vary in size but response rates are usually higher than the open ‘naturally occurring’ online communities (NOOCs). Due to the emerging nature of this type of research inevitably there are issues such as the validity of findings (Stafford and Gonier, 2007) and the consequential maintenance of such communities (Comley, 2008).

This paper reports on an interpretive investigation in collaboration with Virtual Survey Limited (VSL) on the use of Research 2.0 techniques in one online research community they run on behalf of United Biscuits UK Ltd. Specifically, the research adopts a case study approach and aims to explore the motivations behind contributor engagement within online research communities and to identify appropriate stimuli to increase long-term contributor interaction. To accomplish this aim we draw on online communities engagement literature and e-moderation models. Afterwards, we briefly outline details of the initial findings. A preliminary typology of motivational factors based on Snacks.com is proposed and discussed in relation to the literature. This typology has been created to aid understanding of the complex mix of extrinsic and intrinsic motivational factors that contribute to participants’ rationales for engagement in online research communities. This work contributes to the management of online research communities by providing a more detailed understand of why certain members participate in such communities. It also provides guidance for online research community facilitators to assist them in supporting and encouraging those activities that increase participation in online research communities. The observations are discussed in relation to the 5 Stage Model for e-moderation (Salmon, 2004).
2 ONLINE COMMUNITIES

Naturally Occurring Online Communities (NOOCs). When developing NOOCs there are a number of key issues to be considered. According to Wenger, a community consists of three basic elements: firstly the notion of joint enterprise, that the participants share identifiable and common goals; secondly that the participants mutually engage, that they learn and undertake activities together; and thirdly that the participants have a shared repertoire, a set of communal resources that have developed as part of their engagement (Wenger, 1998). Furthermore, trust between community members is a key enabler of community contributions (Ardichvili, Page, and Wentling, 2003). For example, contributors may hesitate to contribute out of fear of criticism, contributors may deliberately or subconsciously provide misleading contributions, they may doubt the importance of their contribution, provide inaccurate contributions or doubt that their potential contributions could be relevant to a specific discussion. Indeed, to remove identified barriers, there is a need to develop various types of trust models, for example knowledge-based or institution-based. The literature also identifies a range of intrinsic and extrinsic motivation factors that influence participants in NOOCs (Nov, 2007). For example, a range of reasons were identified regarding why individuals shared, or failed to share, or engaged with other members of an online community including: self-esteem boosting, altruism, conformist considerations and the moderation processes (McLure and Faraj, 2000). Moreover, others suggest that intrinsic motivators such as a feeling of belonging or the notion of meaningful contribution are much more powerful enablers than extrinsic motivators such as monetary reward (Brandtzæg and Heim, 2008; Nov, 2007). The online activity of participants in NOOCs also provides insight into the success or otherwise of the community. For example, according to Dwyer et al (2004) there are two distinct categories of online behaviour, firstly information seeking as illustrated by passive access and secondly viewing and social engagement as illustrated by participants who undertake active contribution (Dwyer, Zhang, and Hiltz, 2004). The information seeker type implies that the members are interested in updating their own knowledge in relation to a specific area of interest. Hence membership is sustained by the quality of information provided and how it is organised and presented and the speed with which it is updated. The information seeker category reaffirms that it is not sufficient to measure the success of an online community on the active members only. For example, ‘lurkers’, those members who read the community contributions but do not post messages, can be argued to be an important element to community success. Indeed, the approximate percentage of lurkers per online community can be as high as 90% (Mason, 1999). When considering the social engagement category, the social engagement type implies a desire to develop social engagement through the community. Therefore, supporting interpersonal relationships and encouraging social engagement can be argued to be a key objective in developing successful online communities. In fact, social engagement may directly influence contributors’ satisfaction levels and the degree and quality of contact and collaboration with other community members. The role of the moderator is also shown to be critical to the success of NOOCs. For example, Ardichvili et al (2003) suggest that community moderators have to create the right conditions for content generation and dissemination. This aim can be achieved by promoting conditions for an open exchange of ideas and information, creating time and space for dialogue exchange and supporting innovative thinking. However, numerous authors question the notion of the ‘managed’ community arguing that Communities of Practice do not respond well to being managed “outside management efforts may throttle an otherwise thriving Community of Practice” (Ardichvili et al., 2003). Finally, there are a number of Human Computer Interaction (HCI) issues which, it is argued, affect user participation in online communities (Preece et al., 2007). Issues include contributor conceptualisation of the structure and rationale of the online environment, whether an environment is expressive or persuasive and whether an online environment adheres to various usability heuristics.

Online Research Communities (ORCs). Whilst the above literature review gives a rich insight into the key research issues for NOOCs there are key features of ORCs that differentiate them from the more traditional online communities. NOOCs tend to connect people with common interests, in ORCs
the community is closed and members are selected based on specified profiles, whilst these profiles tend to be of people with a common interest (for example frequent flyers) there are also occasions when the research dictates that the participants should be from outside of the niche. NOOCs tend to attract people who are passionate about the subject; whilst this may be the case in ORCs, the aim is to ensure that the participating body represents a broader view of users rather than only brand advocates. The nature of NOOCs means that the community is usually discovered via word of mouth or serendipity, in the case of ORCs, which are closed communities; recruitment to the community is targeted and strictly controlled. This means that NOOCs may have a much larger membership base than ORCs, although engagement in ORCs tends to be much higher. Furthermore, in NOOCs the agenda is self-evident, the community usually has a ‘cause’, whereas in ORCs the agenda is clearly communicated at the recruitment stage, the subject matter is usually dictated by the moderator, which frequently means that there is no shared agenda beyond answering direct questions from the moderator and therefore there are limited opportunities for the users to coalesce. Therefore, whilst in NOOCs, user interaction is usually peer-to-peer, in ORCs there is a tendency for participants to simply respond to the moderator with limited interaction between peers. This reliance on the moderator is also evident in the reluctance in ORCs, unlike NOOCs, for members to self-police; rather participants rely on the moderator to act as mediator and monitor. Furthermore, due to the ‘managed’ nature of ORCs where both moderators and clients are actively ‘listening’ and analysing responses, there is debate on whether some contributions are genuine or whether responses are tailored in an attempt to influence client decisions or attract extrinsic motivators (prizes, payments etc). Note that it is recognised that the incentive scheme has to balance the timeliness, appeal and instant gratification to the participant (Stafford and Gonier, 2007). Finally, whilst initial participant engagement within ORCs may be challenging, primarily due to the fact that the recruited participants may not be as familiar with Research 2.0 tools as those who engage in NOOCs, the evidence suggests that levels of engagement in ORCs are typically much higher than NOOCs (proportion of NOOC forum members who engage <5%, proportion of ORC members who engage 50-70% (Comley, 2008)). A summary of the main differences between the ORCs and NOOCs is presented in table 1.

<table>
<thead>
<tr>
<th></th>
<th>NOOC</th>
<th>ORC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community initiator</td>
<td>Emergent – member lead</td>
<td>Client organisation driven – research agenda driven</td>
</tr>
<tr>
<td>Moderator role</td>
<td>Policing and support</td>
<td>Research directing, steering and probing as well as policing and support. Data analysis and interpretation.</td>
</tr>
<tr>
<td>Participant to participant interaction</td>
<td>High</td>
<td>Medium to low</td>
</tr>
<tr>
<td>Membership selection</td>
<td>Usually self selecting</td>
<td>Usually selected</td>
</tr>
</tbody>
</table>

Table 1: A summary of the differences between NOOCs and ORCs

2.1 Longitudinal Participation in Online Research Communities

In order to develop a model for supporting ORCs to ensure longitudinal engagement, this research has used Salmon’s 5 Stage Model for e-moderation as a foundation. The 5 Stage Model offers practical advice on the use of online communication (Chowcat, 2005; Moule, 2007). This model is based on several years of action research on online conferencing in the Open University UK (Salmon, 2004). Whilst this model was developed with learning communities in mind, it has since been used in a number of other ways to structure online communication processes (Lynch, Heinze, and Scott, 2009). We therefore feel that the model offers a starting point in discussing online research communities.
This model of online community building and facilitation describes a five-stage process mapping the different stages of engaging participants using online communication technology. In the figure demonstrating the model, the level of engagement is indicated by the interactivity column (far right hand side) and the darkness of the colour. For example, stage 4 “Knowledge Construction” is the most interactive of all the stages, this is indicated by the black colour (the amount of interactivity measure on the interactivity column). The development starts from stage 1 “Access and motivation” and progresses up to stage 5 “Development” indicated by the arrowed “learning” line on the left hand side of the model. Each of the stages is subdivided into two triangles representing the roles of the e-moderator and the technical support staff. These roles vary at each stage.

The first stage of the 5 Stage Model is concerned with accessing the system, when participants are issued with access information by the technical support and welcomed by the e-moderator. The second stage focuses on online socialisation of the participants in the community; they are encouraged to familiarise themselves with the environment and socialise with others. The information exchange stage puts more emphasis on interaction and engages participants with the materials. The final two stages are where the participants should already be familiar with their environment and thus are able to proceed with knowledge construction and development. The model is based on a sequential principle that there are certain steps that have to be mastered before higher-level steps can be undertaken. The underlying principle is to use activities to make participants interact with each other and the e-moderator, rather than simply and passively accessing information such as handouts and presentation material. The model assumes that the participants will need to learn how to use the system’s technology and functionality over time.

This 5 Stage Model has resonance with theories about group work, particularly, the working stages such as ‘forming, storming, norming, performing and adjourning’ (Tuckman and Jensen, 1977). These similarities are particularly visible in the interactivity scale of the 5 Stage Model. The 5 Stage Model indicates that the interactivity is reduced at the final stage as is the case with group working stages of Tuckman and Jensen. The purpose of the 5 Stage Model is to address online communication and group work within the constructivist pedagogy (Salmon, 2004). This highlights some of the main benefits of the model such as simplicity (Moule, 2007), grounded in practice, reflects constructivist pedagogy, clarity, and provision of a good navigation tool for facilitators (Chowcat, 2005). However, there are also some disadvantages in the 5 Stage Model, such as the exclusion of other pedagogies and e-learning approaches (Moule, 2007). The model prescribes a course structure, ignores rhythms of participation, isolates e-learning from other learning and finally it is not a model for e-learning per se (Chowcat, 2005), which is an advantage for our investigation into ORCs

3 INTERPRETIVE CASE STUDY RESEARCH DESIGN

The research presented in this paper adopts an interpretive stance. Using an interpretive case study approach (Oates, 2006; Walsham, 2006; Yin, 1994), the research aims to explore the motivations behind contributor engagement within online research communities and to identify appropriate stimuli to increase long-term contributor interaction. The researchers involved in the study area are the authors and members of Virtual Surveys Ltd management team who include individuals with over 25 years experience of running market research studies in commercial settings. The academic researchers are experienced in the field of information systems development, knowledge management, online community moderation and are from an Information Systems background.

The ORC being studied is managed by Virtual Surveys Ltd (VSL) on behalf of United Biscuits UK Ltd (UB). There are two other communities examined in this research, but for the benefit of richness of discussion only the UB community will be described to discuss the preliminary analysis – that is the snackrs.com community. Guided by Myers’ (1997) assertion that interpretive research should present multiple viewpoints of those involved and their different perspectives, the communities were observed over a six-week period and the observations and interpretations made by the researchers were presented back to VSL management for potential improvements to be discussed and incorporated back
into subsequent data collection scenarios. The researchers were actively involved in community membership and were able to create their own posts and reply to other community members. This allowed first hand experience of interaction in the community and provided the researchers with the same level of exposure to email prompts and alerts as any other community member would experience. The presence of the researchers made any major impact on the community members, as members are already accustomed to being observed by a) the actual market research company and b) by the client company stakeholders such as UB.

Since the communities are operated by a commercial market research organisation, this research is undertaken under their strict in-house ethical approval guidelines. Before the start of this research, approval was also gained from stakeholders within the individual online research community client organisations such as United Biscuits. The reconciliation of commercial and academic interests has at times posed certain challenges that impact on the richness of the data reported, for example we are not able to report exact community member numbers. However, collaboration with a market research company also has its benefits such as the ability to use cutting edge market research technology and being allowed access to professional market researchers; for example VSL facilitated the online survey design and creation.

3.1 The United Biscuit Community

The United Biscuits community has approximately 1000 (c. 700 female, c. 300 male) members. Each online research community employs a range of Research 2.0 functionalities including voting polls, discussion forums, virtual focus groups, blog environments and functionality to allow community members to upload personal details, photos, videos and create friendship networks (see Figure 1 for a screenshot example). Each community has a dedicated moderator who posts email requests to members to contribute to the ORC on a specific topic of interest to the client on approximately a twice-weekly basis. The participants are not offered monetary reward for engaging with the online research communities however all respondents in the United Biscuit community are are offered free snacks and gain additional rewards based on ‘quality’ discussion forum contributions. VSL typically release a market research survey and then supply responses to the client within a one-week period. Responses to query requests sent by email to the ORC typically start within minutes of posting a notification of a new query to participants, peak within 24 hours, with no significant additional responses after a one-week period. As is demonstrated by the speed of responses within these commercial ORCs, the volume of data generated within a short period of time means that a longitudinal study was not a prerequisite for gaining a meaningful dataset. Response rates on given topics are on average 2% of the total community with approximately 10 to 45 responses per request. One of the issues of community management is that the majority of community memberships are inactive - for example in snackrs.com over 500 users have not posted a single contribution. Figure 1 is a screen shot taken from the snackrs.com online research forum. This screen shot illustrates that the community website has a conventional navigation bar on the left hand side of the web page and some high level navigation items at the top and bottom. The discussion forums are listed by topics and participants are given some high level information on the number of new topics, number of posts and when the last post was added.
In order to undertake this research study, data was sourced from a number of locations and stakeholders. The data sources influence decisions on potential online research community improvements. This provides an incremental improvement to the community that is continuously studied and reviewed by the research team. The decisions on implementation of any issues rest with the staff of VSL. The initial reflections, which are reported in this research in progress paper, are based on the two academic authors of this paper becoming participants in the research communities for a six-week period. The findings were discussed with the VSL staff and community improvements were initiated. A detailed discussion of these findings is the subject of this research in progress paper. At the beginning of this research a snapshot of all the communities was made which tracks all members, their profiles and number of posts to the community. The data collection and analysis is based on seven stages, extending the 5 Stage Model to include ‘Selection Stage’ and ‘Disengagement Stage’. The ‘Selection Stage’ was used to probe if there are any particular variables that determined individuals’ activities in the community. The ‘Disengagement Stage’ was included to determine any common variables that might have contributed to community members withdrawing.

4 PRELIMINARY DATA DISCUSSION

4.1 United Biscuits Community Typology

In order to improve the levels of engagement, it was perceived as important to identify the motivators for community members to take part in the community. Therefore, the first stage of the research consisted of a detailed analysis of the responses provided by 112 members of the United Biscuits community (known as snackrs.com) in response to the question ‘What are the best things about being a member of this community?’ Because qualitative data analysis is an open and iterative process, categorisation resulted in the emergence or induction of a rich categorisation as the categories were ‘extended’, ‘filled in’, ‘bridged’ and ‘surfaced’ (Lincoln and Guba, 1985). The initial analysis of the responses to the above question indicated that the rationale for participant interaction in the online research communities could be subdivided into seven motivational factors, of concern is the limited number of quotes that appear to suggest that participants are seeking social engagement. Examples of responses against typology type are presented in table 2. Note that some participants provided several responses. This means that someone can be classified as, for example, both information seeker and social engagement seeker at the same time. The typology is provisional and the types will be further refined as the research progresses. A summary of the types analysis and their potential tendencies as identified in the literature is provided in table 3. It is important to note that these are ideal case scenarios and there may be situations that will not conform to these tendencies.
Table 2: Preliminary typology of motivational factors based on Snacks.com

As we can see from this preliminary analysis there is a complex mix of intrinsic and extrinsic motivational factors that contribute to participants’ engaging in online research communities. The typology does not suggest that one member will fit necessarily into a single one of the motivational types, for example some statements of members included several reasons that they felt were their motivators to participate in the online community activities. Yet there were also some individuals who only identified one reason for participation. Since these types are based on individual member’s contributions it is not possible for them to place the entire community into one type, however, the more members of a community that subscribe to one particular motivator, the more this motivator could be used as the main source of activities generation and community management. For example, if the majority of snackrs.com identify with “Freebie seekers” this would mean that the product samples are important to them and stopping this activity could reduce their engagement in the community, on the other hand if samples were sent more frequently they might be tempted to participate more.

4.2 Tendencies of motivational factors

The preliminary typology is further explored in Table 3 in relation to the literature in order to provide insight into the reasons behind participants’ engagement in these communities. The first identified type is the Social Engagement Seeker, these individuals are characterised by comments that indicate interest in interaction with other ORC members. For example, these people might be there because they want to see what others have to say and how they can interact with those interested in a common topic. These individuals highlight the community elements of joint enterprise and collaboration as proposed by Wenger (1998). They are stimulated by intrinsic motivators which contribute to their feelings of belonging to a community, such participant types are more likely to trust others in the community and are likely to be active participants (McLure and Faraj, 2000). Considering the 5 stage model of engagement, this participant type will require minimal support at the initial access stage. They are predisposed to socialise, however the quality of the information provision and critical
reflection as they move to the knowledge construction and development stages may need significant moderator support in order to ensure that a valuable contribution from them is achieved.

<table>
<thead>
<tr>
<th>Nr</th>
<th>Type</th>
<th>Community elements (Wenger, 1998)</th>
<th>Trust between community membership (Ardichvili et al., 2003)</th>
<th>Motivators (Osterloh and Frey, 2000)</th>
<th>Participation (Dwyer et al., 2004)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Social engagement seekers</td>
<td>Yes</td>
<td>Yes</td>
<td>Intrinsic</td>
<td>Active</td>
</tr>
<tr>
<td>2</td>
<td>Power seekers</td>
<td>Not</td>
<td>Not</td>
<td>Intrinsic</td>
<td>Active</td>
</tr>
<tr>
<td>3</td>
<td>Freebie seekers</td>
<td>Not</td>
<td>Not</td>
<td>Extrinsic</td>
<td>Passive</td>
</tr>
<tr>
<td>4</td>
<td>Information seekers</td>
<td>No</td>
<td>Yes</td>
<td>Intrinsic</td>
<td>Passive</td>
</tr>
<tr>
<td>5</td>
<td>Hobbyists</td>
<td>Yes</td>
<td>Yes</td>
<td>Intrinsic</td>
<td>Active</td>
</tr>
<tr>
<td>6</td>
<td>Information hungry</td>
<td>No</td>
<td>Yes</td>
<td>Intrinsic</td>
<td>Active</td>
</tr>
<tr>
<td>7</td>
<td>Geeks</td>
<td>No</td>
<td>Yes</td>
<td>Intrinsic</td>
<td>Passive</td>
</tr>
</tbody>
</table>

*Table 3: Tendencies of the motivational factors*

The second identified type is the Power Seeker, these individuals are characterised by comments indicating a desire to have some influence, usually with the ORC sponsor (e.g. UB). These participants may have a very specific reason for engagement in the community. These individuals have a limited sense of community and, when considering the 5 stage model of engagement, whilst providing detailed knowledge, constructive dialogue may be so topic focused that their contribution to the collection of wider research data may be limited, effectively there may be a need to find mechanisms to move them back to the early socialisation stage of the model if their contributions cease to be useful or of benefit to the development of the wider community.

The third identified type is the Freebie Seeker, these individuals were motivated to join the ORC in order to gain the extrinsic rewards on offer. Logically, the expectation was that the snackrs.com community would have the highest proportion of non-contributing freebie seekers as in this community all registered members of the community were provided with occasional free samples regardless of their level of engagement in the community. However, this community has a low proportion of non-contributing members. As this research progresses, further investigation will be undertaken regarding whether the free samples act as a stimuli for engagement in the community. Considering the 5 stage model, snackrs.com members were also active in terms of knowledge construction and development – there was a great tendency to interact and debate with community development activities occurring, for example a lively informal competition concerned with posting photographs of Halloween activities. This activity was only very loosely related to the research issue (Halloween packaged sweets) yet stimulated co-construction of participant views on the subject. For future work there is a clear need to more fully understand the influence of extrinsic motivators on ORC contributions.

The fourth and sixth identified types are concerned with information gathering – the Information Seeker and the Information Hungry, these individuals are characterised by comments indicating a desire to learn and gain new information. The differentiating factor between the two types is that the Information Hungry specifically want to gain some form of ‘edge’ so that they aware of new developments first. As with Power Seekers, these individuals tended to have limited sense of community and, when considering the 5 stage model of engagement, whilst they may make some attempts at socialising, they tend to move to simple poll responses rather than towards generating useful co-constructed knowledge. Mechanisms need to be established to ensure that they are drawn into socialising and ultimately moved towards community development.
The fifth identified type is the Hobbyists, these individuals are characterised by having a genuine commitment to the ORC topic area and may make ideal participants. The danger is that they become overbearing within the community, consider for example the participant in snackrs.com who had made 243 posts. When considering this participant type against the 5 stage model of engagement the challenge is to ensure that these types become pivotal members rather than bland information providers, their contribution needs to encourage rather than overwhelm other participants. Finally, the seventh identified type is the Geek, these individuals are characterised by having an interest in the technology rather than the topic. Close attention needs to be maintained on their contribution, for example geeks are interested in the technological use of the forum and how it advances their abilities to interact online with other members. For example, the opportunity to upload their own photographs and video content created a technological challenge that some members appeared to enjoy.

4.3 Relevance of the 5 stage model to ORCs

The 5 stage model was developed for educational settings where students are primarily motivated by gaining knowledge and understanding, the moderators role is primarily to assess students’ activities and facilitate their knowledge development. This is a major differentiator with market research communities, where, as identified in our motivational typology above, the motivators for the community members vary. This leads us to re-consider the 5 stages of the model proposed by Salmon (2004). A stage which is not within the control of the educators, but is of primary concern for a commercial online research community, is the motivating factors that influence members’ enrolment in such communities. For example, if a member is displaying “Freebie seeker” characteristics does the community benefit from their presence? Although we are aware that several factors might influence an individual’s contribution and participation, a stage of member selection needs to be added to the existing 5 stage model. This stage would appear before the access and motivation stage and would comprise “Screening” the participants and their intentions. Also, it would be beneficial to highlight the key values of communities at the second stage of the model such as the “Access and Motivation” stage. Practically, this could be implemented by providing a simplified terms and conditions of the site use that highlights the need for expected interaction levels.

Another stage which was not necessarily applicable to the 5 stage model is where participants withdraw their participation. Other processes in educational settings would manage this stage, however, in ORCs it is not practical and could be done much more easily with an exit survey so that data could be captured to establish any trends regarding why members are leaving. This can feed into the recruitment and community management process to identify trends and patterns of members disengaging because perhaps their initial rationale for engaging in the online research community was not being satisfied. This leaves us with an expansion of the five stage model to seven stages, one prior to the engagement “Screening” and one after the community engagement has failed and a member withdraws - this could be referred to as “Disengagement”. Although these stages are already in place in practice in ORCs, the online research community creators tend to focus on the selection of correct sample groups rather than necessarily considering their motivational factors, which we feel are as important to the members’ existence as fitting the right market research segment or participant profile.

Based on our observations, the current interaction of community members tends to be centred around the third stage of the original 5 stage model of “Information giving and receiving”. There is little social interaction between members and where interaction happens, it is primarily between moderators and members, where moderators are trying to clarify the findings for the research brief. Having said that, there are special activities that are designed to be “fun” and to bond community members, but these are not as popular with participants as some research focused discussions. The ideal stage of “Knowledge construction and development”, where participants take over the initiative for product development say by initiating a new biscuit design or packaging for existing biscuits, does not happen in the studied community. This, to some extent, questions the notion of the Web 2.0 phenomena – the technology is there for members to take over the initiative, but they are not stepping over the barrier...
and tend to be more driven by the prompts of moderators and their research agenda. This requires further investigations.

5 CONCLUSIONS AND RECOMMENDATIONS

The current research-in-progress paper has outlined the initial findings of our research study. At this stage it is not possible to recommend firm action in order to improve the engagement of online research community members. However, we are able to provide a preliminary typology based on users’ beliefs. This can be used for by VSL for identification of future online community member types, recruitment, development of incentives and interfaces to suit certain kinds of behaviour. For example, to encourage ‘Social Engagement Seeker’ type of behaviour, more ‘fun’ activities could perhaps be introduced to facilitate community building and not necessarily focus on the core community purpose. This would allow members to get to know each other and allow those who entered the community at a later stage to catch up and see the “human side of interaction”.

The research to date has begun to suggest a profile of an ‘idealised’ ORC participant who would: exhibit community elements, trust the community, have intrinsic motivators and be an active participant. Those individuals who would fit this ideal profile are ‘Social Engagement Seekers’ and ‘Hobbyists’. However, there are drawbacks to such idealised participants. For example, the ‘Social Engagement Seeker’ is primarily motivated by interactions with others hence they could potentially contribute too many off-topic conversations that may not necessarily be of interest to the client. On the other hand the participants who are less engaged in online research communities are those that exhibit no community membership, have no trust in community, are motivated by extrinsic rewards and are passive, for example the ‘Freebie Seeker’. Arguably, there is a need for only active members of the community, who contribute to the discussions, however, those that answer polls and surveys are important to the research company when they are trying to establish a view on a specific topic that can be generalised.

The above conclusions contribute primarily to the online research communities’ moderation literature and secondarily to the moderation of general online communities. The preliminary typology requires further refinement and this is a research direction that we are taking, however, other studies in other communities would benefit this process and allow us to better understand the moderation of online communities. We feel that the 5 Stage e-moderation model is a useful guiding point for the discussion of all online research communities and not only those in educational settings where they were initially developed. However, unlike in educational settings, participants cannot be selected by the educators, in market research settings the participant selection and retention process is crucial and we propose to add these additional “Screening” and “Disengagement” stages to the 5 stage model – making it a seven stage model for online market research communities.

Future work will be based on face-to-face interviews with senior VSL management, online research community system developers and the online research community moderators who are VSL employees. The community members’ views will also be elicited through an online survey created specifically to stimulate discussion on community improvement. Additionally, telephone interviews with a sample of participant community members will be undertaken. Google analytics data (for example times of website access, frequency of contribution, pages viewed, length of engagement on the page) will also be used to provide quantitative data on general participants’ engagement patterns. Finally, participant profiles will be further reviewed.

6 REFERENCES


THE NOTION OF LIFEWORLD APPLIED TO INFORMATION SYSTEMS RESEARCH

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0748.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>IS Philosophy, Shared knowledge, Design research, Research methodology</td>
</tr>
</tbody>
</table>
THE NOTION OF LIFEWORLD APPLIED TO INFORMATION SYSTEMS RESEARCH

Andrew Basden, IRIS, University of Salford, Salford, M5 4WT, U.K.
sbs@basden.demon.co.uk

This paper is dedicated to Heinz Klein, one of our most incisive thinkers whose depth and breadth in IS research is legendary, because it was he who inspired this paper and encouraged me to develop these ideas.

ABSTRACT

Much IS research brings a priori theoretical constructs to its domain of study, and this can generate distorted outcomes. To avoid this danger, IS research should be 'lifeworld-oriented'. Characteristics of the lifeworld are drawn from philosophy and applied to examples of IS research. Surprisingly, both positivist and critical research can be both theorizing and lifeworld-oriented, though in different ways. A proposal is then made for 'lifeworld-oriented IS research', which, by taking into account the lifeworlds of both researcher and researched, can make IS research richer and more relevant.

Keywords Information systems research, lifeworld, diversity and coherence, meaning and norms, background shared knowledge.

1 INTRODUCTION

Geoff Walsham's [2001] work Making a World of Difference: IT in a Global Context, provides an insightful discussion of information systems (IS) use, by individuals, groups, organisations and society. He says his whole professional interest may be encapsulated in the question "Are we making a better world with IT?" [p.251]. With a 'basket of conceptual tools', he identifies ways to make a difference.

But in one case study Walsham's analysis becomes uncharacteristically cavalier. ComCo, a company that produced plotters, had a team of engineers to maintain them, but then outsourced the plotter business. The engineers became agents. The company had an IT system, Traveller, to enable them to communicate. Walsham cites stories from three engineers, called Gary, Keith and Neil. Neil saw his job as fixing machines, not answering customers' questions, and subverted the Traveller system. Keith used the system, and ComCo itself, to further his own goals (setting up his own company). Now listen to the story of Gary [p.68-69]:

"Gary thought that his work had been helped a great deal by the use of the remote Traveller system. ... However, because of the limited memory of Traveller, which restricted the data that could be held on it regarding past jobs, Gary kept his own manual log book with details of work done and other comments that he believed might be useful in the future. Gary thought of himself as closely associated with Comco even though he was not an employee any more. ... In some cases, where agents attacked Comco's policies, he took the side of Comco, and he was keen on recommending ways to improve work practices and reduce costs for Comco. Gary regarded customer relations as a key part of his responsibility as a Comco agent. For example, rather than relying on Comco to supply the right parts for all repairs, he kept his own stock of commonly used spare parts in the boot of his car and at his home."
These snippets ring true as 'everyday' experience. But now listen to Walsham's interpretation of this [p.91]:

"Gary was able to draw on his deep knowledge of plotters to leverage power over Comco who placed high value on his expertise."

There is nothing in the entire story, however, to suggest Gary sought to 'leverage power over Comco'. On the contrary, Gary seems to have exhibited a very different attitude, loyalty to both Comco and customers and self-giving generosity. (If the original sources from which this case study is drawn (a Ph.D. thesis and an internal report) were more nuanced in their discussion, Walsham does not say.)

To make the interpretation Walsham used Foucault's notion of power-knowledge as a 'lens' through which to examine this case. It seems this lens has distorted his view. Use of a lens focuses on certain aspects of the situation being studied, and implies "(It is valid to suppose that) life/reality is like this." The very act of focusing can blind the researcher to other aspects that are just as relevant. This often leads to simplistic explanations [Goldkuhl 2002] and, as here, distortion.

The alternative is research oriented to everyday life, or what philosophers call the lifeworld. The lifeworld attitude in research takes things "as they give themselves to us" [Husserl 1954/1970,p.156]. Husserl proposed a 'principle of principles', that we should let the lifeworld present itself to us, rather than approaching it with a priori theoretical constructs. But what does this mean? It does not mean we can find an 'objective' interpretation; it means we should take a certain attitude in research.

Since Husserl turned a philosophical spotlight on the lifeworld other thinkers have discussed it, sometimes under other names, including 'natural', 'everyday', 'self-evident', 'pre-theoretical' and 'naive' (without negative connotations). The aim of this article is to examine the notion of lifeworld, use this to evaluate examples of IS research, and propose the notion of lifeworld-oriented IS research (LOISR).

2 LIFEWORLD

The lifeworld and everyday life are closely connected, with the lifeworld being the background knowledge that is employed in everyday life. 'Everyday' includes not only home life but also work and even the activity of scientific research, and here designates an attitude rather than a social role. The lifeworld attitude is one that is open to the 'everyday' in whatever role we are playing, in contrast to what may be called the theorizing attitude, which is not.

2.1 How to Approach Understanding Lifeworld and Everyday Life

One way to understand the lifeworld and how to conduct LOISR might be to find an ontology of the things of everyday life. Goldkuhl [2002] offers one:

- Humans
- Human inner worlds (both individualised and intersubjective)
- Human actions (intervention, interpretation, reflection)
- Symbolic objects (signs)
- Artefacts (artificially made material objects and their processes)
- Natural environment (objects and processes).

Not only is this richer than the comparable ontologies of Hartmann [1952] and Bunge [1979], and more geared to everyday life, but it also allows things to exist in several 'realms' simultaneously. For example a business strategy exists as intersubjective inner world, as symbols, and as human actions. Moreover, Goldkuhl suggests that in addition to making such distinctions, scientific conceptualisation also depends on language, understood as Wittgenstinian language games. That is Goldkuhl's theory of how to provide a scientific conceptualisation of everyday life.
Though Goldkuhl's insights are valuable (they are used below), using them as a starting point smuggles a theorizing attitude into study of the lifeworld. It might distort, because it has no grounds for differentiating, for example, fun, justice, love and faith. This is not what is wanted. Instead of beginning with commitment to an ontology, it is preferable to take a lifeworld attitude to understanding both the lifeworld and theoretical thought. The Dutch philosopher, Dooyeweerd [1955/1984], is one thinker who attempted this and pieces of his thought will be employed, along with that of others, to draw together some characteristics of everyday life and lifeworld. (For a systematic introduction to Dooyeweerd see Clouser [2005] or Basden [2008].) These differentiate a lifeworld or everyday attitude from a theorizing one. This is not necessarily a complete picture; pointers are given to further reading.

2.2 Diversity of Everyday Life and Lifeworld

The first thing that strikes us when we take a lifeworld attitude, Dooyeweerd [1955/1984,1,p.3] suggested, is the diversity of aspects or facets of life that we experience, and that these aspects relate to each other. Gadamer also referred to diversity, as a "wealth of modes" [1977,p.191]. Dooyeweerd suggested these include: quantity, space, movement, energy, life, feeling, distinction, formation, expression, sociality, frugality, harmony and beauty, justice, self-giving and faith, but made clear that such lists are every tentative. In the example above, because Walsham focused on formative aspect of power, he was unable to see that of generosity (self-giving). Gary's story exhibits other aspects too, in the records, resources and loyalty, all interwoven with that of generosity.

Each aspect is a sphere of meaning and law, which is irreducibly different from that of other aspects (e.g. self-giving cannot be reduced to formation). It is this irreducibility that gives everyday life its diversity. Yet the aspects are not wholly independent of each other, for example one often needs to form plans when being generous, and in volume II of [1955/1984], Dooyeweerd discussed at length the relationships each aspect has with others. There is a coherence among the aspects that is important when understanding why theory can be problematic, below.

Reflecting on diversity with a lifeworld attitude involves distinguishing the aspects of what is concretely experienced. This makes ontology important. The lifeworld rejects the nominalist theory, prevalent in much IS research, that ontology can be explained away by epistemology, but it does recognise that ontologies are not 'truth'. Unfortunately, the way ontology has been undertaken over the past 2,000 years (e.g. metaphysics, substance concepts, naïve realism, critical realism, systems theory) all impose prior theories about what the world must be like. The lifeworld attitude is not bound by any of these. Things seem to "give themselves to us" in a multi-levelled manner as Goldkuhl [2002] pointed out: consider a sculpture; it is both a piece of marble and a work of art. Schutz and Luckmann [1973] spoke of 'stratification' of the lifeworld. To take account of such stratification while undertaking research, it can be useful to employ a ready-made ontology, and this is where Goldkuhl's [2002] ontology is useful. We might be cautious however, given its origin in theory of scientific conceptualisation. It might be better to employ an ontology whose origin lies more in the lifeworld itself, such as Maslow's [1943] famous set of needs, especially since it includes strata not found in Goldkuhl. But Maslow omitted some found in Goldkuhl.

Sometimes those being studied have their own ontologies. These should not be obliterated by the researcher's prior ontology, nor should they be accepted uncritically. Critical respect for them is needed and one way to achieve this is to use an ontology based on types of meaning rather than types of thing, such as Dooyeweerd's suite of aspects above. Dooyeweerd's suite is not only more comprehensive than others but it emerged from a lifeworld attitude and allows, like Goldkuhl's, for multi-existing. To say something 'exists' in several realms is constituted in it exhibiting several aspects. The sculpture example above is his; it exhibits at least the physical and aesthetic spheres, as he discussed in depth [1955/1984,III,p.110-127].
2.3 Lifeworld as Background

Heidegger [1927/1962] emphasised that the human being is immersed in the world, 'thrown' into it, and that is the lifeworld attitude. In everyday life we "live within" the world, so the knowledge of it that we make use of and rely on as we do so (viz. the lifeworld) must have a background character. "To live is always to live-in-certainty-of-the-world" wrote Husserl [1954/1970,p.142]; in everyday living we place reliance on our lifeworld knowledge. To Husserl, the lifeworld is intuitive knowledge, 'pregiven', taken for granted [1954/1970,p.109]. "By this taken-for-grantedness," say Schutz and Luckmann [1973,p.3-4], who developed his thought, "we designate everything which we experience as unquestionable". It has a 'tacit dimension' [Polanyi, 1967].

In a theoretical attitude, by contrast, we adopt a role of detached observers of the world, distancing ourselves from what is known, questioning what is taken for granted and trying to make aspects explicit. As a result, "The lifeworld ... dissolves ... before our eyes as soon as we try to take it up piece by piece," [Habermas, in Honneth, Knodler-Bunte and Windmann 1981,p.16]. Research involves being an observer and a reflective thinker, but lifeworld-oriented observation and reflection abates the detachment, is sensitive to diversity, and is orientated to action and living rather than pure thinking.

2.4 The Lifeworld, Theory and Critical Stance

What place might theory play in LOISR? Gregor [2006] differentiates five kinds of theory in ISR, used for analysing, explaining, predicting, explaining and predicting, and for design and action. The first, which addresses the question "What is?", is the most basic and is necessary for the development of the other types, because it provides a lens through which they see the world. So LOISR is concerned with the quality of that lens.

If we accept Husserl's 'principle of principles', that we should not come with a priori theory, how can we make use of theory as input to research? Dooyeweerd [1955/1984,III,p.31] argued that "Naive experience may be deepened through ... scientific knowledge, but cannot be destroyed by it." Theory may be brought in at the start of research as long as it deepens without narrowing or causing undue distortion.

The challenge is to minimise such distortion and the 'dissolving' of the lifeworld "as we take it up piece by piece". Using transcendental critique, Dooyeweerd [1955/1984] argued that theoretical thinking necessarily pulls aspects apart, disrupting our view of their mutual interweaving, and Clouser [2005] explains this as high levels of abstraction away from reality, so that we focus exclusively on one aspect and try to ignore its relationship with others. There are two mitigations. One is that the aspects actually remain interwoven even though we do this, so we can always look for inter-aspect relationships, especially those of dependency and analogy [Basden 2008,p.71-72]. The other is what Clouser calls lower abstraction, in which we are aware of the distinctions among aspects (e.g. the beauty vs. the cost of a rose) but do not abstract them away from the types of things that exhibit them. Both Dooyeweerd and Clouser argue that the full operation of theorizing involves not just analytical distinctions and language (as Goldkuhl [2002] claims) but also faith commitments to deep presuppositions about the nature of reality. So theory can never be 'objective'.

2.5 Meaning and Normativity

What is it that is taken for granted? Some of our background knowledge is of relational facts (in Walsham's example, Gary is agent of ComCo). A considerable amount of the lifeworld, however, is meaning (for example, what it is to be an agent) and norms (what we conceive as good and bad, for example that the majority of Gary's day-time activity should be devoted to work for ComCo). Husserl's main concern, a 'crisis' in the European sciences, was loss of meaning and of "norms upon
which man relies" [1954/1970, p.6-7]. Habermas [1987] too acknowledged the lifeworld's meaning and normativity, though his concern was apparent loss of meaning in modern life in general.

Modern thinking has fundamental problems with meaning and normativity because Hume and Kant divorced them from existence (or process), and since that time much Western thinking has presupposed that 'facts' may be studied apart from 'values'. Ethics is relegated to something personal and optional. But the lifeworld knows nothing of the supposed divorce. Husserl sought "truths that are destined to be norms" [1954/1970,p.303], for example "the genuine judge, true honor, true courage and justice" - though it is debated whether he found a way to them. Dooyeweerd, however, might have found a way, because he rejected the Kantian-Humean presuppositions and grounded existence in meaning and law [1955/1984,I,p.4], in the way described above: to be a judge, qua judge, cannot be divorced from the normative notion of justice. However, he sharply differentiated between deep normativity, which transcends us, and norms that are concrete expressions of this, which are usually socially constructed.

2.6 The Social Aspect of the Lifeworld

The lifeworld is shared with others: when "thinking together, valuing, planning, acting together" [Husserl 1954/1970,p.109], we cannot do so successfully unless what is meaningful and normative to one is so to others in largely the same way. The lifeworld has an important social aspect and a strong (though weakly-understood) link with culture and world-view.

This makes language an important issue, as stressed by Goldkubl [2002] and Hirschhein, Klein & Lytytinen [1996]. Shared meanings enable us to understand what the other is saying, and language enables shared meanings and norms to develop. Habermas [1987] argued that this occurs when we critique the truthfulness, sincerity or appropriateness of each others' statements.

There are "highly different lifeworlds in which highly different things pass as unquestioningly self-evident" [Gadamer 1977,p.189] - e.g. those of engineers and judges, right- and left-wingers, adults and children. A person will live in several lifeworlds (a judge may be a left-winger), and some overlap. Some encompass others, e.g. left- and right-wing are mainly within the lifeworld of the Western world-view.

How can there be understanding across different lifeworlds? This question is important in cross-cultural considerations in the Internet age. If all lifeworld meaning is socially constructed via language, it is possible in principle for there to be two lifeworlds that have almost no meaning or normativity in common. As alluded to above, Dooyeweerd [1955/1984] believed socially constructed norms and meanings to be concrete expressions of more fundamental meaning and normativity that transcends humanity. It is this that makes human living possible, including social construction itself, and of which there are distinct yet interrelated spheres or aspects, and which account for multi-existence. These aspects, therefore, are common to all lifeworlds. Dooyeweerd argued that the kernel meanings of aspects can never be grasped by theoretical thought but may be grasped intuitively (e.g. that of justice). If Dooyeweerd is correct, then there might be two sides to the lifeworld: kernel meanings that are intuitively graspable across cultures, and specific meanings that are socially constructed within specific cultures [Basden & Klein 2008]. There are hints of this in Husserl [1954/1970,p.144] when he differentiated "objects ... as substrates of their properties" from "manners of appearance, or manners of givenness".

3 IS RESEARCH AND THE LIFEWORLD

To illustrate these points, how characteristics of the lifeworld have been recognised (or not) in two examples of IS research will now be examined - one that is positivist, which involves quantitative model-testing, and one that is critical, which recognises social structures around IS. Interpretive research, design research and research into frameworks for understanding a field must be left to
another time. It might be expected that a positivist stance would kill any lifeworld orientation and that research from a critical perspective is more lifeworld oriented, but some surprises are in store.

3.1 Research that Involves Quantitative Model Testing

Robert, Dennis & Ahuja [2008] examined the hypothesis that social capital (of three types: structural, relational and cognitive) assists integration of knowledge in a team, which in turn enhances performance of the team, and that this effect is more marked when lean digital environments (LDE) such as synchronous text messaging are used, than in face-to-face communication. They statistically tested this by asking 46 teams of junior-level business school students in the USA to undertake a task both face-to-face and via LDE. The task was to select applicants for university places. Each member of the team was given some information known only to themselves, and the degree to which this was shared and used by others was measured, to achieve a score for knowledge integration. To give teams the chance to build social capital, they had worked together for 6-9 weeks prior to the task and expected to continue working together for another month afterwards. Social capital was measured by questionnaire just before the task. The statistically significant results were: knowledge integration positively relates to team performance, both relational and cognitive capital enhance knowledge integration, and both structural and cognitive capital have more effect on LDE than on face-to-face working. Implications for both research and practice are discussed.

Since this research exhibits a positivist flavour, it is no surprise to find very little lifeworld orientation in the way it is carried out: abstracting certain variables away from the world, reducing social capital to three of its aspects, and relying too heavily on language-mediated information (questionnaires). Abstraction causes problems when the results are used in everyday life, in that a team manager might read them as saying "For good teamwork, the one thing to get right is knowledge integration" and ignore other contributors to team performance, of which there are many. Whereas researchers might have the background knowledge that other relevant constructs were omitted, the team manager might not.

There is just a little lifeworld-orientation in the execution of the research, in that time is given for social capital (not unlike lifeworld) to build up, though this was only a few weeks. Also, the links between social capital and knowledge integration recognise something of lifeworld diversity, in that nine factors are tested. Sadly, all but one of these come from other theory rather than the lifeworld itself, and their operationalisation as questions deemed precise enough to act as metrics robs them of much of their lifeworld meaning. Trust, for example, was reduced to "Given my teammates previous performance I see no reason to doubt their competence and preparation for another team task." Validity of constructs is discussed only in statistical terms and not in terms of their meaningfulness in the domain.

The rest of the paper, however, is more lifeworld oriented. The introduction shows the research to be motivated by a number of lifeworld characteristics, including "larger pools of expertise" (diverse, shared knowledge), "failure to integrate all available information and knowledge" (a concern that presupposes coherence), "knowledge is inherently rooted in ..." (background knowledge). The development of the conceptual apparatus, of what constitutes the three types of social capital, recognises diversity of factors meaningful to teamwork, not only the nine measured but also ones like 'attitude'. (Why were not all measured?) At the end of the paper we find a return to some lifeworld themes, including whether studying students might not be fully representative of real life - though their USA, business school culture is not questioned.

Thus, though the carrying out of the research lacks most lifeworld characteristics, the research itself is situated in the lifeworld. This is not uncommon in better quality positivist research into IS use.
3.2 Critical Social Research

Critical IS research can be more lifeworld-oriented because of its innate normativity. Adam [1998] makes a detailed critical-feminist study of how artificial intelligence (AI) is ‘inscribed’ with masculinity and asks what a ‘feminine’ artificial intelligence would look like. Her approach expresses a number of lifeworld concerns. The ‘masculine’ world view may be characterised as elevating certain aspects of life, especially those of economics, logic and technology. ‘Female’ thinking is more diverse, she says, focusing instead on bodily, sensitive, caring and even aesthetic aspects. ‘Non-Cartesian’ ways of knowing including intuition are emphasised. Adam’s book is normative, in being a call for reinstatement of these lifeworld characteristics in AI and IS research.

However, Adam et al.’s [2006] paper, which portrays critical-feminist approaches to understanding the place of women in the IT industry, is rather different. Against Husserl’s ‘principle of principles’, the authors [p.1], "draw upon the theoretical constructs of the gender and technology literature to theorize the relationship between gender and technical skill and how this impacts conceptions of gender identity.” The conceptual apparatus they employ is constructed from theories: of gender and technology, of power and silence in organisations, of links between masculinity and skill, and that women can take three strategies to working in IT: not to enter, to leave, or to become an ‘it’. The bodily, sensitive, caring, aesthetic aspects in Adam [1998] play little part.

In the data collection, however, much more of a lifeworld attitude is found. The interviews gave respondents freedom to express whatever was meaningful to them; in some, "the flood gates opened". Five of the authors had themselves worked in the IT industry and so could engage with the lifeworld of their respondents. To "let ... respondents speak directly" they quoted their words. Despite the quotations having been heavily selected to support the prior theories, a variety of allusions to background issues come through that are meaningful in the lifeworld.

In the interpretation and discussion this lifeworld orientation evaporates. The intention seems to be to show that the research has satisfied five elements of a theory about what critical research is. While there may a place for that kind of discussion, the research has perhaps missed an opportunity that a full lifeworld approach could offer.

Thus critical IS research might exhibit more or fewer lifeworld characteristics. It is not unusual for critical researchers to situate their research, not in the lifeworld of the researched, but in the debates going on in their own community. This might be why the earlier example drawn from Walsham’s [2001] work likewise exhibits a lifeworld orientation to data collection that is lost in the interpretation stage.

3.3 Comparison

To say that critical research is more lifeworld-oriented than positivist research is oversimplified. A more nuanced picture is revealed here by comparing the patterns of lifeworld and theorizing attitudes displayed by these two examples. In Robert et al. [2008] (positivist) data collection was theorizing but contextualization exhibits lifeworld characteristics, whereas in Adam et al. [2006] (critical) the pattern was reversed, and data collection was sensitive to the lifeworld but its contextualisation exhibits theorizing attitudes. Though the critical stance acknowledges normativity, there can be a danger that it imposes its own a priori norm (emancipation or power) on the situation being researched.

4 LIFEWORLD-ORIENTED INFORMATION SYSTEMS RESEARCH

How, then, can we undertake or recognise lifeworld-oriented information systems research (LOISR)? This paper will venture one definition: ISR is lifeworld-oriented to the extent that all relevant
lifeworlds are respected and taken into account. Two lifeworlds will be briefly considered, those of researcher and researched. (Other lifeworlds, such as of research funding bodies, need separate treatment.)

4.1 The Lifeworld of Researcher

Some lifeworld characteristics will be implemented by following good research practice. Things like courtesy to the researched, not misrepresenting what they tell us, not misusing literature, writing clearly, ensuring good argument, and so on are part of the normativity and meaning of research learned during research apprenticeship. It is the meaningfulness of doing research that makes it fulfilling, the diversity that makes it exciting, and the coherence that enables a project's integrity within the wider research context. All such characteristics of good research relate to the lifeworld of the researcher, though often taken for granted.

Research by its nature involves theorizing. So how can this be integrated with a lifeworld attitude? If Dooyeweerd and Clouser are correct that the human process of theorizing involves abstracting an aspect away from its companions, then some integration is possible. Researchers can bring a theory to the research as long as they remember that while it focuses on one aspect there are others that are important and can impinge on it. Also Clouser's [2005] idea of lower abstraction, in which the aspect is linked with the entity that exhibits it, can ensure appropriate interpretation of research. For example Straub, Limayem & Karahanna-Evaristo [1995] seemed to demonstrate the cross-cultural capability of the Technology Acceptance Model by testing it among Japanese, American and Swiss users - but closer inspection reveals that these were all people of one type, sharing one lifeworld: airline employees.

4.2 The Lifeworld of the Researched

The lifeworld of that which is being researched challenges IS research in a number of ways. It often differs markedly from that of the researcher, so different things are meaningful and normative in each. In LOISR the researcher needs to obtain an immanent, sensitive understanding that does justice to the researched's lifeworld. To do so requires attention to each of the characteristics of the lifeworld, including its diversity, its background nature, its meaning and normativity, and its social aspects.

Diversity of the studied domain challenges IS research in at least two ways. First, it extends beyond what the researcher may be looking for and the stated purpose of the research, so the research must be designed to allow flexible response when unexpected issues or aspects are discovered. Glaser & Strauss' [1967] grounded theory approach was motivated by this. Second, it can extend beyond the known purposes of the researched. In business and government for example overt aims of efficiency, profit or accountability might mix with hidden agendas or pressures. LOISR therefore must be sensitively critical of the researched (this is Klein & Myers' [1999] Principle of Suspicion). IS applications in leisure, social networking and the home exacerbate these challenges.

Walsham's [2001] 'basket of tools' can go some way to meeting both challenges, since each tool might disclose a different aspect - but the whole basket should be used, not just Foucault's notion of power! Another approach is to use an ontology like that offered by Goldkuhl [2002], but there are dangers in this. (1) The ontology might be used to justify the researcher's prior choice of issue rather than stimulating self-criticality and ensuring all categories are given their due. (2) The ontology can be imposed on the researched, whose own ontologies (often informal) are suppressed or explained away.

The author of this paper has found aspectual analysis to be the most useful way of meeting the challenges of diversity. This employs a meaning-based ontology like that of Dooyeweerd, because it more naturally avows unexpected variety. There are many variants [Basden 2008], but at root they all rely on providing a set of ways in which things can be meaningful and good or bad, which are
intuitively grasped by researcher and researched alike and which transcend the purposes of both, so that both are stimulated to consider matters often taken for granted.

That we ‘live within’ the lifeworld means that the context, especially the social and historical context, of the researched must be thoroughly investigated. This is Klein & Myers’ [1999] Principle of Contextualization. It also has implications for design research, having been used by Winograd & Flores [1986] to promote a new approach to IS design that aims at what Polanyi [1967] calls a proximal relationship with the IT tool.

The background nature of the lifeworld challenges LOISR to find ways to explicate the tacit and avoid giving undue priority to the easily-explicit. Because of the limitations of language, the researcher should treat carefully anything mediated by language, including questionnaires, statements of purpose made by the researched, explanations and justifications, discussions, written reports, academic papers and the like. ‘Reading between the lines’ is important in all these, and should be given more priority than it currently is. Knowledge elicitation techniques can be useful because they can probe beneath verbal expressions of knowledge to differentiate understanding (which might be intuitive) from context-dependent knowledge, problem-solving intent and illocutionary intention [Attarwala & Basden 1985, Ngwenyama & Klein 1994, Basden & Klein 2008]. Especially serious is when whole spheres of meaning are overlooked in the interaction between researcher and researched, the spheres of justice, self-giving and faith in particular. Using an ontology based on meaning, such as Dooyeweerd's aspects, as a checklist can help here [Basden 2008], but more sophisticated is Winfield's [2000] multi-aspectual knowledge elicitation (MAKE) technique, which allows the researched to surface aspects in ways that are natural and meaningful to them.

The background nature of the lifeworld also has implications for design research, having been used by Winograd & Flores [1986] to promote a new approach to IS design that aims at what Polanyi [1967] calls a proximal relationship with the IT tool.

The importance of meaning and normativity implies that LOISR will not be content to study entities, structures, processes or causalities, but will give due regard to meaning and normativity. This is recognised in principle by action research, and some insights from it might be useful to LOISR. Critical ISR can be useful because of its normative thrust, but is in danger of imposing its own a priori norm of emancipation or empowerment rather than being sensitive to norms of those being researched.

The social nature of the lifeworld implies not only paying attention to the social relationships pertaining to the researched, but also taking account of the social processes of lifeworld formation, as Robert et al. [2008] above did. It can also affect data modelling, as Hirschheim, Klein & Lyytinen [1996] show. Whereas fact-based approaches assume that relationships between entities is sufficient to model all reality, what they call the rule-based approach recognises that aspects of language and sociality are also important, not just to conceptualise the domain (as Goldkuhl [2002] stresses) but in its very constitution. That language and action cannot be divorced led to the Language Action Perspective, the potential of which for ethnographic analysis is discussed by Klein & Huynh [1999]. This begins to recognise the inter-aspect coherence discussed by Dooyeweerd [1955/1984] and if he is correct we must bring other post-social aspects in too, especially those of faith (loyalty, in Gary's case) and self-giving (generosity).

This brings us back to the original challenge of connecting the lifeworlds of researcher and researched. Klein & Myers [1999] warn that the very data is socially constructed through their interaction. If the two lifeworlds are too similar, this fact can be overlooked, so the researcher should always be questioning their own assumptions. Where the lifeworlds differ another problem emerges, that of multiple interpretations. In both cases, the researcher might let the kernel meanings of the entire range of aspects shed light on the researched area and the researcher's assumptions alike. If their kernel meaning and normativity are intuitively-grasable across lifeworlds, they might offer both critique and mutual understanding, whether the background knowledge of concrete things, norms and meanings is similar or different. This approach has yet to be widely explored, but the author has found it offers a
way by which the researcher can immanently understand the lifeworld of the researched whether he agrees with their concrete norms or not.

5 CONCLUSION

This paper has proposed a new direction for IS research: lifeworld-oriented IS research (LOISR). It has suggested that a theorizing attitude in IS research can be problematic because it narrows the researcher's view and diverts attention away from issues that may be important in the lifeworld of IS use. LOISR takes full account of the lifeworlds of both researched and researcher (and others). To understand this, characteristics of the lifeworld have been discussed, which LOISR should respect: diversity, background character, meaning, normativity and social aspect. Though a number of these characteristics may be detected in extant research methods, they have not previously been put together in the way outlined here.

LOISR does not preclude theorizing, but rather provides a context in which it takes place. If the role of theorizing is to carve out an aspect to deepen our knowledge, theorizing research might be valid when those working in a field are still exploring its inner structure. When a field moves towards maturity then researchers should increasingly recognise the links with all other aspects, and thus move towards LOISR.

The reader might have noticed the number of references to Dooyeweerd. Whereas Husserl introduced the notion of lifeworld, Dooyeweerd is, arguably, the best philosopher of everyday life to date. He was open to issues in a different way because he held presuppositions very different from those that have informed Western thinking for 2,500 years [Basden 2008 ch. II].

An examination of two examples of IS research surprised us: positivist research can exhibit lifeworld characteristics in its contextualisation while critical research can be theorizing. This implies that each stage of ISR needs separate attention, with different benefits arising in each. A lifeworld-oriented (L-O) motivation for the research can widen its appeal, a L-O set of concepts and issues can be easier for readers to understand and trust, a L-O research method yields richer findings, L-O interpretation generates a more faithful picture of the domain, and a L-O discussion makes the conclusions more robust and more widely applicable.

LOISR might also enhance the experience of those engaged in IS research. Respecting the researched's lifeworld with sensitive explication and treatment can make them more favourable towards the research, even when their knowledge is subjected to critique. Researchers are likely to find the activity of research more rewarding and interesting because it is richer and invites fascination. Referring to intuitively-grasped kernel meanings can help researcher and researched understand each other better.

LOISR is a new idea, so it may be too early to know what its drawbacks might be. It could require more effort, especially in handling a wider range of factors and in explication of tacit knowledge. Unless care is taken, research could become characterized by a broad, shallow collecting of an ill-conceived plethora of pieces of information. Attention to the relationships between aspects might help to avert this danger. A number of questions remain, which deserve further consideration. Are all the insights about the lifeworld equally important? Is the above a complete list? Are the insights consistent, given the variety of philosophers from which they have been drawn? Almost certainly there is much yet to be discovered about the lifeworld.

In Richardson & Howcroft [2006] it is suggested that the researcher should not be seen as a mirror, who reflects reality-as-it-is (the positivist, objectivist assumption), but should be seen as a lens or spotlight, which focuses on one part of it, drawing attention to what others might have overlooked (the critical assumption). But perhaps a more apt metaphor for the lifeworld-oriented researcher is a lamp, which makes many types of thing visible. The colour of the light, and what is looked at, are controlled by the researcher, but the lamp nevertheless allows things to present themselves, not only to the
researcher but to all who care to look, including those who live among them. This is the hope of lifeworld-oriented information systems research.

6 REFERENCES


New lenses to investigate media use: the layering process perspective

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0375.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Computer-mediated communication (CMC), Case Study, Human computer interaction (HCI), interpretivist research</td>
</tr>
</tbody>
</table>
New lenses to investigate media use: the layering process perspective

Boukef Charki, Nabila, ESDES Business School, Catholic University of Lyon, 23 Place Carnot, 69 286 Lyon, France, nboukef@univ-catholyon.fr

Kalika, Michel, Ecole de Management Strasbourg, Université Robert Schuman, 61 av de la forêt noire, 67085 Strasbourg, France, michel.kalika@em-strasbourg.eu

Abstract

Managers spend much of their time communicating and use an ever-larger range of communication tools for the purpose. Empirical studies have shown that while ICT tools extend communication opportunities, they do not replace other means of communication. Instead, managers use a set of communication tools in which traditional media coexist alongside ICT tools.

In addition, studying the use of just one medium fails to give us a full picture of managerial communication. To gain a better understanding, we need to examine how a range of different means of communication are used.

To this end, we conducted a case study in a car manufacturing company with data mainly collected through interviews with 36 managers. Our analysis of the data showed that managers use a set of communication tools that form superposed layers, each new ICT being layered over the existing media.

Far from being the result of individual-level rational use, this layering process is socially constructed by the different users depending on their context. Our study identified three forms of layering, namely “subject to constraint”, “planned and emergent”, or “chosen”. We argue that these differences in the layering process can help explain disparities in the outcomes of ICT adoption between organisations.

Keywords: Napoleon effect, Layering process, electronic mail, case study.

Managers spend a large part of their time communicating, and the range of communication tools they use is growing (Te’eni, 2001). However, they are also short of time and need to manage “time famine” (Perlow, 1999, p.57). As information and communication technology (ICT) applications should enable them to reduce traditional lag time in theory, the use of such tools appears to be a crucial factor for managers’ success. In addition, companies have invested massively in ICT applications in the hope of benefiting from their communication capacities (Watson-Manheim and Bélanger, 2007). Their outcomes, however, differ significantly between organisations.

The growing use of ICT applications raises the question of whether or not they replace traditional means of communication (face-to-face, meetings, telephone, etc.). Several studies have explored the issue from a perspective of both complementarity and substitution (Crawford, 1982; Culnan and Markus, 1987; King and Xia, 1997). The results, however, have been contradictory. In contrast, the co-existence of different means of communication that form superposed piles has been noted (Boukef Charki and Kalika, 2006). There is a layering process through which each new ICT tool is layered over the existing media. Kalika (2002) proposed the concept of the ‘Napoleon effect’ to describe this layering process. In baking terms, a Napoleon (also known as a cream slice or millefeuille) is a cake in which pastry, cream or custard, and fondant icing are stacked in layers. A Napoleon is therefore made...
up of a number of layers of different ingredients. The Napoleon effect refers to the superposition of different layers of media. As all media coexist, each new medium is superimposed over the existing ones and thus form the Napoleon effect.

In this paper, we have attempted to further our understanding of the Napoleon effect by examining the use of e-mail, meetings and telephone. Fast becoming ubiquitous (Rice & Gattiker, 2001; Van den Hooff et al., 2005), electronic mail is used to communicate quickly and instantly without worrying about the availability of the recipient (Straub & Karahanna, 1998). Communicating electronically therefore allows greater time management flexibility and would appear to entail fewer constraints than meetings or phone calls. Studying the use of e-mail in comparison with meetings and phone calls is therefore of interest to both academics and managers.

We begin by discussing our theoretical development and describe our research methodology. We then present and discuss our results. The paper concludes with our contributions and areas for further research.

1 THEORETICAL DEVELOPMENT

Comparing the uses of new ICT media and traditional ones is a fertile research area (Watson-Manheim and Bélanger, 2007). We first point out the shift from studying a single medium choice to the management of a set of communication media. We then discuss how the layering process perspective can bridge the gap between existing theories.

1.1 From a choice of medium to management of a set of communication media

Several theories have been put forward to explain how and when a medium is used. Theories on media use are often either single medium theories or multiple media use theories. Most of these theories focus on a single medium choice (media information richness theory, channel expansion theory); few have looked into multiplicity of media use. These theories have been formulated only recently and are not yet well documented.

1.1.1 Single medium choice theories

Single medium theories focus on why a medium is used. Different theories, among them information richness theory, have attempted to account for the choice of medium. This theory provides a rational choice explanation with respect to the individual’s choice of medium (Daft and Lengel, 1984; 1986). Its conclusions have been called into question by many researchers, however. Initially formulated for traditional communication, its application to ICT applications such as e-mail is controversial. In addition, information richness theory offers little explanation of e-mail use (Van den Hooff, 2005). Indeed, the ability of the medium to overcome temporal and spatial constraints and the individual user’s experience with e-mail are often more influential in the choices made. In addition, e-mail use may be motivated by the need to overcome spatial and temporal constraints without concern for the content of the message and its ambiguity (Dimmick et al., 2000; Van den Hooff et al., 2005). While proximity fosters face-to-face communication (Fiol and O'Connor, 2005), e-mail use is often justified by distance (Rice and Aydin, 1991; Straub and Karahanna, 1998). The choice of medium conveys a symbolic meaning, independent of its objective features.

In an extension of information richness theory, the symbolic interactionist perspective (Trevino et al., 1987; 1990) integrates situational constraints and symbolic considerations. At the same time, channel expansion theory (Carlson and Zmud, 1999) highlights the importance of experience. These extensions have improved our understanding of e-mail use while remaining rational choice theories made at individual level. However, these theories cannot explain the fact that two groups with the same context differ in their use of means of communication (Zack and McKenney, 1995).

Fulk et al. (1990) make the assumption that both sense-making and behavior are subject to social influence. Choice of medium relies on a subjective rationality influenced by colleagues’ behavior,
experience, and group standards. This means that e-mail may be used even if there is no correlation between the communication content and the objective characteristics of the medium. This choice may be considered rational by group norms of communication even though it does not correspond to the effectiveness criteria identified by information richness theory (Fulk, 1993). Markus (1994) shows that social definitions relating to e-mail use can be different from those envisaged by information richness theory. Users conform to social standards in determining the behavior to adopt, which implies that their personal choice is limited.

Thus, single medium choice theories have provided different explanations for the choice of medium. Medium use is explained not only by contingency factors but also by situational factors and social influence. Besides, many studies have demonstrated the complementarity of these factors (Wijayanayake and Higa, 1999; Webster and Trevino, 1995; Trevino et al., 2000; Kraut et al., 1998; Van den Hooff et al., 2005).

Still, users are not confronted with the exclusive choice of just one medium, but rather with managing a set of communication means. In this sense, Dennis et al. (2008) argued that communication performance will be improved when individuals use a variety of media to perform a task rather than a single medium. Moreover, studying the use of a single medium cannot give us full understanding of users’ communication. So, recent research has looked into the use of an array of communication media.

1.1.2 Multiple media use theories

Since managers use a range of communication tools, we focus on multiple media usage (Bélanger and Watson-Manheim, 2006; Massey and Montoya-Weiss, 2006; Watson-Manheim and Bélanger, 2007; Lee et al., 2007). Multiple media use theories focus on the management of an array of communication media to accomplish a specific task. The studies that have investigated multiple media use are summarized below (table 1).

<table>
<thead>
<tr>
<th>Authors</th>
<th>Multiple media use focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massey and Montoya-Weiss (2006)</td>
<td>They distinguish between the use of a single medium at a time and the use of multiple media concurrently. They suggest that media utility influences media selection and use.</td>
</tr>
<tr>
<td>Bélanger and Watson-Manheim (2006)</td>
<td>They distinguish different forms of multiple media use combinations which can be sequential (two or more media used over time) or concurrent (two or more different media used at the same time).</td>
</tr>
<tr>
<td>Lee et al. (2007)</td>
<td>They study the notion of a communication portfolio, “a set of ICTs that organization members can use for communication” (p. 135). The communication portfolio can be characterized by its size (the number of ICTs used), its content (specific ICTs used) and its structuring mechanisms (how the different media are used). The authors focus on three structuring mechanisms, which can be sequential, concurrent, or repetitive (the same ICT is used more than once).</td>
</tr>
<tr>
<td>Watson-Manheim and Bélanger (2007)</td>
<td>They propose the communication media repertoire, which is “the collection of communication media and identifiable routines of use for specific communication purposes within a defined user community” (p. 283).</td>
</tr>
</tbody>
</table>

Table 1. Summary of multiple media use studies

For a specific task, users can either use a single medium at a time or multiple media simultaneously (Massey and Montoya-Weiss, 2006; Bélanger and Watson-Manheim, 2006; Lee et al., 2007), sequentially (Bélanger and Watson-Manheim, 2006; Lee et al., 2007), and repetitively (Bélanger and Watson-Manheim, 2006; Lee et al., 2007). In addition, Bélanger and Watson-Manheim (2006) show
that users can combine different ICTs to send the same message which can be either sequential or concurrent.

Given the overwhelming range of communication media available in the workplace and the possible combinations of media (Lee et al., 2007), one wonders how users manage. Multiple media use enables “message clarity” but also leads to redundancy (Lee et al., 2007). This multiplication of media raises overload issues (Bélanger and Watson-Manheim, 2006), since individuals must cope with the escalation of incoming messages. This processing effort needed to deal with the messages or signals is described as communication load which will lead to information overload (Jones et al., 2004). Information overload occurs when not all communication input can be processed and used (Jones et al., 2004).

Information overload is a managerial issue, as the growing use of ICT tools does not obviate the need for more traditional media. Instead, we observe an accumulation of means of communication that can be likened to the layering typical of a napoleon or millefeuille (Kalika et al. 2007; 2008). In other words, as the number of layers increases, the napoleon becomes too much of a mouthful. We thus need to focus not only on multiple media use but also on the coexistence of different media.

1.2 The layering process perspective

The growing use of ICT applications raises the question of whether or not they replace traditional means of communication. Belanger and Watson-Manheim (2006) point out that substitution does not explain the observed media use. Instead, they note innovative media use combinations. Since media coexist, it is necessary to transcend the complementarity and substitution dichotomy. To do so, we propose the layering process perspective as new lenses to examine media use accumulation.

Table 2 compares the layering process perspective to single medium choice theories and multiple media use.

<table>
<thead>
<tr>
<th>Theories</th>
<th>Single medium choice theories</th>
<th>Multiple media use theories</th>
<th>Layering process perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus</td>
<td>Choice of a single medium</td>
<td>Management of a set of communication media</td>
<td>Layering process of communication media</td>
</tr>
<tr>
<td>Investigation</td>
<td>Why a medium is used?</td>
<td>How and when a set of communication media is managed to complete a specific task?</td>
<td>How and why different available communication media coexist? To put it differently, how different layers of media are superimposed?</td>
</tr>
<tr>
<td></td>
<td>Fulk (1993)</td>
<td>Lee et al. (2007)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fulk et al. (1990)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Comparison of the layering process perspective to single medium choice theories and multiple media use

We argue that by exploring the layering process we can improve our understanding of how different media coexist and thus address information overload.

We posit that there is a process by which new media are layered over existing media. Layers of media thus overlap. The added media may or may not modify those already present. We can thus expect to identify different forms of layering depending on how the added layer changes patterns in the usage of
extant media. When the added medium is superimposed on existing layers without being mixed in with them, media accumulation is the result. In other words, as the layering process does not reshape the design of previous layers and does not reduce the existing media use, the napoleon effect is more pronounced, and the resulting concoction becomes, so to speak, harder to swallow. By contrast, when the previous layers are redesigned, the napoleon effect is less pronounced. Since the added layer is mixed in, the napoleon is, as it was, easily ‘eaten’. The array of available media is better managed.

In addition, as managers use a range of communication means, we need to focus on multiple media usage (Massey and Montoya-Weiss, 2006; Watson-Manheim and Bélanger, 2007). Indeed, managers are not confronted with the exclusive choice of just one medium but rather with managing a set of communication means. To improve our understanding of managers’ communication, we focused not only on e-mail use but also on how different media are used.

We demonstrated that the fit between the task and the medium are not the only explanations for the choice of medium. This means that we need to consider situational factors, social influence and the emergent properties of the medium. Moreover, a focus on the use of just one medium is obviously unable to encompass all the complexity of managerial communication. We can improve our understanding by examining how different sets of communication tools are used and we therefore propose the Napoleon effect as a framework for studying media use.

We argue that by exploring the Napoleon effect, we can better understand how managers use different communication media. Their use is explained not only by contingency factors but also by situational factors as well as social influence and the emergent properties of the media. This leads us to expect that the layering process of the new media will vary according to the context, which is characterised by task, organizational environment and the group’s internal system (Desanctis and Poole, 1994). We posit that there is not just one Napoleon effect but several, depending on the context.

Our paper aims to study the Napoleon effect in different contexts. We therefore pose the following research question: how does the Napoleon effect differ between managers in relation to their specific contexts?

2 RESEARCH METHOD

To improve our understanding of the layering process that is intrinsic to the adoption of a new medium in different contexts, we conducted a case study in a French car manufacturing company (Eisenhardt, 1989).

2.1 Research sites

Given that multiple cases facilitate comparisons (Eisenhardt and Graebner, 2007), we conducted a multi-site case study that focused on three departments in the company. According to Miles and Huberman (1994), inter-site analyses allow for enhanced comprehension and clearer explanations of the phenomenon under study, while multiple cases provide grounded empirical findings (Eisenhardt and Graebner, 2007). In addition, the more the number of contexts studied, the greater the validity of the research (Lee, 1989). The importance of case selection was highlighted by Eisenhardt (1989) and we therefore selected three departments which differ with respect to the task carried out and the context of ICT use. While promoting the use of ICT tools is a key concern in the company as a whole, we identified certain differences and characteristic features with respect to each of the three departments studied:

Department 1
Department 1 is a strategic department in the company, dealing with car design and engineering. The work is highly interdependent and requires the involvement of numerous collaborators. A great deal of communication is needed to coordinate both inside and between the different units. In this sense, managers work in an open space and units are co-located.
E-mail is the main ICT tool used by the managers. Disparities were noted between the managers, however, depending on their hierarchical level. Senior managers appeared reluctant to use e-mail and needed the assistance of their secretary to process and print their e-mails.

- "... I followed suit like many of the people here...I was not a pioneer...I had to use it (e-mail) and I used it..."
- "...I can’t always be sitting in front of my computer to read my e-mails, I need my secretary’s help to print them..."

**Department 2**

Two main tasks are carried out in this department, namely the design and sale of parts and accessories. The managers work in an open space.

Department 2 is a pilot department in ICT use.

- "... (name of company) launched a BtoE project and we volunteered to be a pilot department..." (head of department 2).

The e-management project aims to promote new managerial practices through the use of ICT.

- "... we were interested in an experiment which allows us go beyond administrative productivity gains to define new ways of working and new practices." (head of department 2).

The head of the department initiated this project which involves all the local department managers. The e-management project aims to improve coordination mechanisms and define rules to improve e-mail use. Table 1 summarises the characteristics of these objectives.

<table>
<thead>
<tr>
<th><strong>E-management project objectives</strong></th>
<th><strong>Details</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Improve coordination mechanisms</td>
<td>The e-management project aims to reduce the number of meetings and encourage informal face-to-face communication.</td>
</tr>
<tr>
<td>Define rules to improve use of e-mail</td>
<td>The e-management project aims to improve the use of e-mail. Some rules have been adopted to reduce e-mail overload including reducing the mailing list and avoiding the use of CC.</td>
</tr>
</tbody>
</table>

Table 3: Objectives of the e-management project

While the e-management project is not limited to ICT use, it aims to improve coordination through ICT use.

**Department 3**

The managers in this department deal with internal and external corporate communication. E-mail is the main ICT tool used by managers to coordinate the widely dispersed networks so as to manage temporal and geographical constraints.

The characteristics of the three departments studied are summarized at table 4.
2.2 Data collection

Our case study led us to combine several sources of data (Eisenhardt, 1989; Eisenhardt and Graebner, 2007). In addition to interviews, we had access to internal documents and took part in meetings. The meetings were particularly useful in helping us understand the company context as well as that of each department studied. We systematically took notes during the meetings and analysed them later. This mix of methods and sources of data collection enabled us to ensure the internal validity of our research (Miles and Huberman, 1994).

We used diverse data collection methods, but our main sources of analysis were semi-structured interviews (Myers, 1997; Myers and Newman, 2007). We conducted thirty-six in-depth interviews that each lasted one hour on average. Particular emphasis was laid on (1) the work setting, (2) the ICT used to carry out the work, and (3) ICT use vs. traditional media.

2.3 Data analysis

All our interviews, apart from one, were recorded and fully transcribed. We first analysed the media used by each manager, and then in each department. Finally, we compared the differences between the three departments.

Media use analysis
For each manager, we firstly analysed the use of different communication means and then built a media choice matrix.

Intra-case analyses
As recommended by Eisenhardt (1989), we drew up a full description of the characteristics of each department’s context. After analysing the different aspects separately (use and choice), we attempted
to establish links between them. We progressively drew up summaries in which we formulated our
cconclusions in line with our theoretical framework. For each department studied, we built an intra-
department matrix to describe the media use. Our familiarisation with the characteristics of each
department made the inter-department comparison easier (Eisenhardt, 1989).

**Inter-cases analyses**
The inter-department matrix compares the differences and similarities between the three departments
studied as suggested by Eisenhardt (1989). We also highlighted the specificities of each department.
Our conclusions were formulated through successive iterations between theory and our empirical
results. Our interpretations were honed through an analysis of internal documents and the notes taken
during the meetings.
Finally, in order to ensure the plausibility of our findings, our results and our own interpretations were
presented within the company and communicated to the participants.

3 RESULTS

Meetings remain the company’s favourite coordination mechanism, although the number of meetings
is a key issue. Managers frequently complain about meeting overload and its impact on their agenda.

- "... unfortunately, meetings are a disease, a disease of (the name of the company), it’s a well-
  known fact ...".

E-mail is the main ICT tool used by the managers. It extends communication opportunities as it allows
managers to surmount temporal and space constraints.

- "it allows us to dialogue easily with people all over the world whatever the time or the day,
it’s no longer an issue...

Despite its growing use, e-mail has not replaced the need for other communication means. This means
that managers use a set of communication tools in which e-mail is used in addition to meetings and the
telephone but does not supplement them. Rather, it leads to the superposition of different
communication means whereby different media are layered and thus form the Napoleon effect.
Disparities were observed between the three departments in their use of the different communication
means, however. A detailed description of media use in each department is presented below.

3.1 Department 1

E-mail use is tracked through an external memory, which enables the history of the exchanges to be
traced. This is particularly useful with respect to the number of people involved and the strategic tasks
carried out in this department for the company (car design and engineering).

- "... we need to keep track of the information...to keep tabs on the threads...".

E-mail use can reduce face-to-face but not the number of meetings. The managers interviewed often
complained about meeting overload. This problem is accentuated by the impression that many of the
meetings are unjustified.

- "...We spend a lot of time in meetings...I’m sure that 80% of them are unnecessary...

The number of meetings is accounted for by both coordination needs and communication norms.

Indeed, car design and engineering necessitate the involvement of several managers from different
units of this department.

- “...it is a complex product... and it needs the contribution of managers from design,
purchasing, manufacturing, logistics units, etc, ... they all have to work together, and this is
not simple...

Thus, e-mail does not seem to fit the requirements of coordination constraints at this department,
In addition, meetings are privileged coordination means in the department. In this sense, the culture of the company is described as “tribal” encouraging meetings.

- “...a lot of meetings are useless...nothing really happens... no decisions were made...if meetings are more decision-oriented, we will make fewer...”

It turns out that even though meetings are required to insure coordination between different units of the department, the higher number of meetings is an issue in this department. Managers suffer from meetings overload.

E-mail overload is also an issue in this department, so managers have taken to using the telephone instead.

- “to avoid this problem (e-mail overload), we prefer to use the telephone...so that we can have an answer immediately...rather than send an e-mail and increase overload...we place more emphasis on the phone...”

In short, both e-mail overload and the resistance of senior managers to e-mail use are constraints in the use of this medium. More particularly, problems related to e-mail overload reinforce the need for phone calls. Thus, in this case, we do not only observe an accumulation of different layers of media but also an increase of telephone layer. The outcome of this layering process is constrained by the problems of e-mail use and coordination needs in this department. As e-mail layer fails to reshape the existing ones, the Napoleon effect becomes then more pronounced.

3.2 Department 2

While meetings are considered the main means of coordination in the company, managers in this department say there has been an improvement, with less time spent in meetings and a reduction in their number. Meetings have also become more decision-oriented.

- "... it was the nature of the meetings that we wanted to change...”
- “...we ask for a meeting when there is a complex problem which requires that we meet to discuss about the issues”.

This improvement is in line with the aims of the e-management project. However, contrary to the hoped-for outcomes, e-mail is used with colleagues situated in the same open space at the expense of face-to-face. In fact, e-mail is used to communicate with colleagues who are working close by in order to avoid disturbing them.

- "... it’s something that really amazes me, we’re in an open space, and very often you don’t hear any voices, just click, click, click... the noise of fingers on the mouse... it’s obvious that e-mail has replaced face-to-face...”.
- “... we use e-mail heavily everywhere in the company and even in meetings. We use it even more than the telephone. As we are overloaded, we cannot see each other very often, we use e-mail instead...it is rather funny as it conveys closeness...”

Indeed, proximity is perceived as hindering oral communication.

- “moreover, in open space one hears all... if everyone starts to speaking...it quickly becomes cacophony”

We also noted a reduction in phone calls eventhough they are still preferred when there is an emergency and there is a need for interactivity.

In this department then, the use of e-mail has led to improvements in meetings but also a reduction in face-to-face contact. E-mail use in this case reshaped the design of previous layers. Unlike the previous department, the layering process of e-mail layer over the existing ones manages to tighten them. The napoleon effect is then less pronounced than it is the case in department 1.
Although the first outcome of this layering process was planned (meetings are reduced and are better managed), the second (face-to-face reduction) is emergent and even goes against the spirit of the e-management project.

3.3 Department 3

While e-mail is the main ICT tool used by all the department’s managers, we noted that phone calls and face-to-face communication is preferred.

Phone calls are considered to convey greater proximity. They are also often preferred when managers communicate with their international network.

- "... the telephone because it confers a certain degree of proximity and is more friendly..."

This preference for the telephone may be explained by the nature of the tasks carried out, which require greater sensitivity with respect to the relational. In addition, managers prefer face-to-face contact when communicating with co-located workers in the same open space. However, the superposition of the different means of communication has not led to a reduction in the number of meetings, which is considered to be even higher.

- “(meeting) is an issue in the company, there is a mania of meetings...”

In conclusion, it appears that in spite of the growing use of e-mail, phone calls and face-to-face are still preferred. E-mail use does not reshape the design of the existing layers.

We summarize the layering process at each department at table 5.

<table>
<thead>
<tr>
<th>The layering process of e-mail over the telephone</th>
<th>Department 1</th>
<th>Department 2</th>
<th>Department 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-mail overload increases the need for the telephone and thus reinforces its layer.</td>
<td>E-mail layer may tighten telephone layer.</td>
<td>E-mail layer is superimposed over telephone layer without modifying it.</td>
<td></td>
</tr>
<tr>
<td>The layering process of e-mail over meetings</td>
<td>E-mail layer does not reshape meeting layer.</td>
<td>E-mail layer may tighten meeting layer.</td>
<td>E-mail layer does not reshape meeting layer.</td>
</tr>
<tr>
<td>The outcome of the layering process</td>
<td>The layering process results in an accentuation of some previous layers (telephone layer).</td>
<td>The layering process results in a reduction of previous layers.</td>
<td>The layering process results in an accumulation of different layers without modifying the previous ones.</td>
</tr>
</tbody>
</table>

Table 5. Comparison of the layering process between the studied departments

4 DISCUSSION

Managers use a set of communication tools in which both traditional and new media coexist. Our analysis shows that the use of ICT, in our case e-mail, extends the choices available to managers without completely replacing traditional media. All available media form superposed piles or the so-called Napoleon effect. However, we noted some differences in the layering process between the departments studied. Three forms of the Napoleon effect were identified in this study depending on the context. We first analyse the differences in context and then focus on the characteristics of the different forms of Napoleon effect.
4.1 **Differences in the context of each department**

The differences between the three departments studied depended on the characteristics of the departments’ activities, the communication norms and the context of ICT use (Desanctis and Poole, 1994; Markus and Robey, 1988; Robey and Sahai, 1996).

4.1.1 **Characteristics of the departments’ activities**

The three departments differ mainly with respect to the activities carried out, namely car-design (department 1), sales (department 2) and internal and external communication (department 3). Each department therefore has different coordination needs (Lawrence and Lorsch, 1967) and so differs with respect to its use of communication means. In particular, we noted that the large number of interlocutors and the complexity of the manufactured product led to a high prevalence of meetings in department 1. In parallel, the function of department 3 (communication) leads managers to prefer direct communication and this would explain why e-mail has not replaced the telephone. E-mail use therefore effectively appears to be explained by task-related activities (Yoo and Alavi, 2001). In this sense, Germonprez and Zigurs (2009) emphasized the need to consider the fit between the technology and the task. Put it differently, the fit between media used and the task that department undertake explains how the layering process takes place.

4.1.2 **Communication norms**

Along with Hill et al. (2009) we emphasized the importance of communication norms in explaining media use. Moreover, our analysis of e-mail use enabled us to highlight shared practices and communication norms in each of the three contexts studied. We noted, in particular, that meetings remain a privileged means of coordination in department 1, while e-mail is used when communicating with co-located colleagues in department 2, and there is a preference for direct contact in department 3. E-mail use can reinforce or support the emergence of new practices and communication norms that we identified through our analysis of meetings and direct contacts.

The Meeting layer: hard to be removed

Meetings are formal routines (Dosi et al., 2005) through which information is exchanged at different hierarchical levels. In addition to institutional meetings that are planned throughout the year, meetings can either be programmed or not. The number of meetings is a major issue in the company and all the managers interviewed complained about meeting overload. Improvements have been observed in department 2, however. Meetings are better managed and focus more on decision-making. This is largely due to the e-management project, which aims to improve working methods through ICT use. The involvement of local management in the project has helped achieve this objective. The layering process is not spontaneous in this case but has instead resulted from the management’s determination to change working methods. Along with Kalika et al. (2008), we emphasized the “involvement of management” effect through the case of the department 2. Local management was in this case successful at influencing the layering process.

E-mail use can either or not reshapes meetings layer depending on the case. Nevertheless, in all the cases, it can not remove the need for meetings and the result is an accumulation of media. Indeed, even if e-mail is used for a broad range of tasks (Van den Hooff, 2005), it can not remove the need for face-to-face communications. It turns out that while the objective features of e-mail explain the layering process, it fails to explain differences the way it happens. We demonstrated that when the e-mail layer is added, the layering process over meeting can have several forms. These differences are explained by the specificity of each context.

The direct contact layer: beyond the apparent contradiction

An analysis of the use of e-mail in departments 2 and 3 indicates that proximity may be interpreted differently. Indeed, whereas proximity is regarded as a threat to face-to-face communication and
phone calls in department 2, it facilitates face-to-face in department 3. Differences in interpreting proximity are explained by the shared practices and communication norms specific to each context.
- In department 2, managers use e-mail instead of oral communication (face-to-face and phone calls) to avoid disturbing their interlocutors and co-workers.
- In department 3, managers are more sensitive to relational aspects and thus prefer face-to-face communication and phone calls to e-mail use.

The choice of medium depends on the significance users assign to their behaviour, which would either reinforce or support the emergence of new communication norms (Feldman, 1984; Germonprez and Zigurs, 2009). E-mail use does not modify communication norms in department 3 but it does contribute to the emergence of new norms in department 2. Managers prefer to use e-mail even with co-located colleagues. To put it differently, the layering process of e-mail over direct contacts can either or not reshape the design of previous layers.

4.1.3 Context of ICT use

While ICT use is a key concern in the company, differences can also be identified between the three departments with respect to the context of ICT use.
- The managers in department 1 need their secretaries to help them process their e-mails, making this form of communication more problematic as a substitution for other forms. In addition, senior managers are reluctant to use e-mail and tend to use the telephone to overcome e-mail overload.
- Department 2 is a pilot site in ICT use. The e-management project aims to improve working methods by promoting ICT tools. Local management is also highly involved in this project.
- E-mail is the main ICT tool used by all the managers in department 3.

In other words, differences in ICT context explain the specificity of e-mail use in each department and thus e-mail layering over the existing layers.

4.2 Different forms of layering process

In our study, we identified different forms of the layering process that depended on the context of each department, analysing the characteristics of task-based contexts, communication norms and the context of ICT use. We concluded that previous media layering can provide insight into the way a new medium will be layered over the existing ones. The use of a new medium can either reinforce or reshape old layers. Users make sense of the added layer taking into account the opportunities and the constraints of their context. The layering process is then far from being a result of an individual rationale choice.

At the same time, we demonstrated that local managers can influence the layering process. They act as sense-givers (Maitlis, 2005) and thus may redesign the layering process. However, this may happen in way that is contrary to the e-management project objectives and may even go against the spirit of the media (Desanctis and Poole, 1994). It turns out that while the layering process can be reshaped by local management, it can also be interpreted and negotiated by managers depending on their particular objectives (Lamb and Kling, 2003). This interpretation may be different from what is initially expected by local managers. Users reshape the layering process through their interaction through the media and with each other. Thus, while the layering process is rooted in the context of media use (the fit between media and tasks, communication norms and the context of ICT use), we demonstrated that the layering process results also from the negotiation of the sense making of users and sense giving of local managers.

To summarise, whereas the layering process is reshaped by previous layering of media, it is also planned and emergent. The layering process is hence socially constructed. We can expect to identify a large number of layering processes in line with the sense local managers attach to the new medium and the sense understood by individual users, taking into account the layering patterns of existing media. We highlighted three forms of layering patterns in this study:
- The meeting is the main coordination mode and it is difficult to substitute this by electronic communication because of the constraints of the departments’ activities and traditional communication norms. In addition, both e-mail overload and the reluctance of senior managers to use e-mail can explain its non-substitution for meetings and phone calls. Because of these constraints, the use of e-mail does not appear to change how the previous layers are used, but instead tends to strengthen them. We posit that the layering process of different media is subject to specific constraints, and in this case the Napoleon effect is seen to be subject to constraints.

- E-mail use leads to improvements in meetings which are subsequently better managed. In addition, e-mail is preferred to face-to-face communication and phone calls. While it can lead to new communication opportunities, the use of e-mail with co-located colleagues goes against the spirit of the “e-management” project. In this case, e-mail use changes how previous layers were designed. The process of layering is therefore thought through in that it is planned within the frame of the e-management project but is also emergent. The Napoleon effect thus appears to be both planned and emergent.

- Despite the growing use of e-mail, phone calls and face-to-face communication are preferred. Both the characteristics of the task carried out and the communication norms encourage direct interaction at the expense of electronic exchanges. In this case, e-mail use does not change how previous layers are designed. The Napoleon effect may consequently be said to be chosen.

5 CONCLUSIONS

In spite of the limitations of subjective measurements (managers’ perception), which exclude objective factors such as the variation in the number of meetings and phone calls (Limayem et al., 1997), we identified both theoretical and managerial contributions. The focus on the Napoleon effect enabled us to improve our understanding not only of the choice of medium but also of media use. It provided us with greater insight for understanding how different means of communication are used by studying the layering process that underlies their adoption and use. However, different forms of layering process can be identified, and our study ascertained three forms in particular, i.e. the Napoleon effect may either be subject to constraints, planned and emergent or chosen, depending on the context. We demonstrated that the Napoleon effect depends on both the sense-making of users and the sense-giving of local managers. The layering process is thus interpreted and negotiated by users (Lamb and Kling, 2003), which can either be in line with or go against the spirit of the media (Desanctis and Poole, 1994). The Napoleon effect is the result of a negotiated use of the media, depending on the design of previous layers. This negotiation can either modify or reinforce the existing media layering. Far from being the result of rational use at individual-level, the Napoleon effect is in fact socially constructed by the different users depending on their context. This can explain the disparities between organisations when they adopt a new medium. As we pointed out earlier, the use of a new medium can have different and even contradictory effects depending on the context.

We also demonstrated that the layering process can improve working methods. In this case, it necessitates the involvement of management and a rethinking of the coordination mechanisms. This may provide solutions to companies aiming to improve their use of communication means.

Our study was conducted on a set of communication means that included e-mail, meetings or face-to-face and phone calls. Future studies could examine a more diversified set of communication means. In addition, our focus on the Napoleon effect is static. Further research could examine the Napoleon effect at different points in time. A future line of research could be to analyse how the Napoleon effect would evolve with the introduction of new ICT tools such as mobile technology.
References


# INTERNET SOCIAL NETWORKING – DISTINGUISHING THE PHENOMENON FROM ITS MANIFESTATIONS IN WEB SITES

<table>
<thead>
<tr>
<th><strong>Journal:</strong></th>
<th><em>17th European Conference on Information Systems</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manuscript ID:</strong></td>
<td>ECIS2009-0647.R1</td>
</tr>
<tr>
<td><strong>Submission Type:</strong></td>
<td>Research Paper</td>
</tr>
<tr>
<td><strong>Keyword:</strong></td>
<td>Social Networking (e.g. Facebook, second life), Web 2.0, Virtual community, Computer-mediated communication (CMC)</td>
</tr>
</tbody>
</table>
INTERNET SOCIAL NETWORKING – DISTINGUISHING THE PHENOMENON FROM ITS MANIFESTATIONS IN WEB SITES

Richter, Daniel, University of Liechtenstein, Fürst-Franz-Josef-Straße, FL-9490 Vaduz, Liechtenstein, daniel.richter@ercis.de

Riemer, Kai, European Research Center for Information Systems (ERCIS), University of Münster, Leonardo-Campus 3, D-48149 Münster, Germany, kai.riemer@ercis.de

vom Brocke, Jan, University of Liechtenstein, Fürst-Franz-Josef-Straße, FL-9490 Vaduz, Liechtenstein, jan.vom.brocke@hochschule.li

Große Böckmann, Stefan, University of Liechtenstein, Fürst-Franz-Josef-Straße, FL-9490 Vaduz, Liechtenstein, stefan.grosse-boeckmann@ercis.de

Abstract

Social Networking Sites (SNS) are one of the most popular business models on the Internet at the moment. At the same time, Social Networking is increasingly interesting as a topic of research in Information Systems. Drawing on existing research in the field, in this paper we propose to distinguish ISN (Internet Social Networking) as a phenomenon from its concrete manifestations in the various SNSs in the marketplace. On the basis of this distinction we take to the classification of SNSs grounded in real-life marketplace variety. In doing so, we identify seven different classes of SNSs. We argue that a typology of SNSs is useful for shaping our understanding of the diverse nature of ISN as existing in concrete manifestations. Most importantly, our classification makes accessible existing research for conceptually sound meta-analysis research. In order to fully grasp the phenomenon of ISN we also propose to include in the definition web sites that feature only certain aspects of ISN, while networking is not their core feature. Using our classification we discuss future research directions.

Keywords: Social Networking (e.g. Facebook, second life), Web 2.0, Virtual community, Computer-mediated communication (CMC).
1 INTRODUCTION

Social Networking is not a new phenomenon. But with the rise and widespread diffusion of the Internet and its enabling technologies, the topic has gained significant momentum over the past years. Since the emergence of the first web sites that supported the creating and inter-linking of user profiles in the late 1990s (Boyd & Ellison, 2007) social network sites have mushroomed. Today, Social Network Sites (SNSs) are among the most frequently visited sites on the Internet. In popular traffic statistics, SNSs consistently rank among the top sites just behind the ubiquitous search engines (Alexa.com, 2008). Social network sites such as MySpace (www.myspace.com) or Facebook (www.facebook.com) account for more than one hundred million members (Schonfeld, 2008; Zuckerberg, 2008). Moreover, various new, smaller and more specialised SNSs have emerged over the past few years, further signifying the success of this new business model (Costa, 2008; Green, 2008). Also, a new industry is currently developing around the phenomenon; some providers deliver technologies (like www.elgg.org) or even full services (like www.ning.com) that enable web site providers to integrate their own social network services, which essentially makes possible and drives the emergence of a plethora of new specialised SNSs. Consequently, thousands of SNSs have been created over the last year alone, providing a networking-space for smaller communities, which are for example characterised by geographical proximity or special interests.

At the same time as SNSs have emerged on the Internet, the phenomenon has gained increasing attention from the scientific community. As a consequence, a considerable body of research on SNSs has been created (Boyd & Ellison, 2007). Studies have researched different social network sites (e.g. Choi, 2006; Schaefer, 2008), have focused on different aspects of Social Networking on the Internet (e.g. Ellison & Steinfield & Lampe, 2007) and shaped our understanding of the specifics of SNSs (e.g. Donath & Boyd, 2004; Kreps, 2008), such as impression management (profiling) (e.g. Lampe & Ellison & Steinfield, 2007; Rosen, 2007), privacy issues (e.g. Govani & Pashley, 2005; Gross & Acquisti, 2005) or the proliferation of special interest networks (e.g. Mellins, 2008; Pledger & Howard & Thomas & Reitberger, 2008). However, in doing so, existing research is most often quite focused: studies are either limited on researching a single SNS (most often Facebook) or they concentrate on one or few aspects of Social Networking. Few studies have carried out comparisons across cases. Mostly, these are limited to three (e.g. Ahn & Han & Kwak & Moon & Jeong, 2007; Dwyer & Hiltz & Passerini, 2007) or two cases (e.g. Ahn et al., 2007; Kumar & Novak & Tomkins, 2006). Some studies also compare users with none-users of SNSs (e.g. Hargittai, 2007; Ofcom, 2008).

While a lot of empirical work has been carried out in the field, few efforts have been made to align or integrate existing research findings. Research to date remains rather scattered, not only across various studies, but also across communities. This limits its impact in providing a better understanding of the general phenomenon of Social Networking on the Internet. At this point in time, more research is needed that takes stock of what has been achieved so far and that also tries to elicit general streams of understanding from the existing body of work by comparing the findings of existing research.

Against this backdrop, we argue that we need to develop a more differentiated conceptual understanding of Social Networking on the Internet. More specifically, we propose to distinguish between the phenomenon of Internet Social Networking (ISN) and its concrete manifestations in existing Social Network Sites (SNS). We argue that current definitions do not provide a differentiated enough understanding, which however is necessary to fully grasp the phenomenon and the variety of manifestations in the respective web sites (Beer, 2008). For example, if we want to draw conclusions from comparing the findings of existing studies, we need to understand the variations in the underlying manifestations of the phenomenon as influenced by the concrete SNSs. Today, even though similarities can be found in the findings of studies based on specific SNSs, these results can not be readily compared or generalised (Hargittai, 2007). Essentially, such comparisons are problematic without a good conceptual understanding of the existing diversity in manifestations of ISN across different SNS. For example, we would expect that user behaviour, and thus the proliferation and role
of phenomena such as user profiling, turns out quite differently on SNSs like LinkedIn (a business-related SNS) when compared to Facebook (a SNS with a focus on students).

Henceforth, we propose that we need to derive a conceptual understanding of the existing variation in manifestations of ISN across different types of SNSs. Essentially, by identifying different types of SNSs we are able to take into account the various SNS characteristics in setting up future research studies. Therefore, in this paper we set out to deriving a typology of SNSs as the basis for future meta-analysis type research and for new cross-case empirical studies. In doing so, our main argument is that user behaviour on SNSs can vary greatly depending on the characteristics of the site such as domain (O'Murchu & Breskin & Decker, 2004) or target group focus (compare Friendster example in Boyd, 2006b), the culture of the member community (compare MySpace example in Rosen, 2007 or Friendster example in Boyd, 2006b), or the role of Social Networking Features (SNFs) on the site.

More specifically, if we want to deduce findings regarding ISN as a phenomenon from researching (a set of) specific SNSs, we need to understand the specific characteristics of the latter and their possible impact on user behaviour. In essence, we have to be able to describe and analyse the concrete manifestation of ISN on a specific social networking platform. We argue that, for doing so, we firstly need a set of criteria to suitably describe SNSs and to further distinguish SNSs into different classes in order to branch out future research in an increasingly complex and diverse field.

Our paper proceeds as follows. In section two, we begin by providing a brief discussion of existing definitions of SNS. Based on this we set out to distinguishing ISN from SNSs. Moreover, we give a short overview of existing research in the field. In section three we derive, based on extant literature, a set of criteria for describing SNSs. This is followed by a classification of a set of 60 social network sites on the basis of which we arrive at a typology of SNSs featuring seven distinct types. In section four we discuss the types and derive distinct sets of research questions for each of them, which illustrates not only the value in investigating ISN phenomena across a set of different types of SNSs, but also provides possible directions for future research. We conclude the paper with a short summary.

2 LITERATURE REVIEW

We propose to distinguish between the phenomenon of Internet Social Networking (ISN) and its concrete manifestations in existing Social Networking Sites (SNSs). In the following, we will first define ISN and give examples of observations on this phenomenon and then derive an SNS definition.

2.1 Internet Social Networking

ISN can be understood as the phenomenon of Social Networking on the Internet. Hence, the concept subsumes all activities by Internet users with regard to extending or maintaining their social network. We draw on social network analysis theory for further characterising the concept (Carton & Wellman, 1999). A social network thus is defined as a set of individuals who establish with each other links of some kind, such as acquaintance or friendship (Newman, 2003). As such, the individuals and their activities in the social network are interdependent and the linkages are channels for transfer of immaterial resources (Wasserman & Faust, 1994). Hence, our unit of analysis is not the individual, but the collection of individuals represented in the social network as well as the linkages among them.

Existing ISN research typically investigates phenomena such as the self-presentation of people by way of SNS profiles (e.g. Kreps, 2008; Lampe et al., 2007) or friends lists (e.g. Donath & Boyd, 2004; Rosen, 2007) or phenomena such as social browsing and searching (e.g. Lampe & Ellison & Steinfield, 2006), which describe the ways in which users utilize their networks of friends. Other phenomena, more concerned with the utility of ISN, are the maintenance of weak tie networks (e.g. Paul & Brier, 2001; Schaefer, 2008) or the initiation of new contacts through web sites (e.g. Ellison & Heino & Gibbs, 2006; Thew, 2008). In doing so, research into ISN is typically targeted at observations made in SNSs, as the concrete manifestation of ISN on the Internet.
2.2 Social Network Sites as manifestation of ISN

Boyd and Ellison define Social Network Sites as “web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system. The nature and nomenclature of these connections may vary from site to site” (Boyd & Ellison, 2007). While others have criticised this definition on the grounds that it is too wide and includes all sites that feature social network of any kind (and not just as core features) (Beer, 2008), we agree with the wider interpretation by Boyd and Ellison. Otherwise, research into many sites offering SNFs as non-core features would be left out.

We argue that research on sites where SNFs are not the core features (e.g. Holme & Edling & Liljeros, 2004; Kumar et al., 2006; Maia & Almeida & Almeida, 2008) can contribute significantly to a better understanding of the phenomenon of ISN in general. It should be taken into account that ISN might manifest in a different way on these sites. Examples are the Chinese SNS QQ (www.qq.com), the most popular Korean SNS (Choi, 2006) Cyworld (www.cyworld.com) and the French SkyRock (www.skyrock.com), which have started as Instant Messaging Service (QICQ), discussion forum or Weblog respectively. Hence, research into these sites might contribute to understanding how a social network starts to form and how this is reflected in the concrete usage of the site. One popular example today is YouTube (www.youtube.com), the world’s leading video sharing website. Also, at Youtube the SNFs are mostly focused on the development, presentation and community rating of video content (Maia et al., 2008). Therefore it offers a very specific view on ISN as a phenomenon.

Henceforth, we follow the above definition and regard all websites that implement any features for enabling Social Networking as SNSs. In this definition all sites are included that support/enable Social Networking regardless of whether it is the core/defining or a none-core feature of the site. It is worth mentioning that no consensus exists in the literature with regard to how narrow or wide the term SNS should be defined (e.g. should it include technologies such as weblogs). We argue that our proposed distinction between the phenomenon of ISN and SNSs can help to overcome this problem as it offers the opportunity to classify both core and none-core SNSs. By classifying SNSs using a set of criteria such as domain (O'Murchu et al., 2004), target group focus (Boyd, 2006b), the culture of the member community (compare MySpace example in Rosen, 2007), or the role played by SNFs on the site, the impact different SNSs and their specific characteristics have on the phenomenon of ISN can be investigated more specifically. By spelling out the specifics of different manifestations, i.e. different types of SNSs, research into phenomena such as self_representation, social grooming or social browsing might gain credibility. Comparison studies using cross-case approaches are then able to compare ISN across different types of SNSs.

In the following section we will draw on the existing body of research on SNSs. We will show that research is fragmented and few efforts have been made to align observations made based on different SNSs. In doing so, we will motivate the necessity of a classification of SNSs.

2.3 Fragmented research in an emerging research field

Existing research is most often characterised by a focus on one specific SNS (e.g. Schaefer, 2008). Most studies also concentrate on specific aspects of this SNS, like the self_presentation in profiles (e.g. Lampe et al., 2007; Liu, 2007), the value of friendship_links (e.g. Fono & Raynes_Goldie, 2006; Rosen, 2007), the network_structure (e.g. Holme et al., 2004; Maia et al., 2008), the relation to offline networks (e.g. Choi, 2006; Ellison et al., 2007) or security/privacy issues (e.g. Govani & Pashley, 2005; Gross & Acquisti, 2005). Only some studies research multiple SNSs (e.g. Ahn et al., 2007; Dwyer et al., 2007; Lenhart & Madden, 2007) or take a more general view on the phenomenon of ISN (e.g. Donath & Boyd, 2004; Kreps, 2008).

Even though similarities can be found in the results gained with different SNSs these results cannot be easily generalized (Hargittai, 2007). The existing body of research might entice scholars to mixing the
results gained from the various, often very different, SNSs in order to form a general understanding of ISN, which however is likely to lead to problematic conclusions. The vast majority of SNS studies deal with Facebook. Other SNSs that have been repeatedly researched are MySpace (e.g. Boyd, 2006a; Dwyer et al., 2007), CyWorld (e.g. Choi, 2006), LinkedIn (e.g. Thew, 2008) and Xing (www.xing.com) (e.g. Schaefer, 2008). Only little research exists on other SNSs. Such a concentration on very few SNSs potentially prohibits identification of specific kinds of behaviour that might only exist on certain other SNSs.

Comparing the results of studies on SNSs it becomes obvious that the ISN phenomenon does not manifest in the same way across SNSs, but that significant differences can be observed in site usage. For example, people seem to use business and leisure-related SNSs not only for different reasons, but also in a different ways. On business-related SNSs for example the presentation of oneself can serve as a potential gateway to new employments or business contacts (King, 2006). Therefore, self representation is done very carefully (Schaefer, 2008). This does manifest in the profile description as well as in the choice of contacts (Thew, 2008). The usage of leisure-time SNSs in contrast can be characterised as being much more playful (Sledgianowski & Kulviwat, 2008). Profiles are filled up with information about favourite music, artists or TV-shows, personal heroes and every kind of pictures (Liu, 2007). The choice of contacts is much more unstrained. The search for a romantic partner by users of dating-related SNSs again can result in a different form of self-presentation; users tend to present their “ideal self”. Due to risks of misrepresentation, users also adopt various strategies to proof the credibility of their profile (Ellison et al., 2006). This implies that a comparison of results on the usage of different SNSs needs a conceptual understanding of their differences.

All in all, we face a highly fragmented marketplace of SNSs. Sites are characterised by different focus and functional settings, by an integration in different real world communities, and by a significant variance in user count. This fragmentation, heterogeneity and diversity are not mirrored sufficiently in research yet. Moreover, current research seems to be unaware of this diversity. In contrast, current research seems to treat all SNSs as one of a kind. While existing research has pioneered the field in achieving a first overview of the marketplace and initial understandings of ISN phenomena, we think that now is the time to take stock of the actual variety of SNSs existing in the marketplace in order to take ISN research to the next level. In the next section we carry out a classification of SNS, before we discuss the role of our classification in future research.

3 CLASSIFICATION OF SOCIAL NETWORK SITES

For classifying SNSs we first analysed the current body of research in order to derive criteria for differentiation. Subsequently, we identified SNSs in the marketplace and applied the criteria to them. We were able to derive six criteria for classification and we identified seven different classes of SNSs.

3.1 Criteria for Classification

Based on an extensive literature review, we identified six criteria for classifying SNSs, all of which are briefly described in the following sections.

**Relationship Notion:** One core element of SNSs is displaying the connections that exist between users (Donath & Boyd, 2004). In doing so, the relationships between two users do not have the same label or the same notion across platforms. Popular labels for relationships are for example “friends” or “contacts” (Boyd & Ellison, 2007). But more importantly these relationships can reflect very different understandings of who is a friend or contact online. The meaning of friends online and offline can be both diametric or exactly the same (Boyd, 2006a). Even though the relationship notion might be specific to every single user (Ofcom, 2008; Schaefer, 2008) it is also highly dependent on the SNS. Two major aspects have to be mentioned: (a) the culture on the SNS and (b) the main purpose of the SNS. The emerging culture (a) found on an SNS can shape the understanding of the value of a relationship. On MySpace for example, encouraged by the system functionality, the collection of as
many friends as possible is much more common than on most other SNSs (Rosen, 2007). Differences in the interpretation of ‘relationships’ have also been observed across SNSs with different purposes (b), like business, leisure-time, dating or online-gaming SNSs (Boyd, 2006a; O'Murchu et al., 2004). Relationships can be characterised as business-partners or colleagues on business-related SNSs, or as friends or acquaintances on leisure-time related SNSs. On online gaming-related SNSs relationships are more with ‘playfellows’ than real-life friends. The characteristic of “relationship notion” thereby offers to differentiate between SNSs according to different interpretations of relationships influenced by site purpose and emerging culture.

Purpose of Usage: Whether the intrinsic motivation for SNSs usage is largely hedonistic, altruistic or utility-based, is expressed in the characteristic “purpose of use”. Whereas SNSs like Facebook or MySpace are mostly used for hedonistic purpose (Sledgianowski & Kulviwat, 2008), SNSs concentrating on business matters like LinkedIn or Monster.com have a clear focus on utility (Thew, 2008). They can act as mediators for job assignments or useful business contacts. The usage of SNSs that focus primarily on information exchange or the dissemination of multimedia content is mostly hedonistic in nature. Exceptions are the usage of SNSs concentrating on health issues, being utility based, or of SNSs seeking for social good like TakingItGlobal.org, which is rather altruistic in nature.

Role of SNFs: SNSs can be differentiated by SNFs being the core features, like in the case of Facebook, MySpace or LinkedIn, or none-core features like on YouTube or Flickr (Beer, 2008; Boyd & Ellison, 2007). Sites not implementing SNFs as the core features can use SNFs as added value or as an enabler. Good examples for using SNFs as added value are dating sites whose focus is on the mediation of contacts to potential dates. The building up of a social network among users can provide additional quality information about users as the social network of a person can give insights into that person’s life and personality (Donath & Boyd, 2004). Thereby information on the social network of a person can help identify whether s/he is a possible partner or not. Websites focusing on health issues can use SNFs as an enabler to foster information exchange between professionals and patients (Kamel Boulos & Wheelert, 2007). An example for a professional health network is theijjs.com, an SNS build around the International Journal of Surgery.

Mode of Usage: The way people adopt the functionality of an SNS can vary greatly depending on the context of the site as well as the community represented on it. This can be illustrated by again comparing leisure-time and business-related SNS. In leisure-related SNSs, users act relatively informal and unconcerned regarding their self-presentation (Govani & Pashley, 2005; Liu, 2007; Sledgianowski & Kulviwat, 2008). The messaging functionalities are used intensively and relationship requests are only very rarely denied (Ito & Horst & Bittanti & Boyd & Herr-Stephenson & Lange & Pascoe & Robinson, 2008; Ofcom, 2008). On business-related SNS users seem to care more about self-presentation. The profile needs to be appealing to potential employers or business partners (King, 2006). Messaging functions are only rarely used (Thew, 2008). Usage is concentrated on a favourable self-presentation whereas usage of leisure-related SNSs focuses on interaction with others. Hence, we differentiate between the modes of usage “interaction” and “self-portrayal”.

Target and Domain Focus: The most popular SNSs today count more than one hundred million users (Schonfeld, 2008; Zuckerberg, 2008). SNSs like MySpace or Facebook, limited to a specific audience in the beginning (music/students), opened up to attract more and more users eventually seeking market domination. A large number of users is not only advantageous for the provider, but also for users, who are able to connect to more other users. But this also implies that users can only present one identity to all possible contacts, be it business contacts, parents or co-students. This can have negative effects on the communication taking place within a specific social group on a SNS (Boyd, 2006b). Hence, even though hidden behind the success of the “big players” an increasing number of domain and target group specific SNSs have emerged (Green, 2008). Open source platforms, such as elgg.org, or full service providers, like ning.com allow social groups to easily create their own SNSs. On ning.com alone thousands of SNSs have already been created, concentrating on specific target groups like fire fighters (firefighternation.ning.com) or college students (playboyu.com), or on domains like sports (streeball.com) or gothic living (fairiesvampires.com).
3.2 Classification

Using the above-described criteria we characterized a total of 60 SNSs. Through techniques of grouping and displaying (see Miles & Huberman, 1994), we were able to derive seven main classes/types of SNSs. Figure 1 gives an overview providing four examples for each category.

Public-SNSs do not limit their potential audience by any means. Some started with a focus on domain-topics or target-groups but abandoned these restrictions in favour of further growth. Consequently, the SNSs with the biggest user group (Facebook, MySpace, QQ or CyWorld) can be found in this class. Public-SNSs are core-SNSs that are used for interaction with a hedonistic purpose. The self-presentation is playful (Sledgianowski & Kulviwat, 2008) and the friends online do often match with friends offline (Ellison et al., 2007; Paul & Brier, 2001) even though this differs slightly between the certain sites of this class (Rosen, 2007).

Getting into contact with potential business-partners and looking for employments or employees are the predominant usage patterns on Business-SNSs (Purpose of Use: utility). These SNSs are less used for communication. Characteristic is a very careful design of the self-presentation concerning the profile as well as the contact list (Schaefer, 2008; Thew, 2008). Contacts can be real-world contacts, distant acquaintances or even users not known from the offline world (Thew, 2008). SNFs are usually the core features with the exception of recruiting sites, which use SNFs as additional features. The most popular SNS in this class is LinkedIn.com with more than 30 Million users.

Content-SNS focus on the production, sharing and discussion of content such as pictures (Flickr.com), video (YouTube.com), music (imeem.com) or stories (tokoni.com). The manifestation of ISN on this type of SNS has not been widely researched as yet. SNFs on the sites are mainly used for enriching the discussion on the content (role of SNFs: adding).

Target-SNSs are targeted at certain real-world groups like students (studiVZ), mothers and mothers-to-be (cafemom.com) or people of a certain age (like platinnetz.de: over 50/ or clubpenguin.com: children). With the notable exception of focusing on a specific target group these SNSs are very similar to Public-SNSs. The major Public-SNSs Facebook started as a Target-SNSs focused on students until it was opened to a wider public. However, it is to be expected, that the bounded character of social networks will lead to different behaviour on the SNSs, as users tent to feel less observed in such environments (Boyd, 2006b).

Domain-SNS aim at bringing together people on a specific domain-topic, like sports (ruku.com: rugby), travel (passportstamp.com) or health (patientslikeme.com). Like Target-SNS, Domain-SNS are comparable to Public-SNS. But again the focus on a certain topic domain can influence ISN manifestation on the SNS. For example, Domain-SNS concerned with health issues, where patients can find assistance, will most likely be used in very specific ways. Little research has been conducted on Domain-SNSs so far.

Users of Activity-SNSs usually follow a certain goal, which can be meeting a romantic partner, a playfellow for an online game, or business contacts for setting up a project. Hence, Activity-SNSs are strongly grounded in real life and aim at mediating contacts in a specific domain. The purpose of use can vary depending on the domain like charity (altruistic), business (utility) or online gaming (hedonistic). Many Activity-SNSs are none-core SNSs.

Micro-SNSs can be very different in their character and very closely related to other SNS-classes. The major aspect for differentiation is neither a domain focus nor a specific aim, but their size. Micro-SNSs usually only have a few hundred to a couple of thousand users. They are used to accompany conferences (fowa-miami.crowdvine.com), for supporters of a sports team (spiritoff12.com) or are dedicated to students of a specific university (community.brighton.ac.uk). The development of Micro-SNSs is only possible due to providers delivering the technologies or services (see above).
<table>
<thead>
<tr>
<th>Relationship</th>
<th>Purpose of Usage</th>
<th>Role of SNS</th>
<th>Mode of Usage</th>
<th>Target Group Focus</th>
<th>Domain Focus</th>
<th>Public-SNSs</th>
<th>Business-SNSs</th>
<th>Content-SNSs</th>
<th>Target-SNSs</th>
<th>Domain-SNSs</th>
<th>Activity-SNSs</th>
<th>Micro-SNSs</th>
</tr>
</thead>
<tbody>
<tr>
<td>friends, real world</td>
<td>hedonistic</td>
<td>core</td>
<td>interaction</td>
<td>none</td>
<td>none</td>
<td>facebook.com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>friends, often informal</td>
<td>hedonistic</td>
<td>core</td>
<td>interaction</td>
<td>none</td>
<td>none</td>
<td>mySpace.com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>friends</td>
<td>hedonistic</td>
<td>core</td>
<td>interaction</td>
<td>none</td>
<td>none</td>
<td>cyWorld.com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>orkut.com</td>
</tr>
<tr>
<td>business contact</td>
<td>utility</td>
<td>core</td>
<td>self-portrayal</td>
<td>none</td>
<td>none</td>
<td>xing.com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>linkedIn.com</td>
</tr>
<tr>
<td>business contact</td>
<td>utility</td>
<td>core</td>
<td>self-portrayal</td>
<td>none</td>
<td>none</td>
<td>monster.com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>plaxo.com</td>
</tr>
<tr>
<td>business contact</td>
<td>utility</td>
<td>core</td>
<td>self-portrayal</td>
<td>none</td>
<td>none</td>
<td>tokon.com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>imeem.com</td>
</tr>
<tr>
<td>friends</td>
<td>hedonistic</td>
<td>enabling</td>
<td>Interaction, self-portrayal</td>
<td>none</td>
<td>stories</td>
<td>YouTube.com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>flickr.com</td>
</tr>
<tr>
<td>friends</td>
<td>hedonistic</td>
<td>adding</td>
<td>interaction</td>
<td>none</td>
<td>musik</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>friends</td>
<td>hedonistic</td>
<td>adding</td>
<td>interaction</td>
<td>none</td>
<td>videography</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>friends</td>
<td>hedonistic</td>
<td>adding</td>
<td>interaction</td>
<td>none</td>
<td>pictures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>friends</td>
<td>hedonistic</td>
<td>core</td>
<td>interaction</td>
<td>mothers and mothers to be</td>
<td></td>
<td>cafemom.com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>playboy.com</td>
</tr>
<tr>
<td>friends, co-students</td>
<td>hedonistic</td>
<td>core</td>
<td>interaction</td>
<td>college students</td>
<td></td>
<td>studiVZ.net</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>platinnetz.de</td>
</tr>
<tr>
<td>friends, co-students</td>
<td>hedonistic</td>
<td>core</td>
<td>interaction</td>
<td>students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>friends</td>
<td>hedonistic</td>
<td>core</td>
<td>interaction</td>
<td>over 50</td>
<td>none</td>
<td>bodybuilding.com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ruck.com</td>
</tr>
<tr>
<td>friends</td>
<td>hedonistic</td>
<td>core</td>
<td>interaction</td>
<td>none</td>
<td>travel</td>
<td>patientslike.me</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>passport.com</td>
</tr>
<tr>
<td>friends, dates, playfellows</td>
<td>hedonistic</td>
<td>adding</td>
<td>self-portrayal</td>
<td>none, singles</td>
<td>dating</td>
<td>match.com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>world.europe.com</td>
</tr>
<tr>
<td>business contact</td>
<td>altruistic</td>
<td>enabling</td>
<td>Interaction, self-portrayal</td>
<td>none</td>
<td>social good</td>
<td>takingitglobal.org</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>popetent.net</td>
</tr>
<tr>
<td>business contact</td>
<td>utility</td>
<td>adding</td>
<td>self-portrayal</td>
<td>freelance, customers</td>
<td>videography</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>friends, co-fans</td>
<td>hedonistic</td>
<td>core</td>
<td>interaction</td>
<td>none</td>
<td>Seattle Seahawks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>spiritof12.com</td>
</tr>
<tr>
<td>friends, co-students</td>
<td>hedonistic</td>
<td>core</td>
<td>interaction</td>
<td>students in Brighton</td>
<td>university of brighton</td>
<td>Future of Webapps</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>community.brighton.ac.uk</td>
</tr>
<tr>
<td>co-participants</td>
<td>utility</td>
<td>adding</td>
<td>interaction</td>
<td>conference participants</td>
<td></td>
<td>fowa-miami.crowdwise.com</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mykicksonfire.com</td>
</tr>
<tr>
<td>friends</td>
<td>hedonistic</td>
<td>core</td>
<td>interaction</td>
<td>none</td>
<td>Sneakers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Classification of Social Network Sites
4 DISCUSSION

We differentiated between ISN as a phenomenon and SNSs as its concrete manifestations, which can take different forms. In order to explore the latter, we classified SNSs and identified seven classes. With this we seek to build ground for more structured research on 1) the phenomenon of ISN and 2) the specifics of certain SNSs and types of SNSs.

4.1 Research on ISN

Until now, our understanding of the phenomenon of ISN is rather limited. This is despite a considerable body of research on SNSs. We face a highly fragmented marketplace with different classes of SNSs. This fragmentation is not reflected in current research. Rather, SNSs are treated as one of a kind. To better understand the phenomenon we need to understand the specific characteristics of a set of SNSs and their possible impact on user behaviour. With our classification we want to provide a framework for integrating current research results in meta-studies. It is also intended as a means to sensitise future research to the, often quite subtle, differences between certain SNS as object of research. Also, when we want to deduce findings on ISN as a phenomenon from researching (a set of) specific SNS we need to understand their specific characteristics, as they are likely to impact on user behaviour and the concrete manifestation of ISN on the platform.

A recent study by the Office of Communication (UK) classified different types of users and non-users (Ofcom, 2008). One example is the “intellectual rejecter” as a non-user who thinks of ISN as “a waste of time”. However, this might not be true for utility-based SNSs. The different types of users characterised could be a starting point for investigating user types in different SNS classes. Also, an analysis of profile data across the different classes promises to reveal different strategies of SNS usage. Henceforth, whenever we want to compare or draw conclusions from existing research, the classification might be useful as a device for judging, validating and interpreting these results, knowing that the type of SNSs, from which the specific findings have been derived, exhibits certain characteristics. Moreover, each type of SNS (as manifestation of ISN) might offer a unique, different, even novel perspective on ISN as a phenomenon.

4.2 Research on SNS classes

The current body of research is mostly concerned with Public-SNSs (cf. the overview in Boyd & Ellison, 2007). Attention is also paid to Business-SNSs (e.g. Schaefer, 2008; Thew, 2008). On the other types, such as Content- (e.g. Maia et al., 2008), Activity- (e.g. Ellison et al., 2006), Target- (e.g. Byrne, 2007) and Domain-SNSs (e.g. Mellins, 2008; Ploderer et al., 2008) only single studies have been conducted. To our knowledge no research has been conducted so far on the relatively new class of Micro-SNSs. In the following, we present some directions for further research on the SNS classes.

Recent traffic statistics on the main Public-SNSs like Facebook and MySpace show, that the number of page views per user is decreasing on these sites (Alexa.com, 2008). This is in contrast to the constant growth in user numbers and reach, which is in line with the platforms’ search for market domination (Schonfeld, 2008). However, it has been argued that the growth of the social network, especially in new user groups, can be a threat to SNSs (cf. Facebook example in Boyd, 2006b). However, it is yet to be seen how the constant growth in user numbers might ultimately change the main Public-SNSs. Research also needs to be conducted on how growth influences user behaviour. Another issue most prominently discussed in the context of major Public-SNS is one of security (recent examples are: Boyd, 2008; Felt & Hooimeijer & Evans & Weimer, 2008; Ybarra & Mitchell, 2008). Users present large amounts of data on themselves without fully comprehending what they are doing (Govani & Pasley, 2005; Gross & Acquisti, 2005). The benefits SNSs provide to their users are very closely interlinked with their potential abuse. Even though a considerable body of research exists on this topic, more research is needed in order to understand how the benefits can be achieved, while minimising the potential risks at the same time.
Users seem to be more careful in their self-presentation on Business-SNSs. This observation should be substantiated by investigating the more concrete differences in the adoption and use of SNFs, for example with regard to the information published in profiles on Business-SNSs in comparison to other SNS classes. Furthermore, Business-SNS claim to improve inter- and intra-company communication in providing contact to potential business-partners. However, this claim still warrants verification.

In Content-SNSs SNFs are provided as additional features and therefore it can be expected that they are used differently compared to core-SNSs. Also the concrete usage patterns are likely to differ depending on the content that is shared on the Content-SNSs. The variety of possibilities for using SNFs as an added functionality and how usage behaviour is affected thereby should also be investigated. Serving a real-world group, Target-SNSs can provide insights on processes of transferring real world contacts online. Especially interesting is a comparison of the behaviour of users on Target-SNSs and other SNSs. Potentially, self-presentation might be less important on such SNSs, when taking into account that the online contacts are often already known from a real world context. Profiles might be more ironic and playful drawing on the knowledge existing from real world contexts.

Until now, only a few findings concerning the adoption of SNFs in Domain-SNSs can be identified. It is unclear whether Domain-SNSs are similar to Target-SNSs or whether site usage is mainly focused on discussing domain matters. It is likely that within the class of Domain-SNSs considerable differences in the adoption might be observable, considering the variety of possible domains (health issues vs. sports discipline). Especially on health related Domain-SNSs the quality of exchanged information and of services should be analysed and threats of misuse should be identified. Addressing a very important part of life, research should search for evidence whether the usage of health-related Domain-SNSs can be helpful or might even be problematic.

Theoretically, Activity-SNSs have a high potential for creating new real-world relationships as users seek to find partners for certain activities. This characteristic of Activity-SNSs can bear significant risks in light of the current security discussion; a recent homicide case was linked to the usage of such an SNS (Ulrich, 2008). Further research might investigate strategies for minimising such risks for the users. With regard to the ISN phenomenon, Activity-SNSs can give insights into different strategies for attracting potential partners. These strategies might in turn influence the design of user profiles, as well as the communication behaviour. While some research exists, in-depth analyses are still missing.

Micro-SNSs have only a relatively small user group. Nevertheless, users take the effort of maintaining an additional profile even for this limited audience. The underlying motives might be subject of future research. However, it is still unclear whether the proliferation of Micro-SNSs is only a short-lived phenomenon. One possibility might be that some survive and grow and others just disappear. A second scenario could be that usage of interfaces (like provided for example by elgg.org) leads to an interconnection of various Micro-SNSs, forming a community comparable to major Public-SNSs, but more diverse in terms of user base.

5 CONCLUSION

We proposed to distinguish Internet Social Networking (ISN) as a phenomenon from its concrete manifestations in the various Social Network Sites (SNSs) in the marketplace, through the implementation of specific Social Networking Features (SNFs). We provided definitions for ISN and SNSs and more importantly proposed a classification of SNSs grounded in real-life marketplace variety. We derived seven SNS-classes: Public-SNSs, Business-SNSs, Content-SNSs, Target-SNSs, Domain-SNSs, Activity-SNSs and Micro-SNSs. Finally, we identified future research directions for both the phenomenon of ISN in general and with regard to the various SNS classes. Most importantly, meta-research seems appropriate in order to align the already existing findings. Consequently, we hope that our paper might indirectly contribute to a better understanding of the ISN phenomenon and that our classification is useful as a sensitising device in comparing research on various SNSs. Our classification should make existing research accessible for conceptually sound meta-analyses. Our characterization of SNSs is based on the current body of literature and an analysis of the marketplace.
While our classification provides a state of the art overview of the SNS market, it is likely to further change and differentiate with the further proliferation of the marketplace. The market of SNSs is still a “moving target”, therefore our classification needs to be revised and updated from time to time.

References


CRIMINALISING FANTASIES: THE REGULATION OF VIRTUAL CHILD PORNOGRAPHY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0536.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Critical theory, Ethics, Information society, Regulation / Deregulation</td>
</tr>
</tbody>
</table>
Abstract

Law enforcement agencies are currently concerned with the increasing availability of virtual child pornography. Virtual child pornography refers to computer-generated images where no real child is involved, but which nevertheless raise important ethical and legal issues. There is an ongoing debate whether the possession of virtual child pornography should be criminalised. As this paper will show the criminalisation of such images is a controversial issue with arguments for and against making it a criminal offence. This paper aims to contribute to the debate by offering a definition of virtual child pornography and considering the legal and policy arguments that are put forward as justifications for the regulation of such material. This paper contributes to the theoretical field of critical information systems research by highlighting how the main interest of critical research, namely emancipation can be furthered or hindered by legal developments. We will explore how the idea of emancipation can be applied in the area of virtual child pornography. Furthermore the paper aims to contribute to the debate of this topical issue and thereby support policy developments on a national and international level.

Keywords: Virtual child pornography, regulation, critical information systems research, emancipation
1 INTRODUCTION

Recently there has been concern in the media that those with a sexual interest in children are ‘acting out’ their fantasies in virtual worlds such as Second Life. For example, in an area within Second Life, called ‘Wonderland’, young “children” were offering sex in a playground. The young children were in this context not real children, but graphical representations and the playground was a virtual playground created with computer software in this technological world. This incident demonstrates one form of virtual child pornography, namely computer-generated images (CGI), which is created wholly through computers and does not involve the abuse of real children but nevertheless raise important legal and ethical issues (Oswell, 2006). The term virtual child pornography refers to CGI drawings, paintings and cartoons portraying sexual representation of children. It should be noted that ‘virtual child pornography’ and ‘pseudo-photographs’, also called ‘morphed images’, are not the same thing. Pseudo-photographs consist of a photograph or a collage of photographs manipulated with digital techniques (Taylor and Quayle, 2003). As Gillespie (2008) points out, the quality of virtual child pornography varies strongly but some of the CGI appear to be of ‘photo-quality’ and can hence appear to be very realistic. It also highlights the importance of placing virtual child pornography in its wider context where technological developments enable easy sophisticated production of this kind of material. In addition to the Second Life incidents concerns are expressed by the law enforcement agencies and welfare organisations about the increasing availability of virtual child pornography, which has given rise to the question whether the possession of this kind of material should be criminalised (Home Office, 2007). Ireland is one of the few countries to have introduced legislation specifically encompassing the production and possession of virtual child pornography (sections 5 and 6, Child Trafficking and Pornography Act 1998) although it has rarely been used in practice. The Swedish legislation regarding child pornography criminalises the production, distribution and possession of ‘virtual’ child pornography through its formulation ‘anybody that depicts a child in a pornographic image’, ‘distribute such image’ or ‘possess such image’1. The UK government recently launched a consultation process suggesting that the law in England and Wales could be altered to criminalise possession of such material (Home Office, 2007). The government highlights in their report that there is no existing evidence suggesting a direct link between possession of this kind of material and the increased risk of sexual offending against children (Home Office, 2007) whereas some researchers (Quayle et al, 2008) have noted that virtual child pornography such as CGI can be used as part of the solicitation process and that the production, distribution and possession of child pornography, whether virtual or not, could be considered an exploitation of children. While most people agree that child pornography should be illegal there are critical voices raised about the criminalisation of virtual child pornography that does not demonstrate a direct link of harm to children (Williams, 2004). One of the main arguments in the debate regarding this issue is whether it is appropriate for the law to regulate peoples fantasies and imaginations through the creation of a ‘thought crime’ (Home Office, 2007) or whether this is a threat to the individual’s privacy and right of freedom of expression. Privacy and freedom of expression, although important civil liberties, have been in the focus of the civil liberties and technology debate for more than a century (Wong, 2005). One can however argue that they should not be considered as absolute rights and that they should be balanced with other rights, and not take the precedence over other rights, such as the right of the child not to be exploited or abused (Eneman, 2005).

The above references to current legislation and consultation processes in different countries show that virtual child pornography is considered an important issue. It seems to be intrinsically linked to the abuse of children, which is one of the most problematic areas of misuse of information and communication technologies (ICT). Virtual child pornography touches on deeply held moral

---

1 The prohibition of such depiction does not include drawings or paintings produced for artistic merit. For further information see the Swedish Code of Statutes, the Penal Code, Chapter 16, Crime against public order, 10 a §.
convictions, where subjective opinion and moral position tend to dominate the debate rather than rational arguments (Taylor and Quayle, 2003; Sandywell, 2006; Luck, 2008). This affects the way we perceive technology and its use in society with such tensions being evident in the desire for technology-related legislation (Gillespie, 2008) which means it is a playing field for social conflicts and negotiations. Even though we know that ICT is used for harmful content such as the dissemination of child pornography (Eneman, 2009; Gillespie, 2008), little academic research exists in relation to the use of ICT to produce and disseminate harmful content. The specific subject in this paper, virtual child pornography, has not previously been researched using a critical lens in the area of information systems or ICT. Adam (2005) emphasizes the importance of adding a critical perspective when discussing crimes against children such as child pornography, to avoid that subjective moral opinions take precedence over rationale arguments in the debate. Finally, it is an important area for critical investigations because it relies on a number of assumptions that are worth exploring and it is related to power relationships that call for a more detailed analysis. This paper addresses therefore the question of virtual child pornography from a critical angle (Stahl, 2008) and aims to contribute to the debate by offering a definition of virtual child pornography and considering the legal and policy arguments that are put forward as justifications for the regulation of such material.

This paper contributes to the theoretical field of critical information systems research (CISR) by highlighting how the main interest of critical research, namely emancipation (Alvesson and Deetz, 2000; Cecez-Kecmanovic, 2005; Stahl, 2008), can be furthered or hindered by legal developments. We will explore how the idea of emancipation can be applied in the area of virtual child pornography. This will entail a discussion of the stakeholders whose interests are touched upon. We will ask what emancipation can mean (Stahl, 2008) in the context of virtual child pornography and how the current legal framework and planned developments thereof are conducive to it. Furthermore the paper aims to contribute to the debate of this topical issue and thereby support policy developments on a national and international level. This paper is based on an analysis of the consultation paper (Home Office, 2007) examining whether England and Wales should criminalise possession of virtual child pornography. In our analysis of this document we have focused upon the arguments used for and against criminalising possession of virtual child pornography. This paper is organised as follows: in the next section the theoretical framework, critical information systems research (CISR), is presented with a focus on the theoretical concept of emancipation. Section three describes the phenomenon virtual child pornography. This is followed by a discussion in section four and then the paper is closed by a conclusion in section five.

2 THEORETICAL FRAMEWORK

An emerging research field within the IS discipline, relevant for this paper, is critical information systems research (CISR) (Howcroft and Trauth, 2005). CISR can be seen as a reaction to the mainstream IS research which tends to assume that technological innovation is ‘inherently desirable’ and beneficial to all (McGrath, 2005). The use of ICT for the creation of virtual child pornography is an illustrative example of how ICT can be used for harmful purposes and raises challenges for society (Gillespie, 2008). CISR is based on ‘critical theory’ (Klein and Huynh, 2004; Croon-Fors, 2006). ‘Critical theory’ is not a unified theory ‘but rather a set of loosely linked principles’ (Klecun, 2005), with a commonality ‘to change the status quo and promote emancipation’ (Alvesson and Deetz, 2000; Stahl, 2008). Critical IS researchers use a wide spectrum of critical social theories (for example Bourdieu, Foucault, Habermas) to critically question established assumptions about the technology, its use and its implications (Cecez-Kecmanovic, 2005, Willecocks, 2006). Engaging in CISR entails the study of the research object with the aid of concepts relevant to critical theory, for example domination, power, control, emancipation etc (Cecez-Kecmanovic, 2005). In addition to using relevant concepts, the research object should be placed in a wider historical, political, social and economic context (Alvesson and Deetz, 2000). By critically questioning ‘social realities’ and providing alternative insights how these ‘realities’ are historically, politically and socially constructed and strongly shaped by asymmetries of power in society, we are able to move beyond established
definitions and assumptions and can achieve emancipation from traditional existing structures (Cecez-Kecmanovic, 2005; Alvesson and Deetz, 2000). The central aim in critical IS research is its deep interest of emancipation (Stahl, 2008; Alvesson and Deetz, 2000). Stahl (2008) highlights however that there are fundamental problems involved with the critical intention to emancipate, for example when the research subjects don’t prefer to be emancipated. The emancipatory perspective in this study refers to emancipation on a societal as well as an individual level.

It is important to note that there are numerous other critical movements that refer to particular areas of research. With regards to our topic of virtual child pornography which is sometimes subject to legislation, it is worth pointing out that an established tradition of critical studies in the area of law exists (cf. Kelman, 1987; Unger, 1986; Mansell et al., 1999). These critical legal studies have a similar interest in emancipation and build on a similar set of theories, but they often emphasise other aspects, which we cannot reflect in full in this paper. Our aim here is not to give a comprehensive review of the critical literature but to show that a critical angle can provide a better understanding of the social consequences of ICT. Using the central concept of emancipation and related critical topics, such as ideology or rationality, we hope to show that seemingly uncontentious issues such as legislation of a morally repugnant act such as the production or possession of visual child pornography are in fact complex and deserve detailed attention. We use some of the thoughts of critical theory and investigate in how far they enlighten our understanding of virtual child pornography and its regulation. Our aim is not to come to a final judgment on whether or not virtual child pornography should be banned but rather to show how critical issues could be framed. The result of this will be a wider understanding of some of the problems associated with ICT use in society, which may influence public perception as much as policy making and further research.

3 WHAT IS VIRTUAL CHILD PORNOGRAPHY?

It is necessary to define what virtual child pornography is. This is not necessarily easy because, to an extent, “child pornography” itself is not readily defined, even within International law. The UN Optional Protocol to the Convention on the Rights of the Child on the Sale of Children, Child Prostitution and Child Pornography has perhaps the widest definition when it talks about any representation “by whatever means” of a sexual representation of a child. This could conceivably include non-visual depictions, for example text or audio files. This would tally with some definitions of child pornography produced by, for example, Interpol (see Gillespie, 2008). Other International instruments are more restrictive and focus on, for example, visual depictions (see, most notably, the Council of Europe Convention on Cybercrime, Council Framework Decision 2004/68/JHA on Combating the Sexual Exploitation of Children and Child Pornography and the Council of Europe Convention on the Protection of Children Against Sexual Exploitation and Sexual Abuse) and most legal systems similarly base themselves on visual depictions. Further problems with the definition of child pornography include the age of a child which has traditionally differed between jurisdictions (usually based on the age at which a child can consent to sexual intercourse). Whilst some international instruments continue to adopt this approach, others (most notably the UN Optional Protocol and Council Framework 2004/68/JHA) reject this approach and suggest that “child” should mean any child under the age of majority (universally set at 18 under the UN Convention on the Rights of the Child). When the Council Framework is fully implemented this will have the effect of standardising the age of a “child” to 18 across the member countries of the European Union (Gillespie, 2008 but see Gillespie, 2004 for a discussion on the propriety of such a change).

Until comparatively recently, visual depictions meant, in essence, photographs but the development of ICT blur the boundary between photographs and other forms of realistic visual depictions. Child Pornography has traditionally been seen as the record of child sexual abuse or exploitation of a real child (Taylor and Quayle, 2003) but advances in ICT meant that the depictions may not necessarily involve a real child or show it being abused. For example, morphed images became increasingly recognised. A morphed image is where a real photograph is digitally altered. This can take a number
forms and may continue to involve the exploitation of a child. For example, the alteration could involve superimposing the face of a child onto an (adult) pornographic picture, or removing the image of a child (for example eating an ice-cream) and superimposing it onto a pornographic picture so as to make it look like the child is participating in a sex act. Such images can still be considered sexually exploitative since although no abuse has necessarily occurred to the child, it remains exploited since a recognisable picture of the child is displayed purporting to show it engaging in sexual activities.

A second form of morphing is to take an adult pornographic image and, using a graphic manipulation package, transform the image into a child. So, for example, the hips of the female could be slimmed, breast size be reduced and pubic hair airbrushed out or thinned. The resulting image would appear to show a child. Again, where it shows a recognisable person this could be said to be exploitative since a person recognising the person portrayed may think (wrongly) that the person posed sexually whilst a child.

Criminalising such exploitative images is perhaps less problematic since they amount to an interference in the personal integrity of the subject shown. However ICT also allows for the creation of purely-fictitious images. Complex and realistic Computer Generated Images (CGI) of children can be produced. Such images differ from morphed images in that a real person is never involved. Whilst it is, of course, conceivable that a real person could be modelled, it is equally likely that a person will simply create a representation of their desired body-type. Accordingly CGI child pornography cannot realistically be considered abusive since there is no ‘real’ child. It is submitted that the lack of exploitation allows for a differentiation to be drawn between morphed and CGI images. Accordingly for the purposes of this paper we will consider “virtual child pornography” to exclude morphed images, and instead to constitute solely those images that are wholly generated by a computer, i.e. a “virtual child”.

The term “virtual child pornography” will be used because although the term “child pornography” has fallen out of favour with many law enforcement and child protection organisations, it remains a term readily used within International law (see the UN Optional Protocol and related European instruments discussed above). It is not necessary for us in this paper to define “pornography” save to state that we are concerned with images that purport to show a “child” involved in sexual activity or an image that focuses on the genitalia of a child (in other words, in effect adopting the definition used by the UN Optional Protocol (Article 2(c)).

3.1 Arguments for and against the criminalisation or legalisation of virtual child pornography

Now that a definition of virtual child pornography has been established it is possible to consider whether virtual child pornography should be criminalised. Most legal systems criminalise the production and dissemination of child pornography and, whilst some have argued to the contrary (see, for example, Ost, 2002), most systems also criminalise the possession of child pornography. Certainly in International law, all four international instruments identified above require member states to criminalise the production, dissemination and possession of child pornography.

The justification for criminalising child pornography is that it is inextricably harmful to children. The production of child pornography requires a child to be abused and those who possess the images are considered to be fueling the demand for the production of new images to be traded (Taylor and Quayle, 2003). Similar arguments exist for morphed images where it can be said that a child or other person will be exploited and that this can cause secondary harm through psychological distress (Palmer, 2005; Gillespie, 2008). However can the same be said to be true of virtual child pornography where there is no real person to be identified or caused either primary or secondary harm?

It is notable that International law differs in its treatment of virtual child pornography. Whilst the UN Optional Protocol undoubtedly includes virtual child pornography through its definition which encompasses the production of visual representations by any means and also expressly includes simulated sex acts, other international instruments are less clear. For example, the Council of Europe
Convention on Cybercrime expressly states that member States can decline to criminalise fictitious child pornography (Article 9.4), something replicated in the Council Framework Decision (Article 3.2(a), (c)) and Convention on Sexual Exploitation on Children (Article 20.3).

Some have argued that virtual child pornography can be harmful especially since it can be used, for example, in the grooming process (Quayle et al, 2006) but the question is whether this justifies the criminalisation of such images or whether it would be better to criminalise the use of such images. For example, if the concern is that such images can be used in the grooming process then surely the better strategy would be to criminalise grooming rather than the images? Certainly this argument is put forward by some who suggest that in the absence of evidence to the contrary, it is difficult to justify the criminalisation of the possession of images just on the basis that they might be misused (Williams, 2004). The counter-argument is that proving grooming is not necessarily easy (Gillespie, 2008) and that tackling the images rather than their use would, at least, allow them to be removed from circulation, thereby potentially preventing their use in grooming.

The tension that exists over the criminalisation of virtual child pornography can be best summarised as the competing interests of free speech and the protection of the individual (children). A difficulty in discussing this issue is identifying the precise meaning of these concepts. The European model of “free speech” is traditionally considered to be encapsulated within Article 10 of the European Convention on Human Rights, the “freedom of expression”. The other competing form is that contained within the United States Constitution, with the First Amendment being considered to be one of the most important constitutional safeguards.

Within Europe there has been no direct challenge to the criminalisation of virtual child pornography although challenges have been brought against various forms of pornography, including child pornography. The standard way of dealing with such matters was set out by the European Court of Human Rights (ECtHR) in Handyside v United Kingdom (1976) 1 EHRR 737 where it was held that a direct nexus of harm did not need to be shown in order to interfere with the right to freedom of expression, and that interference to uphold the morals of society would suffice where it was proportionate (see, also, Müller v Switzerland (1991) 13 EHRR 212). Others have suggested that criminalising the possession of pornography or providing an objective definition for pornography amounts to an interference with Article 8 of the ECHR but the ECtHR has again been relatively robust in rejecting such challenges (see, for example, Wingrove v United Kingdom (1997) 24 EHRR 1 and O’Carroll v United Kingdom (2005) 41 EHRR SE 1).

However the same is not true within North America. Perhaps the most notable case to rule on this matter was Ashcroft v Free Speech Coalition (2002) 122 S.Ct 1389 but the Supreme Court of Canada has also addressed this matter in R v Sharpe [2001] 1 SCR 45. The US Supreme Court held that legislation that purported to criminalise virtual child pornography was unconstitutional since it infringed the First Amendment. The Court held that virtual child pornography did not lead to any harm being caused to a child and rejected the suggestion that such material could encourage sex offenders to abuse children although, it should be noted, that this was in part because the US Government did not adduce cogent evidence to support such a link.

In terms of the law (for these purposes defined universally) it can be seen that there is no agreement as to whether virtual child pornography should be criminalised. In this paper we seek to set out the competing reasons why virtual child pornography should, or should not, be criminalised.

4 DISCUSSION

In this section we will critically analyse the emancipatory effects of the arguments, presented in the previous section, regarding criminalising or legalising virtual child pornography.
One of the arguments brought forward to support the criminalisation of virtual child pornography is that such images are used in the grooming process with real children (Quayle et al, 2006; Home Office, 2007). It is accepted that part of the “grooming” stage of the cycle of abuse (the cycle of abuse is a model that shows how an offender will seek to befriend a child so as to permit acquiescence to sexual acts: see Gillespie, 2008) may be to show children pornography to help “normalise” sexual activity (Taylor and Quayle, 2003). Virtual child pornography could be used to show images that purport to show children being happy to participate in such activity (in a way similar to the way it is known that morphed photographs do). This means that virtual child pornography could be used to harm and abuse real children. The US Supreme Court was, however, unconvinced by such approaches. It is questionable whether this argument justifies the criminalisation of such images or whether it would be better to criminalise the use of such images. Otherwise the criminalisation tends to be created on the basis that they might be misused (Williams, 2004).

Some authors have suggested that it is becoming increasingly difficult for law enforcement to differentiate between real, morphed and virtual child pornography (Edwards, 1995). If this is correct, then it could be argued that one reason for criminalising such material is to ensure that vital investigative time is not wasted in searching for victims that are not, in fact, real. It could also be difficult for successful convictions if the prosecutors have to prove for each image that it depicts a real and not a virtual child. It is suggested that this can create situations where the prosecuted can operate with impunity (Levy, 2002). A counter-argument is that the same sophisticated technology that can be used to create virtual child pornography, could also be used by law enforcement to distinguish between real and virtual child pornography. Another complex and important issue is how to determine the age of a fictional character. The risk of subjective interpretations of the age of a computer-generated representation is problematic.

Interestingly in Ashcroft v Free Speech Coalition the Supreme Court held that they accepted in principle that this could act as a justification for criminalising such material (2002) 122 S.Ct. 1389 at 1406). The Attorney-General had not, in that case, managed to identify a real case where such confusion had occurred but as technology develops it becomes easier to identify cases where such confusion occurs, potentially raising this as a live issue.

There are many things innocent in themselves...such as cartoons, video games and candy, that might be used for immoral purposes, yet we would not expect those to be prohibited because they can be misused. (Ashcroft v Free Speech Coalition (2002) 122 S.Ct. 1389 at 1402 per Kennedy J.)

Some would argue however that there is a difference between morally neutral 'things' such as cartoons and sweets and those that are morally 'doubtful', and that it is easier to justify an interference with the possession of such matters, this would seem to accord with the European model of the treatment of freedom of expression discussed later. However this does highlight the question as to whether the justification for criminalising such matters is based on a freedom of abuse or an interference to uphold the morality of society.

There have been cases where the police have been unable to prosecute because the suspects were only found in possession of drawings and cartoons: no illegal photographs or pseudo-photographs were discovered. As it is currently legal to possess this material (accept from Ireland), the police then have to return it to the person in possession of it regardless of what it portrays. Making it a criminal offence allows the police to take the images away from the offender and out of public circulation.

Moreover, there is increasing concern among child welfare experts and the law enforcement that the problem will increase as technology becomes more sophisticated. There are already indications that
websites featuring animated images depicting child sexual abuse are on the increase. These are often hosted abroad, beyond our jurisdiction. As technology advances a gap in the law is starting to emerge.

4.2 Problematic empirical assumptions about relationship between VCP and abuse

It is also suggested that the possession and circulating of such images can legitimise and reinforce inappropriate perceptions of children, and while being unregulated, allow the development of a sense of social acceptance towards child sexual abuse. In addition, one could argue that virtual child pornography leads to secondary victimisation where victims of child abuse re-live their abuse when they are exposed to such images. There is however still a lack of empirical evidence showing any direct link between the possession of such images and sexual offending against children (Quayle et al, 2006). This means that we need more research about the connection between looking at computer-generated images and committing sexual abuse to be able to use that as argument for criminalising such images. Levy (2002) uses the following argument to support the legalising of virtual child pornography: "there is every reason to think that if virtual child pornography is legal, pornographers will abandon production of actual images of children in favor of it ... why take the risk of a jail term for producing actual pornography when a simulacrum can be produced legally?" One major problem with this argument is that it is not empirically underpinned. This viewpoint is relevant to our argument because it shows that the uncontroversial aim of protecting children from abuse can easily be turned into an argument for the legalisation of virtual child pornography. The problem with this turn of the argument is that it relies on similar empirical assumptions as the opposing argument and that the empirical studies available do not conclusively support one or the other.

A particular difficulty in this area is deciding what the research and studies shows. The Williams Committee concluded that it was not possible to show any link between pornography and sexual violence beyond that which occurs in the making of the images (Williams, 1979). Whilst law enforcement officers, and those who work closely with them, believe that the link between the possession of pornography and further abuse of children is proven (Long, 2002), others believe that the link is more tenuous (Taylor and Quayle, 2003). Within the United States of America and Canada, this absence of a link was considered an important reason for questioning the legal justification for criminalising the possession of such material (R v Sharp [2001] 1 SCR 45 at [28]).

4.3 Emancipation through freedom of fear and moral convergence

While it is difficult to come to an unambiguous conclusion whether criminalising or legalising of virtual child pornography is more conducive to the freedom of children from abuse, a different argument can be made for the criminalising of virtual child pornography from the viewpoint of a shared moral point of view. It is difficult to deny that the vast majority of members of the public in most countries are morally opposed to virtual child pornography. The root of this opposition may be in a desire to protect children, but its expression is independent of such considerations. To put it differently, most people think that child pornography is repulsive, whether or not its production involved the abuse of real children. Criminalising virtual child pornography on this basis can be seen as an attempt to cast into law the moral perceptions that a populations holds. This is arguably one possible goal of legislation, which supports shared moral norms.

A related argument could then be that outlawing virtual child pornography can lead to freedom of fear of the very existence of such material. Parents may fear that their children view such material or that its existence may increase the probability of their children being abused. Other members of society may just find the thought of the existence repulsive and live a better life just because the possibility of its existence is reduced.

Following this type of argument, it is easier to understand why the UK government, for example, thinks that criminalisation of virtual child pornography may be justified by the "inappropriate feelings
towards children" that it may engender (Home Office 2007). The issue is no longer related to children and their protection but to a majority view of what is considered normal and appropriate.

The problems of this approach are easy to see. Moral preferences change over time and are culturally relative. There are numerous examples of sexual moralities (Foucault, 1998) that have changed drastically over time. Homosexuality can serve as an example of sexual practices that were seen as abjectly repulsive by a majority of the population and therefore outlawed. If moral convergence or consensus is the main reason of criminalising virtual child pornography, then one needs to understand that such preferences are subject to change over time and that regulations may change as a consequence.

4.4 Emancipation as expressed in freedom of speech and expression including in VCP

It has been noted that virtual child pornography could be considered an important aspect of freedom of speech. “[p]aedophilia may be a form of expression involving fantasies and imaginings which may be explicitly important to minority sexual groups, the paedophiles.” (Akdeniz, 1997). Certainly such arguments would chime with judicial expressions of freedom of expression:

"[Article 10(1)] is applicable not only to “information” or “ideas” that are favourably received or regarded as inoffensive or as a matter of indifference, but particularly to those that shock, offend or disturb the State or any sector of the population. There is no point guaranteeing this freedom only as long as it is used in accordance with accepted opinion.”

Otto-Preminger Institute v Austria (1995) 19 EHRR 34

It became clear in the case of Ashcroft v. The Free Speech Coalition that the issue of criminalising virtual child pornography is controversial. The court ruled that it would be a ‘thought crime’ to criminalise such images, since it is digitally created images without a real victim and therefore protected by freedom of expression. Freedom of expression therefore goes beyond the right to publish or hold "nice" images or information and should include material that society may object to. That said, freedom of speech and freedom of expression are both considered to be qualified rights but the key question that arises therefore is whether criminalising virtual child pornography can be said to be necessary and proportionate.

4.5 Ideologies and Hidden Agendas

One of the issues that critical research is keenly aware of is the question of ideology and hidden agendas. Ideologies can be defined as one-sided perceptions of certain social realities that benefit some and disadvantage others (Freeden, 2003; Hawkes, 2003; McLellan, 1995). Ideologies are typically removed from critical discourses via a range of different mechanisms. Central among such mechanisms is the use of language in a way that renders it impossible for participants in discourses to question particular assumptions. In the case of virtual child pornography, one could argue that a central term that betrays ideological assumptions is the use of the term virtual child pornography itself. By using the term "child pornography" in conjunction with "virtual", it becomes very difficult to argue that the artefacts and the way they are produced can under any circumstances be acceptable. The strength of the moral argument against child abuse is such that even the attempt to ignore it is doomed to failure. We are thus witnessing a use of language that impregnates itself against questioning. "Virtual child pornography" carries strong moral connotations and thereby forces discourses in a particular direction. This hardening of discourses is an issue that critical research typically aims to expose. The use of language for ideological purposes is not confined to virtual child pornography. In the area of ICT metaphors such as that of a "virus" for self-replicating programs or of "piracy" for the infringement of intellectual property regulations have been noted before (Stahl, 2007).
Luck (2008) highlights that there are ethical and moral aspects involved when discussing virtual child pornography. He argues that even though most people think 'murder' is wrong, 'virtual murder' in computer games is socially accepted since no 'real' person is actually 'murdered' within a computer game. Luck (2008) then applies this argumentation in the context of 'virtual paedophilia' and question if we should view 'virtual paedophilia' in the same way as with virtual murders since no 'real children are actually being abused in virtual child pornography. His conclusion is that there is a 'possible inconsistency in the social acceptance of virtual immoral acts' when comparing 'virtual murder' in computer games and 'virtual paedophilia'. This logic is questionable since Luck does not differentiate between "killing" and "murder", which are distinct concepts with different meanings. Murder is the unlawful killing of an innocent person. Virtual killing is normally premised on the basis that your player is a 'good' person (police, soldier or some other authority person etc) killing the 'bad' person. While Luck's argument may not hold due to conceptual inaccuracies, the question of consistency of treatment of moral and legal issues in virtual environments remains pertinent.

Notwithstanding the validity of Luck's argument, one can observe the use of language with regards to virtual child pornography that suggests that there are hidden agendas. A prime example of this is the UK consultation document (Home Office, 2007). It is clear that its proponents are in favour of outlawing virtual child pornography. The document is clear in that such an approach would be a departure from existing legislation: "A change to the definitions [of child pornography so as to include virtual material] […] would be a significant shift in the thinking behind the existing legislation which was designed to protect real children from abuse" (p. 6). The report states this in the knowledge that its authors are "… unaware of any specific research into whether there is a link between accessing these fantasy images of child sexual abuse and the commission of offences against children,…" (p.1). The purpose is thus to penalise "inappropriate views about children" (p.1).

Despite this clear indication that the proposed legislation is one that aims to reinforce moral convictions, the authors nevertheless imply that it will contribute directly to the protection of children. In their foreword to the report, the authors say that "Technology has advanced to the point that photographs of real children being abused can be manipulated into cartoons or other depictions. Possession of these images is not caught under existing legislation…” (p. i).

What we are observing here is the use of rhetorical devices to achieve an aim. There is nothing fundamentally wrong with rhetoric but in this case it is used to promote legislative aims that would require an ethical and legal underpinning which is possibly at odds with established views. This is a typical instance of interest to critical research because it exemplifies the use of hidden agendas and ideologies for particular purposes. It would thus be an interesting exercise to investigate the history of the consultation document. Who launched it and what did they have to gain by it? This is a complex question that lies beyond the scope of this conceptual paper, but it is a question worth following up.

5 CONCLUSION

The aim of this paper has been to highlight some of the issues and open questions surrounding virtual child pornography. The topic was chosen because it is related to questions of use, misuse, and regulation of ICT. It is an interesting and highly emotional topic that is not often treated in much depth by information systems scholars, even though computer misuse is often seen as a major problem for the organisational use of ICT.

Using a critical lens we have argued that the topic of virtual child pornography requires detailed understanding of technology as well as its social, ethical, and legal context. The theoretical contribution of the paper is to show that critical ideas can help shed a better light on such issues. The central idea of emancipation may help to better understand some of the arguments surrounding questions such as the legalisation or criminalisation of virtual child pornography.

A critical reflection of research is typically seen as a hallmark of critical research. In this spirit, we need to acknowledge that the paper has not covered all possible angles. One aspect we have omitted
was the economic one. Critical research typically aims to cover the social context of its topic and the economic environment is often seen as central for the explanation of social phenomena. It would certainly be plausible in the area of virtual child pornography given the dominance of adult pornography on the internet and its importance as a piloting technology for most e-commerce interactions.

Another possible improvement would be to provide a more fine-grained ethical analysis of the issues. If virtual child pornography is about moral beliefs, then the question of emancipation should cover the issue of ethical justification of such moral beliefs. The US supreme court arguably followed a utilitarian line of argument when it ruled that free speech was more important than the moral downsides of virtual child pornography. However, such a view would need to be analysed in more depth, which should lead to a more in-depth understanding of the meaning and realisation of virtual child pornography. Furthermore, there are empirical questions about the relationship between consumption of virtual child pornography and abuse of children. A critical piece of research would need to investigate in more depth how such relationships can be established and in what way empirical results can be used as a basis of legislation.

Despite these limitations, we nevertheless believe that this paper makes a substantial contribution to the debate surrounding virtual child pornography while, at the same time, having shown the relevance of critical research when applied to ICT.

References:


**Legal references**

Child Trafficking and Pornography Act 1998

Obscene Publications Acts of 1959 and 1964
A SOCIAL NETWORK ANALYSIS OF THE CO-AUTHORSHIP NETWORK OF THE AUSTRALASIAN CONFERENCE OF INFORMATION SYSTEMS FROM 1990 TO 2006

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0084.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Social Network Analysis, Bibliometrics, Collaboration, Community</td>
</tr>
</tbody>
</table>
A SOCIAL NETWORK ANALYSIS OF THE CO-AUTHORSHIP NETWORK OF THE AUSTRALASIAN CONFERENCE OF INFORMATION SYSTEMS FROM 1990 TO 2006

Cheong, France, RMIT University, School of Business IT, 239 Bourke Street, Melbourne 3001, Australia, france.cheong@rmit.edu.au

Corbitt, Brian, RMIT University, School of Business IT, 239 Bourke Street, Melbourne 3001, Australia, brian.corbitt@rmit.edu.au

Abstract

Using bibliographic data extracted from an Endnote database, social network analysis techniques were used to generate and analyse a network of co-authors with the aim of developing an understanding of the research community that produces the research knowledge published by the Australasian Conference on Information Systems (ACIS). The ACIS community was found to be a healthy small-world community that kept evolving in order to provide an environment that supports collaboration and sharing of ideas between researchers. It was also found that, unlike a similar analysis of the European Conference (ECIS), the Australasian scene was not dominated by a couple of key researchers as quite a significant number of popular researchers were identified.

Keywords: social network analysis, Information Systems discipline, co-authorship, collaboration
1 INTRODUCTION

Interaction between researchers is well known to be the essence of research practice. Researchers interact not only to communicate research activities but also to collaborate with each other to co-produce research and co-author research results (Melin & Persson 1996). Since collaboration has the potential to promote research activity, productivity and impact, it should be encouraged, supported and monitored. Although it has been argued that co-authorship is no more than a partial indicator of collaboration, Laudel (2002) found that a major part of collaboration is not acknowledged as co-authors. Several studies (for instance, Patel 1973) have shown that there is a positive correlation between collaboration and co-authorship. In fact, co-authorship is one of the most tangible and documented forms of research collaboration (Glänzel & Schubert 2004).

A co-authorship network is a social network consisting of a collection of researchers each of whom is connected to one or more other researchers if they have co-authored one or more papers. This is based on the reasonable assumption that researchers who co-author a paper are acquainted with each other, although there are many researchers who know each other quite well but have never written a paper together. Such a network can be represented as a set of nodes (or vertices) denoting co-authors joined by edges (or links) denoting research acquaintance.

Social Network Analysis (SNA) is a sociological approach for analysing patterns of relationships and interactions between social actors in order to discover underlying social structure such as: central nodes that act as hubs, leaders or gatekeepers; highly connected groups; and patterns of interactions between groups (Wasserman & Faust 1994). SNA has been used to study social interaction in a wide range of domains. Examples include: collaboration networks (Newman 2001a), directors of companies (Davis & Greve 1997; Davis, Yoo & Baker 2003), organisational behaviour (Borgatti & Foster 2003), inter-organisational relations (Stuart 1998), computer-mediated communications (Garton, Haythornthwaite & Wellman 1999), and many others.

In this study, we propose to use SNA to study the community of researchers who publish their papers in the Australasian Conference of Information Systems (ACIS) in order to reveal interesting patterns and features within this academic community. With the help of SNA, we hope to develop an understanding of the research community that produces the research knowledge published by ACIS by answering the following: Is the network a random structure or does it display recognisable properties? Is the community highly clustered around a few high profile researchers or is influence spread among a number of researchers? Who are the influential members of this community? What are the weaknesses or strengths of this network?

Clarke (2008) has recently completed a retrospective review of the Information Systems discipline in Australia based on new research and revisiting the work of Culnan (1986; 1987), Land (1992), Barki et al. (1993), Avgerou (1999), Pervan and Cecez-Kecmanovic (2001), Galliers and Whitley (2002) and Banker and Kauffman (2004). He notes that IS emerged as a discipline in the 1960s in Australia, formally from the old Caulfield Institute of Technology (now embedded within Monash University) and has grown now to a community of some 700 with explosive development from the 1970s onwards. Clarke notes further that “the emergence of the IS discipline was in historical terms brisk, but to an observer at the time would have appeared laboured and wayward”. However, by the end of the 1980s all but the two oldest universities (Melbourne and Sydney) had specialist organisational units in IS. What is clear from this research and others reported in a newly edited book by Gable et al. (2008) on the Information Systems discipline in Australia is that the discipline and its associated conference (ACIS) has grown significantly reaching its peak after 2000 and which subsequently has generally shown a decline in research output, student graduates and student demand, albeit at a time of skills shortages in IT in Australia generally. The ACIS conference emerged in 1990 as a response to the growth in research activity in Australian universities.
2 RELATED WORK

The idea of studying research collaboration patterns using bibliographic data is not new as there is a substantial body of literature in Information Science dealing with co-authorship patterns (Crane 1972; Persson & Beckmann 1995; van Raan 1990). Using co-authorship networks to study collaboration patterns between researchers is also not a new idea since with the availability of large bibliographic databases, it is relatively easy to construct large social networks with high reliability. These networks are true social networks, in the sense that it is very likely that two authors who write a paper together are acquainted with each other (Newman & Park 2003).

Scientific collaboration networks were studied for three disciplines, namely: physics, biomedical research and computer science using bibliographic data from four databases for the period 1995-1999 (Newman 2001b). In all three networks, a giant component of researchers was found to exist in which there is only a short path of intermediate acquaintances between any two researchers, hence all networks studied displayed the “small world” property. Some differences found between the disciplines studied were: (1) on average, researchers from experimental disciplines have larger number of collaborators than those from theoretical disciplines (largest average number of collaborators found in high-energy physics), and (2) the degree of network clustering is much lower in biomedicine than in the other disciplines (indicating less social organization in biomedicine). A similar study was performed for the disciplines of mathematics and neuro-science using bibliographic records from an electronic database for the eight-year period from 1991 to 1998 (Barabási et al. 2002).

Research collaboration within the Information Systems discipline has also been studied as social network analysis has been performed for both the International Conference on Information Systems (ICIS) (Xu & Chau 2006) and the European Conference on Information Systems (ECIS) (Vidgen, Henneberg & Naudé 2007). Social network analysis of ICIS was conducted using bibliographic data for the period 1980 to 2005 available from the Association of Information Systems to study the social identity of the discipline (Xu & Chau 2006). Among other things, results showed that: (1) the community of international IS researchers is well connected and they frequently interact with each other, (2) there exists a giant component of well-connected and most productive authors, and (3) the network has evolved healthily over time with the addition of new members and the improved connection among members.

The ECIS analysis was performed using bibliographic data from an Endnote database available from the London School of Economics for the period 1993 to 2005 (Vidgen, Henneberg & Naudé 2007). Research contributions were separated into research papers and panels and two networks were generated and analysed. While the panel network displayed small world properties, unlike other collaboration networks, the co-authorship network displayed only a few “small world” properties and hence a lesser sense of social cohesion than would be expected. Although social network analysis of the Information System discipline has been performed at the international and European levels, to the best of our knowledge, it has not been attempted for the Australasian scene yet, other than a study of frequency of publication locations by Australian authors in IS by Sellitto (2007), hence the motivation for the present work.

3 METHODOLOGY

Social network analysis (SNA) has emerged as a key technique in the social and behavioural sciences as well as in other major disciplines (Wasserman & Faust 1994). The main focus of SNA is on the relationships among social entities (e.g. communications among members of a group) and it makes use of a variety of statistical and visual analyses to achieve this. Although, social networks were initially studied in the social sciences, such studies were restricted to rather small systems viewing these networks as static graphs consisting of nodes representing individuals and links representing various quantifiable social interactions. In contrast, recent approaches rooted in statistical physics focus more
on large networks searching for universalities both in the topology of the network and in the dynamics
governing its evolution (Barabási et al. 2002).

Recently, SNA has been increasingly used as a structured way to analyse the extent of informal
relationships (among people, teams, departments, or even organisations) within various formally
defined groups (Cross et al. 2001). SNA makes visible these otherwise invisible patterns of
interaction, to identify important groups in order to facilitate effective collaboration (Cross, Borgatti &
Parker 2003). Thus, SNA helps to identify and assess the health of strategically important networks in
an organisation. In the context of this study, we are using SNA to gain an understanding of the nodes
(co-authors) and relationships (those who wrote a paper together) in the co-authorship network.
Clearly, there are many different metrics that can be used to assess such networks. At an aggregate
level, we will analyse the network as a whole in order to identify important major groups or
components within the community of researchers, and for the giant or core component we will use
measures that can give an indication of the productivity of the network (i.e. density of the network),
speed of communication within the network (diameter of the network), etc. At a lower level, we will
analyse the nodes of the network using several measures of centrality to find out who the most popular
and influential researchers are within the ACIS community.

4 DATA COLLECTION AND PROCESSING

The bibliometric data used in this study is based on bibliographic data extracted from an Endnote
database available from the ISWorld Net Research and Scholarship page\(^1\) of the Association for
Information Systems. The Endnote database contained all research papers published by ACIS from
1990 to 2006.

The contents of the Endnote database were exported in XML format in order to facilitate further
processing. Since we are only interested in co-authored publications, all papers written by single
authors were ignored. This was achieved by writing a small Java program to extract a list of co-authors
on a paper-by-paper basis. Another custom-written Java program was then used to convert this list of
authors into a network file to the DL format which is readable by UCINet (Borgatti, Everett &
Freeman 2002), the software used for most of the social network analysis in this study. Apart from
generating the DL file, the Java utility was programmed to output a list of authors sorted in
alphabetical order which was visually inspected to discover typographical errors (e.g. Peta Dark
instead of Peta Darke) and inconsistencies in authors’ names, especially those with middle initials who
used them part of the time (e.g. Brian Corbitt instead of Brian J. Corbitt) and those with aliases (e.g.
Kit Dampney instead of C. N. G. (Kit) Dampney). A more subtle typographical error was the use of a
left apostrophe instead of a right apostrophe (e.g. Paul O’Brien instead of Paul O’Brien) in the Endnote
database. Data cleaning was a highly iterative activity and consumed a large part of the data
processing activities. Once the co-authorship data was in UCINet’s DL format, various statistical
analyses were performed using UCINet at network and co-author levels and the results of these
analyses are reported and discussed in the following sections. Visualisation of the co-authorship
network (or parts of the network) was performed using Pajek (Batagelj & Mrvar 1998), another
popular SNA software.

5 NETWORK ANALYSIS

Recent analysis by Gable (2008) shows that at the ACIS conference ten universities (Monash,
Melbourne, Edith Cowan, Curtin, Deakin, QUT, Wollongong, UNSW, Tasmania and Victoria) have
contributed the vast majority of papers. This paper reports on the broader perspective of the people
involved as a number of key people have moved from one university to another as universities try to
shore up quality and buy expertise to strengthen or develop existing or new departments. Table 1

\(^1\) http://home.aisnet.org/displaycommon.cfm?an=1&subarticlenbr=395

Proceedings ECIS 2009
shows the evolution of the ACIS community during the period 1990 to 2006. The cumulative number of papers presented at the conference grew from 15 in 1990 to 1333 in 2006 while the cumulative number of co-authored papers grew from nine to 820 during that time frame. As of 2006, the percentage of co-authored papers represents 69% of the total number of papers.

The size of a social network is denoted by the number of actors or nodes (co-authors in this case) and it gives an indication of the likelihood of interaction between nodes; the bigger the network, the greater the likelihood of interaction between co-authors. However, the bigger the network, the more difficult it becomes for everyone to be connected with each other and when the network is not fully connected, it contains a number of sub-networks (called components) for which there are no paths between nodes from one sub-network to another sub-network. The number of co-authors in the ACIS network grew from 19 in 1990 to 1256 in 2006 while the number of co-authors in the main component (the largest sub-network in which there is a path from a co-author to any other co-author) grew from three in 1990 to 587 in 2006. As of 2006, the percentage of co-authors in the main component represents 46% of the total number of co-authors. It should be noted that the main component is not a fully-connected network (i.e., everyone is not connected to everyone). The degree of connectedness of a network (or sub-network) is given by the density measure, which is the percentage of the number of actual connections over the total number of possible connections. The density of the main component dropped from 33.33% in 1990 to 0.24% in 2006.

Another interesting feature of a network is an indication of the amount of time for a communication to pass through the network. A commonly-used measure is the diameter of the network; the shorter the diameter, the faster the diffusion of communications. The diameter of a network is measured by the longest geodesic distance in the network with the geodesic distance being the shortest path (in terms of number of links or connections) between any two nodes. So far, the diameter of the network has grown to 9, slightly more than what would be expected from a “small-world” network. One of the main characteristics of a small-world is the so-called “Six Degrees of Separation” phenomenon in which it is claimed that everybody on the planet is separated by only six other people (Milgram 1967; Watts 1999).

The structure of the ACIS network of co-authors displays small world properties because co-authors are well-connected, and are close to each other. Hence, information and knowledge can be transferred effectively in the network, although in practice the flow of information might be different from the established formal network structure, and is thus acknowledged as a limitation of the study (Cross, Borgatti & Parker 2002).

Table 1: Evolution of ACIS community (1990-2006)

<table>
<thead>
<tr>
<th>Year</th>
<th># Papers (cumulative)</th>
<th># Co-authored papers (cumulative)</th>
<th># Actors in co-authorship network</th>
<th># Actors in main component</th>
<th>Density of main component</th>
<th>Diameter of main component</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>15</td>
<td>9</td>
<td>19</td>
<td>3</td>
<td>0.3333</td>
<td>1</td>
</tr>
<tr>
<td>1991</td>
<td>44</td>
<td>24</td>
<td>48</td>
<td>4</td>
<td>0.2500</td>
<td>1</td>
</tr>
<tr>
<td>1992</td>
<td>89</td>
<td>40</td>
<td>84</td>
<td>4</td>
<td>0.2500</td>
<td>2</td>
</tr>
<tr>
<td>1993</td>
<td>149</td>
<td>69</td>
<td>142</td>
<td>5</td>
<td>0.2000</td>
<td>1</td>
</tr>
<tr>
<td>1994</td>
<td>205</td>
<td>104</td>
<td>209</td>
<td>8</td>
<td>0.1429</td>
<td>2</td>
</tr>
<tr>
<td>1995</td>
<td>268</td>
<td>144</td>
<td>264</td>
<td>13</td>
<td>0.0962</td>
<td>2</td>
</tr>
<tr>
<td>1996</td>
<td>340</td>
<td>179</td>
<td>324</td>
<td>13</td>
<td>0.0962</td>
<td>2</td>
</tr>
<tr>
<td>1997</td>
<td>401</td>
<td>216</td>
<td>372</td>
<td>25</td>
<td>0.0617</td>
<td>3</td>
</tr>
<tr>
<td>1998</td>
<td>461</td>
<td>259</td>
<td>420</td>
<td>29</td>
<td>0.0505</td>
<td>2</td>
</tr>
<tr>
<td>1999</td>
<td>563</td>
<td>330</td>
<td>516</td>
<td>51</td>
<td>0.0263</td>
<td>4</td>
</tr>
<tr>
<td>2000</td>
<td>657</td>
<td>405</td>
<td>628</td>
<td>110</td>
<td>0.0123</td>
<td>5</td>
</tr>
<tr>
<td>2001</td>
<td>741</td>
<td>462</td>
<td>694</td>
<td>170</td>
<td>0.0077</td>
<td>5</td>
</tr>
<tr>
<td>2002</td>
<td>845</td>
<td>543</td>
<td>795</td>
<td>170</td>
<td>0.0077</td>
<td>5</td>
</tr>
<tr>
<td>2003</td>
<td>991</td>
<td>660</td>
<td>940</td>
<td>273</td>
<td>0.0046</td>
<td>7</td>
</tr>
<tr>
<td>2004</td>
<td>1111</td>
<td>758</td>
<td>1058</td>
<td>410</td>
<td>0.0033</td>
<td>8</td>
</tr>
<tr>
<td>2005</td>
<td>1223</td>
<td>837</td>
<td>1146</td>
<td>467</td>
<td>0.0030</td>
<td>9</td>
</tr>
<tr>
<td>2006</td>
<td>1333</td>
<td>920</td>
<td>1256</td>
<td>587</td>
<td>0.0024</td>
<td>9</td>
</tr>
</tbody>
</table>
In order to give an idea of the evolution of the ACIS co-authorship in time, cumulative networks were drawn from 1990 to 2005 and assembled in Figure 1. Since the main component was rather small in the early years (1990-1994), the first four pictures show the complete network (with the main component in the middle of the network) while the rest of the pictures (1995-2005) show only the main component. It can be seen that the main component started to become significant around the year 2000.

![Figure 1: ACIS Network evolution (1990-2005)](image)

The state of the ACIS network in 2006 (most current at the time of writing) represented using its main component is shown in Figure 2. We chose to represent the co-authorship network as a directed network i.e. directed links from the main author to his/her co-authors. The thickness of the links gives an indication of the number of co-authored papers between a main author and the particular co-author. Although the network shows the high popularity of certain individuals (e.g. Graeme Shanks), it also shows that the scene is not dominated by a few individuals as there is quite a range of well-connected individuals. A more detailed analysis of popular individuals follows in the next section.

6 EGO ANALYSIS

After having analysed the characteristics of the ACIS network as an entity, we now analyse it in terms of the individual actors or “egos” that make up the nodes of the network. More specifically, co-authors are analysed in terms of their centrality in the ACIS network. The idea of centrality of individuals was one of the earliest used by social network analysts and the origins of this idea can be found in the sociometric concept of the star i.e. the most popular person or the person at the centre of attention (Scott 2007). Thus, a central actor is one at the centre of a number of connections i.e. an actor with a large number of direct links with other actors.
Centrality is measured by the degree of the various nodes in the network, with degree representing the number of other nodes to which a node is adjacent. This measure of centrality is known as local centrality since indirect connections to the particular node are ignored. Thus, the notion of centrality has been extended to global centrality (Freeman 1979) to include the distant connections of the nodes. This is measured by the closeness of the nodes to other nodes expressed in terms of the distances among the various nodes. Betweenness (Freeman 1979) is another centrality measure which measures the extent to which a particular node lies between the various other nodes of the network. A node of relatively low degree may play an important intermediary role (e.g. broker, gatekeeper, etc) and hence be a central node in the network. Eigenvector (Bonacich 1972) is another measure of centrality proposed based on the belief that the centrality of a particular node cannot be assessed in isolation from the centrality of all the other nodes to which it is connected. Centrality scores are assigned to nodes based on the principle that connections to high-score nodes contribute more to the score of the particular node than connections to low-score nodes.

The term structural hole was coined by Burt (1992) to refer to some important aspects of positional advantage (or disadvantage) of actors in a network. He developed a number of measures to explain how and why the ways actors are connected affect their constraints and opportunities and hence their behaviour. Table 2 shows the top 30 actors ranked on the centrality measures previously discussed, namely: (1) degree, (2) betweenness, (3) closeness, (4) eigenvector, (5) and structural holes.

Table 2: Centrality measures of actors in main component

<table>
<thead>
<tr>
<th>Rank</th>
<th>Name</th>
<th>Degree</th>
<th>Betweenness</th>
<th>Closeness</th>
<th>Eigenvector</th>
<th>Structural Hole</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rosemann M</td>
<td>15</td>
<td>9</td>
<td>17</td>
<td>6</td>
<td>McKay J</td>
</tr>
<tr>
<td>2</td>
<td>Shanks G</td>
<td>14</td>
<td>10</td>
<td>9</td>
<td>18</td>
<td>Campbell J</td>
</tr>
<tr>
<td>3</td>
<td>Carroll JM</td>
<td>13</td>
<td>11</td>
<td>8</td>
<td>19</td>
<td>Cater-Steel A</td>
</tr>
<tr>
<td>4</td>
<td>Love PED</td>
<td>9</td>
<td>17</td>
<td>Arnott DR</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Marshall P</td>
<td>9</td>
<td>18</td>
<td>Campbell J</td>
<td>6</td>
<td>Rahim MM</td>
</tr>
<tr>
<td>6</td>
<td>Johnstone MN</td>
<td>8</td>
<td>19</td>
<td>Cater-Steel A</td>
<td>6</td>
<td>Rouse AC</td>
</tr>
</tbody>
</table>

The term structural hole was coined by Burt (1992) to refer to some important aspects of positional advantage (or disadvantage) of actors in a network. He developed a number of measures to explain how and why the ways actors are connected affect their constraints and opportunities and hence their behaviour. Table 2 shows the top 30 actors ranked on the centrality measures previously discussed, namely: (1) degree, (2) betweenness, (3) closeness, (4) eigenvector, (5) and structural holes.

Table 2: Centrality measures of actors in main component

<table>
<thead>
<tr>
<th>Rank</th>
<th>Name</th>
<th>Degree</th>
<th>Betweenness</th>
<th>Closeness</th>
<th>Eigenvector</th>
<th>Structural Hole</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rosemann M</td>
<td>15</td>
<td>9</td>
<td>17</td>
<td>6</td>
<td>McKay J</td>
</tr>
<tr>
<td>2</td>
<td>Shanks G</td>
<td>14</td>
<td>10</td>
<td>9</td>
<td>18</td>
<td>Campbell J</td>
</tr>
<tr>
<td>3</td>
<td>Carroll JM</td>
<td>13</td>
<td>11</td>
<td>8</td>
<td>19</td>
<td>Cater-Steel A</td>
</tr>
<tr>
<td>4</td>
<td>Love PED</td>
<td>9</td>
<td>17</td>
<td>Arnott DR</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Marshall P</td>
<td>9</td>
<td>18</td>
<td>Campbell J</td>
<td>6</td>
<td>Rahim MM</td>
</tr>
<tr>
<td>6</td>
<td>Johnstone MN</td>
<td>8</td>
<td>19</td>
<td>Cater-Steel A</td>
<td>6</td>
<td>Rouse AC</td>
</tr>
</tbody>
</table>
Since we chose to represent our co-authorship network as a directed network (because the author selected the co-author for writing the paper), a centrality degree analysis yielded two scores: out degree (number of connections sent out i.e. as main author) and in degree (number of connections received i.e. as co-author). The first part of Table 2 shows the ranking of the top 30 individuals on the out degree score while the second part of the table ranks individuals by the in degree score. The top

<table>
<thead>
<tr>
<th>Rank</th>
<th>Out Degree</th>
<th>In Degree</th>
<th>Name</th>
<th>Score</th>
<th>Rank</th>
<th>Out Degree</th>
<th>In Degree</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Shanks G</td>
<td>50</td>
<td>Bentley J</td>
<td>12</td>
<td>1</td>
<td>Gregor S</td>
<td>17</td>
<td>Marshall P</td>
</tr>
<tr>
<td>2</td>
<td>Carroll JM</td>
<td>40</td>
<td>Howard S</td>
<td>18</td>
<td>12</td>
<td>Lane MS</td>
<td>8</td>
<td>Marshall P</td>
</tr>
<tr>
<td>3</td>
<td>Arnott DR</td>
<td>34</td>
<td>Campbell J</td>
<td>19</td>
<td>10</td>
<td>Kuzic J</td>
<td>8</td>
<td>Timbrell GT</td>
</tr>
<tr>
<td>4</td>
<td>Pervan GP</td>
<td>29</td>
<td>Gable GG</td>
<td>20</td>
<td>19</td>
<td>Prananto A</td>
<td>8</td>
<td>Steele PM</td>
</tr>
<tr>
<td>5</td>
<td>Fisher J</td>
<td>20</td>
<td>Lin C</td>
<td>21</td>
<td>17</td>
<td>Seddon PB</td>
<td>7</td>
<td>Standing C</td>
</tr>
<tr>
<td>6</td>
<td>Rosemann M</td>
<td>19</td>
<td>Bursture F</td>
<td>14</td>
<td>20</td>
<td>Darke P</td>
<td>8</td>
<td>Standing C</td>
</tr>
<tr>
<td>7</td>
<td>Dawson L</td>
<td>16</td>
<td>Swatman PMC</td>
<td>23</td>
<td>19</td>
<td>Keller S</td>
<td>7</td>
<td>Hasen H</td>
</tr>
<tr>
<td>8</td>
<td>Smith R</td>
<td>16</td>
<td>Howard S</td>
<td>24</td>
<td>11</td>
<td>Yoong P</td>
<td>7</td>
<td>Schepers H</td>
</tr>
<tr>
<td>9</td>
<td>Shanks G</td>
<td>22</td>
<td>Calvert C</td>
<td>19</td>
<td>17</td>
<td>Nguyen L</td>
<td>18</td>
<td>Chaiyasut P</td>
</tr>
<tr>
<td>10</td>
<td>Arnott DR</td>
<td>21</td>
<td>Corbitt BJ</td>
<td>19</td>
<td>19</td>
<td>Simson G</td>
<td>18</td>
<td>Cheong K</td>
</tr>
<tr>
<td>11</td>
<td>Carroll JM</td>
<td>32</td>
<td>Weber R</td>
<td>19</td>
<td>19</td>
<td>Howard S</td>
<td>18</td>
<td>Moore J</td>
</tr>
<tr>
<td>12</td>
<td>Fisher J</td>
<td>20</td>
<td>Tansley E</td>
<td>20</td>
<td>19</td>
<td>Rahim MM</td>
<td>18</td>
<td>Nuredini J</td>
</tr>
<tr>
<td>13</td>
<td>Dawson L</td>
<td>16</td>
<td>Staehr LJ</td>
<td>21</td>
<td>17</td>
<td>Darke P</td>
<td>18</td>
<td>Rembch M</td>
</tr>
<tr>
<td>14</td>
<td>O'Donnell PA</td>
<td>20</td>
<td>Gibbs MR</td>
<td>22</td>
<td>18</td>
<td>Schackleton P</td>
<td>18</td>
<td>Tobin D</td>
</tr>
<tr>
<td>15</td>
<td>Rosemann M</td>
<td>19</td>
<td>Jayaganesh M</td>
<td>23</td>
<td>18</td>
<td>Giannoccoro A</td>
<td>18</td>
<td>Cheong K</td>
</tr>
<tr>
<td>16</td>
<td>Pervan GP</td>
<td>19</td>
<td>Moody DL</td>
<td>24</td>
<td>18</td>
<td>Hodgson B</td>
<td>18</td>
<td>Chaiyasut P</td>
</tr>
<tr>
<td>17</td>
<td>Shanks G</td>
<td>75</td>
<td>Darke P</td>
<td>20</td>
<td>17</td>
<td>Staehr LJ</td>
<td>15</td>
<td>Cheong K</td>
</tr>
<tr>
<td>18</td>
<td>Rosemann M</td>
<td>48</td>
<td>Indulska M</td>
<td>17</td>
<td>18</td>
<td>Gibbs MR</td>
<td>15</td>
<td>Moore J</td>
</tr>
<tr>
<td>19</td>
<td>Carroll JM</td>
<td>34</td>
<td>Simson G</td>
<td>17</td>
<td>19</td>
<td>Recker J</td>
<td>14</td>
<td>Nuredini J</td>
</tr>
<tr>
<td>20</td>
<td>Arnott DR</td>
<td>30</td>
<td>Giannoccoro A</td>
<td>20</td>
<td>17</td>
<td>Tansley E</td>
<td>14</td>
<td>Darke P</td>
</tr>
<tr>
<td>22</td>
<td>O'Donnell PA</td>
<td>21</td>
<td>Fisher J</td>
<td>22</td>
<td>16</td>
<td>Rahim MM</td>
<td>14</td>
<td>Darke P</td>
</tr>
<tr>
<td>23</td>
<td>Corbitt BJ</td>
<td>21</td>
<td>Green P</td>
<td>23</td>
<td>16</td>
<td>Swatman PMC</td>
<td>14</td>
<td>Darke P</td>
</tr>
<tr>
<td>24</td>
<td>Moody DL</td>
<td>20</td>
<td>Nguyen L</td>
<td>24</td>
<td>15</td>
<td>Chaiyasut P</td>
<td>13</td>
<td>Darke P</td>
</tr>
<tr>
<td>25</td>
<td>Shanks G</td>
<td>22</td>
<td>Arnott DR</td>
<td>12</td>
<td>17</td>
<td>Love PED</td>
<td>8</td>
<td>Quaddus MA</td>
</tr>
<tr>
<td>26</td>
<td>Rosemann M</td>
<td>18</td>
<td>Seddon PB</td>
<td>11</td>
<td>18</td>
<td>Gregor S</td>
<td>8</td>
<td>Shackleton P</td>
</tr>
<tr>
<td>27</td>
<td>Carroll JM</td>
<td>16</td>
<td>Gable GG</td>
<td>11</td>
<td>19</td>
<td>Smith R</td>
<td>8</td>
<td>Lane MS</td>
</tr>
<tr>
<td>28</td>
<td>Bursture F</td>
<td>14</td>
<td>Corbitt BJ</td>
<td>11</td>
<td>20</td>
<td>Zaslavsky AB</td>
<td>8</td>
<td>Dr. Lichtenstein S</td>
</tr>
<tr>
<td>29</td>
<td>Fisher J</td>
<td>13</td>
<td>Swatman PMC</td>
<td>10</td>
<td>21</td>
<td>Moody DL</td>
<td>8</td>
<td>Dr. Lichtenstein S</td>
</tr>
<tr>
<td>30</td>
<td>Pervan GP</td>
<td>12</td>
<td>Campbell J</td>
<td>9</td>
<td>22</td>
<td>Cater-Steel A</td>
<td>8</td>
<td>Timbrell GT</td>
</tr>
</tbody>
</table>

Since we chose to represent our co-authorship network as a directed network (because the author selected the co-author for writing the paper), a centrality degree analysis yielded two scores: out degree (number of connections sent out i.e. as main author) and in degree (number of connections received i.e. as co-author). The first part of Table 2 shows the ranking of the top 30 individuals on the out degree score while the second part of the table ranks individuals by the in degree score. The top
scorers in terms of out degree (main author) are: Michael Rosemann closely followed by Graeme Shanks, Jennie Carroll, Dubravka Cecez-Kecmanovic and Julie Fisher. The individuals with high out degree scores can be thought of as having high influence in the network while those with high in degree scores as prestigious or popular individuals. The most prestigious individual is Graeme Shanks followed by Michael Rosemann.

In regards to betweenness centrality, the top individuals are: Graeme Shanks followed by Jennie Carroll. Thus, Shanks and Carroll can be viewed as leaders in the ACIS network since being on the shortest paths between other individuals they are able to control the flow of information in the network. The leading individuals in terms of closeness centrality are: Graeme Shanks closely followed by David Arnott, Jennie Carroll, Julie Fisher, Linda Dawson and Peter O'Donnell. Since closeness centrality measures the distance of an individual to all others in the network, the closer an individual is to others, the more favoured that individual is. Individuals with high closeness scores are likely to receive information more quickly than others as there are fewer intermediaries between them. Graeme Shanks is by far the leading individual when the eigenvector centrality criterion is used. This means that he is connected to many other individuals who are well connected and thus is most likely to receive new ideas.

Structural holes was measured in terms of Effective size of the network i.e. the number of connections an individual has, minus the average number of connections that each individual has to other individuals. Graeme Shanks followed by Michael Rosemann again led on this criterion suggesting that they have more opportunities to act as brokers or coordinators. From the ego analysis, it can be seen that, unlike the ECIS community, influence in ACIS is not limited to a few individuals (Vidgen, Henneberg & Naudé 2007).

7 VISUAL ANALYSIS

The top-ranking 30 actors for each centrality criteria mentioned in Table 2 were merged and an ego network (sub-network) made up of only these actors and their collaborators extracted from the main component sub-network in an attempt to visually identify any leading individuals. The resulting network is shown in Figure 3. From Figure 3, it can be seen that Graeme Shanks and Michael Rosemann are significant individuals. It can also be seen that there are other popular individuals, such as (limited to a few names in alphabetical order as the list can be quite long): Jennie Carroll, Dubravka Cecez-Kecmanovic, Brian Corbitt, Julie Fisher, Guy Gable, Michael Lane, Graham Pervan, Peter Seddon, Craig Standing, Paul Swatman, and Paula Swatman. Figure 3 further reinforces the previous finding (from the ego analysis) that there are quite a number of key researchers in the ACIS community. Another possible limitation of the study is that the cumulative approach used to include authors in the network does not remove authors who for some reason or another are no longer present. However, in the current study this was not the case as the main actors identified are known to be still active.

8 DISCUSSION

The key findings of this study of the community of ACIS researchers are: (1) the total number of papers presented at the conference has been constantly growing since the establishment of the conference, (2) currently the percentage of co-authored papers represents 69% of the total number of papers, (3) the network contains a significantly large main component which includes 47% (587 individuals) of the total number of co-authors, (4) the main component exhibits small-world characteristics (nodes that are well-connected and close to each other), (5) although Graeme Shanks and Michael Rosemann seem to be very popular individuals, they are closely followed by a number of other popular individuals.

The positive evolution of the main component of the ACIS network coupled with the presence of a number of key individuals (rather than a few) are evidence of the healthy status of the ACIS
community. They are proof of the ability of the community to attract new members over the years and to produce new generations of popular researchers. It is worth noting that although popular researchers play an important role in the ACIS network, other researchers are also important as without them is no ACIS community.

Figure 3: Main actors sub-network

The existence of a number of key researchers in the ACIS community provides several advantages. Firstly, it ensures the diversity of research within the community as a large number of popular researchers would decrease the likelihood of performing similar research. Secondly, the ACIS community is resilient since removing a few key persons from the community (e.g. retirement, etc) will not cause it to fall apart as other key persons will ensure its continued existence. Thirdly, succession planning is a smooth and effective process as the large number of current generation popular researchers train their doctoral students and junior collaborators to form a large pool of researchers from which new popular researchers will emerge to lead the community in the future.

The structural properties of the ACIS network also indicates the existence of some potential problems with the community. Since the diameter of the network is slightly wider than desirable (nine instead of six), it is possible that information might not travel quickly enough for effective collaboration. However, the diameter of the ACIS network is still better than the diameter of 31 for the ECIS network (Vidgen, Henneberg & Naudé 2007). Although the presence of many key researchers guarantees free and open debate which is the lifeblood of academia, the uptake of new ideas might meet with more opposition (e.g. political resistance, resistance to change, etc) because there are more people to convince and hence greater likelihood of disagreements.
9 CONCLUSION

In this paper, we used SNA to study the interactions between co-authors of research papers presented at ACIS. SNA provides techniques to analyse the structure of a network as an entity as well as with techniques to analyse individual nodes (egos) and their place in the network. Using SNA metrics and visualisation techniques we were able to reveal structural characteristics of the ACIS co-authors community and identify influential members of this community. The ACIS community was found to be a healthy small-world community that kept evolving in order to provide an environment that supports collaboration and sharing of ideas between researchers. It was also found that unlike Europe, the Australasian scene was not dominated by a couple of key researchers as quite a number of such people were identified.

Future work that could be undertaken to provide a better understanding of the ACIS community includes: (1) identification of the various groups that exist in the network and their research topics (using keyword analysis), and (2) incorporating institutional information in the analysis. Since most researchers publish in more than one conference or journal, the analysis of bibliographical data from ACIS cannot give a complete picture of the Australasian IS authorship patterns. Thus, for a more complete coverage of the IS discipline in Australasian, the boundary of the network should be extended to include other IS-related conferences and journals.

References


USING INTERPRETIVE STRUCTURAL MODELING TO UNCOVER SHARED MENTAL MODELS IN IS RESEARCH

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0710.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Research methodology, Means-end analysis, Matrix analysis, Theory Building</td>
</tr>
</tbody>
</table>
USING INTERPRETIVE STRUCTURAL MODELING TO UNCOVER SHARED MENTAL MODELS IN IS RESEARCH

Abstract

The role of grounded approaches has been advocated for long in IS research. However, the inherent subjectivity of such approaches and the apparent lack of a basis to validate or even replicate such research has often been the subject of debate among IS researchers. As a result, many IS researchers tend to fall back on variance-theoretic approaches to conceptualize, design and operationalize their research. In this paper, we show how a grounded approach, interpretive structural modeling (ISM), can be used to qualitatively elicit individual cognitive structures. Further, we show how it can be applied to derive the shared aspects of such a structure across many individuals. We use the well-known technology acceptance model (TAM) to demonstrate the utility of our approach. We conclude the paper by discussing the strengths and weaknesses of this approach.

Keywords: Interpretive structural modeling, information system, user acceptance, inductive research, research methods

1. INTRODUCTION

ISM is a graph-theoretic method that belongs to the causal mapping family of approaches. It can be used to address problems that are complex and subjective. The ISM approach is useful when a multilevel research design is required where the outcome of the research cannot be predicted based on available research (Klein and Kozlowski, 2000) – implying the use of both theory- and data-driven approaches for research.

Our objective in this paper is to demonstrate the effectiveness of interpretive structural modeling (ISM) in carrying out inductive research in information systems (IS). We have focused on inductive research (data-driven) because of the relative paucity of such research in the IS literature. While we believe that inductive approaches have the potential to contribute significantly to both theory and practice, we are also aware that theories that are grounded in data tend to be relatively harder to defend. The essential premise for the use of such approaches by a researcher tends to be that existing theories do not account for the complex phenomena that a researcher faces. It is implicit in inductive approaches that a careful study of the phenomenon will reveal the hidden patterns that a researcher believes exist.

Using the ISM approach, we suggest a way to efficiently elicit and synthesize user responses with respect to complex phenomena. We demonstrate that ISM can help to operationalize research approaches that can be considered a grounded theory by using graphical techniques to extract underlying structures from data. These structures would form what is often the outcome of grounded theory – revealed thought patterns. We believe that individuals share aspects of cognitive structures that they form about technology and its use. Such shared aspects across individuals can be used to develop theoretical models which can be tested and validated subsequently. We refer to the shared aspects of cognitive structure across individuals as shared mental models in this paper.

Proceedings ECIS 2009
We apply non-directed and non-model-driven approach of ISM to generate a well-accepted and well-known theoretical model, technology acceptance model (TAM). We would like to clarify that the focus of this paper is not to replicate or revalidate TAM model but to just use it as an example to demonstrate the application of ISM approach to generate theoretical models in IS field. By showing that such theory-generation is possible, we hope to persuade IS researchers to employ this technique appropriately in settings or phenomena for which theories do not yet exist. While other causal mapping techniques have been used in IS research (Nelson et al., 2000; Tan and Hunter, 2002), the ISM approach is different in that it is relatively more efficient (in some cases) and lends itself to being replicated more effectively.

Since we are not sure about the level of familiarity with ISM methodology among IS researchers, we introduce ISM methodology first. We then describe how the characteristics of problems (especially complexity and subjectivity) that IS researchers face make the problems well suited for being scrutinized with ISM. We then take up a well-tested theory in IS research (the technology acceptance model - TAM) and show how ISM can be used effectively to develop a TAM. After analyzing the results, we discuss the implications of ISM for IS research in general and elaborate on its strengths and weaknesses. Finally, we suggest areas of IS research that could benefit from the use of ISM.

The contribution of this paper lies not so much in the novelty of the particular finding – but in the novelty and potential of data collection and analysis and how ISM approach can be used in IS research. Since we were able to recreate the TAM structure, we are encouraged to suggest ISM as a viable research approach for many IS research problems that are inherently inductive and qualitative in nature.

2. EXPLANATORY FRAMEWORKS IN IS RESEARCH

We start with the premise that no IS research is either completely inductive or completely deductive. To that extent, development and/or extension of the theoretical framework is often the precursor to empirical support for that theory. However, a major problem with existing approaches to theory development (and research, in general) in IS is the fragmented adhocracy, a result of the federated research framework at work (Landry and Banville, 1992; Hirschheim et al., 1996). Given the richness of the field and the absence of normative or prescriptive frameworks, researchers and consumers of research (who are primarily other researchers) tend to align themselves with a well-established set of ideas or work on a well-known problem. Researchers usually adopt approaches that tend to reconfirm existing theories in a different context or marginally extend them. In doing so, researchers protect themselves from criticisms from other groups that do not agree with their assumptions or beliefs. This framework certainly allows for a thousand flowers to bloom – and enriches the IS field. However, in being overly theory driven, IS researchers may end up playing to the wrong gallery – that of other researchers. However, if IS is an applied discipline, then practitioner-driven research can also be effectively and rigorously incorporated into the IS research process. Stated differently, research that is grounded in data, and which need not be subjected to the researcher’s interpretation, can also be useful to investigate multiple phenomena.

The notion of causality in IS research has long been held to the same standards as those of its stronger and better-established disciplines like psychology and economics. The plurality of perspectives in IS research has certainly led to stronger criticism and a shared awareness/need for rigor. This plurality has also resulted in a variety of “explanation types” in IS research. Since the field of IS is built from both natural and artificial scientific disciplines, Hovorka et al. (2003) argue that explanation types depend on the reference disciples through which research phenomena are understood and research agendas are shaped. Hovorka et al. (2003) provide the following types of explanation types: descriptive/structural explanation, covering-law explanation, statistical relevance explanation, pragmatic explanation and functional explanation. While majority of IS research was categorized as statistical relevance (35%), a significant proportion of explanation in IS research was categorized as descriptive/structural explanation (25%) and framework or model-based (23%). The emergence of process-theory (Soh and Markus, 1995; Crowston,
2000; Kanungo, 2003) perspective in IS research points to the need for alternate methodologies to support descriptive/structural explanation option.

Our approach, in this paper, can be considered to belong to the descriptive and pragmatic explanation categories. We take advantage of the fact that human knowledge, "consists of models constructed by human beings" (Warfield, 1998). Our approach focuses on modeling complex entities created by the multiple interactions of components by abstracting from certain details of structure and components, and concentrating on the dynamics (or linkages) that define the behaviors, properties, and relationships that are internal or external to the system.

3. INTERPRETIVE STRUCTURAL MODELING (ISM)

ISM falls into the soft operations research (OR) family of approaches. Soft OR methods can be used to augment traditional quantitative methods, but do not replace traditional tools and techniques (Glasgow, 2000). ISM is a process that helps groups of people in structuring their collective knowledge. The term ISM refers to the systematic application of graph theory in such a way that theoretical, conceptual, and computational leverage is exploited to efficiently construct a directed graph, or network representation, of the complex pattern of a contextual relationship among a set of elements. In other words, it helps to identify structure within a system of related elements. It may represent this information either by a digraph (directed graph) or by a matrix. Interpretive Structural Modeling results in a “directed graphic representation of a particular relationship among all pairs of elements in a set to aid in structuring a complex issue area” (Porter, et al., 1980).

There are three broad steps for developing an interpretive structural model. Step 1: ISM begins with an issue or problem (Hansen et al., 1979). Step 2: The next step is to identify the elements that comprise the issue context are listed. Step 3: In the third step, pairs of elements are compared graphically or in a relation matrix, using a contextual relationship, which is mostly a verb or a verb phrase. Typical generic verbs are “influences” or “causes” and verb phrase are “leads to” “is more important than”. Following the selection of the contextual relationship, a graphic representation of the mental model is constructed using the approach described later in the subsequent paragraphs. Mizuno (1988) describes the relationship diagram as a tool that “clarifies intertwined causal relationships in complex problems or situations in order to find appropriate solutions (p. 87).” The relationship diagram, therefore, provides a visual means of mapping out the causal and/or associated relationships in the development of a coherent theory (Anderson et al., 1994). Warfield and Perino (1999) elaborate on the utility of ISM further as the representation of a problematique because it captures the richness and the variety of complex phenomena. A problematique is a graphical portrayal – a structural model – of relationships among members of a set of problems (Warfield and Perino, 1999).

Application of ISM Approach

Having discussed the ISM methodology, we now demonstrate the application of ISM approach to uncover shared mental models. The shared mental model can be treated as a tentative theoretical framework because it captures how respondents commonly understand and explain a phenomenon under consideration. We applied ISM to a well-studied phenomenon in IS – information system use. We took this approach because we wanted to demonstrate the effectiveness of this technique by validating our

1 A process framework is denoted by \( Y = F(C) \) where \( Y \) is the set of outcomes or consequences of a process, \( C \) is the set of considerations or elements in the process, and \( F \) is the network linking the considerations to each other and to the outcomes. Process models are often considered to be complementary to models that lend themselves to variance-theoretic approaches – in other words statistical models.

2 This tentative theory can then be subjected to variance theoretic approaches. For instance, a model generated using ISM could be statistically validated (or, for that matter, invalided).
results with a well-tested theory. We used ISM (Sage, 1977; Warfield, 1973, 1974) to collect, analyze and synthesize the data. Following the three broad steps described above, we first identified the problem at hand. Our problem was to understand IS use behavior at the individual level. The IS usage context that we focused on was spreadsheet usage. We selected spreadsheet usage because it is a well-known and ubiquitously available application and yet there is enough variety in terms of its use and acceptance in different usage contexts.

The next step was to identify and list the elements that are relevant in the problem context. For this we chose to provide the respondents with a superset of elements from Venkatesh et al’s (2003) unified technology acceptance model shown in Table 3. The expectation was that not every user will find every element useful or relevant in the context of IS use. We added an additional element, IT-enabled productivity, to the list of elements from Unified Theory of Acceptance and Use of Technology (UTAUT). This theory is an extension of TAM. Users were given the option of providing any other elements that they believed would influence IS use. As stated before, all research can be construed as part inductive and part deductive. For this research, this step was useful to provide an initial list of elements that each individual user considered to be important to the aspect of IS use. In doing so, we were also able to obtain a set of elements that were common to all respondents.

<table>
<thead>
<tr>
<th>No.</th>
<th>Element</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Performance expectancy (PE)</td>
<td>Performance expectancy is defined as the degree to which an individual believes that using the system will help him or her to attain gains in job performance.</td>
</tr>
<tr>
<td>2</td>
<td>Effort expectancy (EE)</td>
<td>Effort expectancy is defined as the degree of ease associated with the use of the system.</td>
</tr>
<tr>
<td>3</td>
<td>Social influence (SI)</td>
<td>Social influence is defined as the degree to which an individual perceives that important others believe he or she should use the system.</td>
</tr>
<tr>
<td>4</td>
<td>Facilitating conditions (FC)</td>
<td>Facilitating conditions are defined as the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system.</td>
</tr>
<tr>
<td>5</td>
<td>Behavioral intention (BI)</td>
<td>Intention to use the information system</td>
</tr>
<tr>
<td>6</td>
<td>Use behavior (Use)</td>
<td>Actual use (time spent using the information system)</td>
</tr>
<tr>
<td>7</td>
<td>Self-efficacy (SE)</td>
<td>Judgment of one’s ability to use a technology (e. g., computer) to accomplish a particular job or task.</td>
</tr>
<tr>
<td>8</td>
<td>Anxiety (AN)</td>
<td>Evoking anxious or emotional reactions when it comes to performing a behavior (e. g., using an information system).</td>
</tr>
<tr>
<td>9</td>
<td>IT-enabled productivity (IP)</td>
<td>Actual improvements or gains in job performance as a result of using the information system</td>
</tr>
</tbody>
</table>

Table 3. Set of Elements Used in the Study

The third step was to compare pairs of elements graphically or in a matrix. The contextual relationship that we used in this study was “influences.” This forms the essence of the inductive process – where each user performs pair-wise comparisons among elements of the set of variables and a final structure emerges. It is important to reemphasize at this point that although we limited ourselves to the nine elements of UTAUT, we did not specify a research model nor did we specify any variables to be dependent or independent. This is what makes this approach inductive and the theory emergent.

The data elicitation protocol was based on a structured interview. Every respondent was provided with a 9 x 9 matrix shown in Appendix A. The user was instructed that she would have to fill out the upper triangular only. To do that, the user would engage in a pair-wise comparison of elements. For instance, to
compare the PE and EE pair the user would answer “yes” to only one of the following three questions: Does PE influence EE? Does EE influence PE? Are EE and PE unrelated? If the element in the row led to the column element, it was coded as \( \rightarrow \). If the element in the column led to the row element, it was coded as \( \leftarrow \). Lack of a relationship was coded as O. While filling out this the researcher (or the research assistant) would also document the reason(s) for why the respondent chose a particular relationship between two elements. These would typically be direct quotes from the respondent explaining her response. As shown in Appendix B, every respondent was requested to make 36 pair-wise comparisons.

We collected pair-wise comparison data from 88 individuals. These individuals were selected randomly from four organizations to which graduate research assistants were provided access. The average time for an interview was one and a half hours. The interviews typically started with the researcher explaining to the respondent the study protocol. Most of the time was used up by the pair-wise comparisons and an explanation of the constructs along the way. The interview typically ended with the interviewer collecting data on the respondent’s gender, age (age range), experience with computer use (in years) and voluntariness of use of spreadsheets (descriptive). These data items have not been used in this research paper. The interviewer also collected respondents’ justification for their inputs to pair-wise comparisons. This was, typically, a single sentence and, sometimes, a small paragraph.

4. ANALYSIS AND RESULTS

While we collected data based on all the elements shown in Table 3, in this paper we report on a subset of the data – one that pertains to TAM. Table 4 shows how responses were distributed. For instance, we can see that 33 out of 88 respondents believed that effort expectancy influences performance expectancy. In the same cell, 29 out of 88 respondents believed that performance expectancy influences effort expectancy and 26 out of 88 respondents believed that there is no relationship between performance expectancy and effort expectancy.

<table>
<thead>
<tr>
<th></th>
<th>PE</th>
<th>EE</th>
<th>BI</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE</td>
<td>( \rightarrow = 33 )</td>
<td>( \leftarrow = 26 )</td>
<td>( \rightarrow = 31 )</td>
<td>( \leftarrow = 37 )</td>
</tr>
<tr>
<td></td>
<td>( \rightarrow = 29 )</td>
<td>( \leftarrow = 47 )</td>
<td>( \rightarrow = 13 )</td>
<td>( \leftarrow = 13 )</td>
</tr>
<tr>
<td>EE</td>
<td>( \rightarrow = 15 )</td>
<td>( \leftarrow = 27 )</td>
<td>( \rightarrow = 65 )</td>
<td>( \leftarrow = 45 )</td>
</tr>
<tr>
<td></td>
<td>( \rightarrow = 7 )</td>
<td>( \leftarrow = 9 )</td>
<td>( \rightarrow = 13 )</td>
<td>( \leftarrow = 14 )</td>
</tr>
<tr>
<td>BI</td>
<td>( \rightarrow = 50 )</td>
<td>( \leftarrow = 14 )</td>
<td>( \rightarrow = 50 )</td>
<td>( \leftarrow = 50 )</td>
</tr>
</tbody>
</table>

Table 4. Frequency of Responses (numbers show the frequency distribution of responses)

Each individual’s response results in a directed graph. That graph captures how an individual understands the linkages between the elements and can be considered to be the individual’s mental model. The shared mental model across individuals is captured by the degree of overlap across all the individual directed graphs. In order to seek out those common patterns from these data, we employed the straightforward counting technique for aggregating data across individuals based on Kanungo et al., (1999). In order to retain the “shared” component of the mental models we had to define a minimal level of sharing. This
Based on the 50% cut-off, the following relationships (PE \( \rightarrow \) BI, 54.65%; EE \( \rightarrow \) BI, 74.71%; EE \( \rightarrow \) Use, 55.56%; and BI \( \rightarrow \) Use, 64.93%) emerged as being common across a majority of respondents. For the other relationships, no clear relationship emerged as dominant and hence that lack of clarity was coded as the absence of an agreed upon relationship. The final relationship matrix we obtained is shown in Table 5.

<table>
<thead>
<tr>
<th></th>
<th>PE</th>
<th>EE</th>
<th>BI</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE</td>
<td>O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EE</td>
<td></td>
<td>O</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BI</td>
<td></td>
<td></td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Use</td>
<td></td>
<td></td>
<td></td>
<td>O</td>
</tr>
</tbody>
</table>

*Table 5. Final Relationship Matrix*

This translates into the binary relationship matrix shown in Table 6. The elements in the diagonal are 1 because every element (from a reachability\(^3\) standpoint) can “reach” itself. As mentioned before, the “V” coding implies that the row variable influences the column variable and not vice versa. So, for instance, in the case of PE and BI, PE influences BI. That means that the element in row 1 and column 3 (excluding row and column headings) will be 1, while the element in row 3 and column 1 will be 0. The lack of a well-agreed relationship is coded as zeros.

<table>
<thead>
<tr>
<th></th>
<th>PE</th>
<th>EE</th>
<th>BI</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>EE</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>BI</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Use</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

*Table 6. Binary Relationship Matrix (1 implies a relationship exists)*

Next, we identified the levels associated with each element by identifying the reachability and antecedent sets. This iterative process is shown in Tables 7 and 8. Essentially, for all the reachability sets that are proper subsets of antecedent sets, we associate the same level and eliminate those variables or elements for the next iteration. We do this till we have no more reachability and antecedent sets to compare.

<table>
<thead>
<tr>
<th></th>
<th>R(ti)</th>
<th>A(ti)</th>
<th>R(ti) ( \cap ) A(ti)</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1, 3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2, 3, 4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1, 2, 3</td>
<td>3, 4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>2, 3, 4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

*Table 7. The Reachability Set and the Antecedent Set*

\(^3\)“Reachability”, in this case, has to do with relations between elements. Relations between elements are assumed transitive in ISM. In other words, if A “leads to” B and B “leads to C”, then A “leads to C”. From a causality standpoint, every element if perfectly correlated with itself.
Table 7 shows that PE and BE are at the same “level” (level 1) in the hierarchy of the elements that need to be structured. Table 8 shows that BI is at level 2. This partitioning of elements into levels creates the “structural” model that adds value to the graph by preventing it from being a non-directed graph.

<table>
<thead>
<tr>
<th>e_i</th>
<th>R(ti)</th>
<th>A(ti)</th>
<th>R(ti) ∩ A(ti)</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 [BI]</td>
<td>3</td>
<td>3, 4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4 [Use]</td>
<td>3, 4</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

*Table 8. The Reachability Set and the Antecedent Set -II*

The final levels associated with the elements are shown in Table 9. These levels are used to draw the final graph shown in Figure 3.

<table>
<thead>
<tr>
<th>Level</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PE, EE</td>
</tr>
<tr>
<td>2</td>
<td>BI</td>
</tr>
<tr>
<td>3</td>
<td>Use</td>
</tr>
</tbody>
</table>

*Table 9. Final Levels for Elements*

The final structure is constructed using information from Table 7 and Table 4 (steps shown in Appendix B). In this case, there are no transitivities to be removed; hence we retain the graph shown in Figure 3. It is to be noted that this “model” emerged from the data as it were as opposed to us framing the relationships in any predefined manner.

![Figure 3. Final Influence Structure](image)

Given the contextual relationship (*influences*), we can read this diagram to convey the following: performance expectancy (PE) and effort expectancy (EE) influence behavioral intention (BI), which influences IS use (Use). While Figure 3 shows the consensus structure, it is important to keep in mind that there are 88 (number of respondents) possible graphs.

5. **DISCUSSION**

The key objective of this paper was to show how a non-directed and qualitative approach could be used to replicate results from a validated line of research. The final result in Figure 3 shows that the model that emerges from the ISM process is structurally identical to the one suggested by Davis (1989). The primary contribution of this research paper lies in demonstrating that ISM is an efficient and effective method to undertake research that is aimed at theory development based on an inductive approach. In the remainder of this section, we discuss the scientific contributions and implications, practical implications, and limitations of our work.

The structure of equations, variables, and parameters of module is visualized by the ISM hierarchy (Warfield 1976). Since the directed graph consisting of extracted linkages does not explain the whole systematic order of cause-effect relationships, a researcher may not be able to grasp how to calculate an output variable from other input variables and parameters. The structural analysis by ISM classifies...
variables and parameters according to the hierarchical levels, which are obtained by finding a set of nodes that cannot reach any other nodes except the set itself. The hierarchically organized directed graph ensures that only linkages from a lower level to an upper level are included in the entire graph; however there is no reverse directional arc. Nodes at the same level tend to imply that they codetermine or co-influence elements in the subsequent level.

An important part of the entire exercise needs to be underscored at this stage. There was no a priori definition of a dependent variable. Nor was there any a priori definition of an independent variable. However, as argued by Bougon and Weick (1977), who used a variant of this technique as causal maps, the variables on the left, middle and right can be treated as the set of givens, means and ends respectively. As a result, such a model, once it emerges from research, can subsequently be subjected to further empirical scrutiny by subjecting each element pair to tests of correlations individually or using structural equation modeling or path analysis.

It is also important to note from our data collection and analysis process that we have provided a robust framework for stepwise refinement and synthesis. Both the ability to do stepwise refinement and synthesize multiple inputs are important for inductive research. Stepwise refinement is important from the standpoint of localized attention to a specific phenomenon at any given point in time. When a respondent deals with pairs of constructs, it is hoped that she is concentrating on those two constructs only (and operationalizing the ceteris paribus assumption). The essential idea is to build a larger conceptual model piece by piece. Two types of synthesis have also been demonstrated in this paper. The first is the synthesis of pair-wise information into a larger graph and the second type of synthesis is the aggregation of multiple respondents’ viewpoints into a single graph.

Depending on the nature of the contextual relationship, the derived ISM can be considered to be a causal graph or a causal structure. In this study, given the contextual relationship that we have chosen (“influences”), it would be appropriate to consider the emergent graph as a causal model. However, in case we had used “is more important than” as the contextual relation, then the emergent graph would be more meaningful as a priority structure and it would not be even appropriate to consider it to be a causal structure. This is a framework that allows qualitative research to be efficiently replicated. One of the major challenges of qualitative research is that it often has a significant interpretive component. Here the interpretation is left almost entirely to the respondent and the researcher can focus on addressing the rigor of the research protocol.

Like other methodologies, ISM too has its weaknesses. One weakness of this approach includes respondent fatigue. We have found that comparing 36 pairs of elements got the respondents bored – especially toward the later stages of the pair-wise comparison process. In addition, some respondents could not really shut out other elements while dealing with a specific element pair. For instance, a respondent, while comparing PE and EE stated that PE influences EE and her explanation was that “I find the spreadsheet easy to use because I use it a lot; and I use it a lot because it improves my job performance.” While collecting data, we tended to avoid “educating” the respondent in real-time and “contaminating” the data.

It is also natural for other researchers to question the validity of this approach and, in particular, question the relevance of the cut-off value of 50%. Our argument is that if at least fifty percent of respondents agree on something, then there is something of significance there. Just as in the case of p-values (probability of making a type I error) in inferential statistics, if researchers want additional stringency they can reduce the alpha value (maximum allowable type-I error) from 0.05 to 0.01, we could, in our case, increase the threshold to 60 or even 70 percent. However, we have found that it is revealing for the researcher to start with a lower threshold and incrementally increase the threshold to unravel more resilient graph or causal structures.

A third weakness of this approach, as it has been presented here, is that there is no mention of the strength of the relationships between variables. However, there are multiple resolution frameworks for this
problem. In the context of causal or influence maps, there are many approaches that can be used to impute the strength of the causal or relational connection. Techniques like social networks and matrix algebra (Axelrod, 1976; Carley and Palmquist, 1992), system dynamics (Eden et al., 1992), relation algebra (Chaib-Draa, 2002), neural networks (Rossi et al., 1983), and Bayesian probabilities (Nadkarni and Shenoy, 2004) have been used.

We feel that IS researchers adopting an emic stance to IS research can use this approach to complement traditional research approaches. An *emic* analysis of phenomenon is based on internal structural or functional elements of a particular cultural or organizational system. An *etic* analysis is based on predetermined general concepts external to that cultural system (Lovelace, 1984). Since, we have adopted an *emic* perspective that provides the "insider's" or "native's" interpretation of or "reasons" for his or her customs/beliefs, this specific perspective can and should be used to compare and contrast with the *etic* perspective which is the external researcher's interpretation of the same beliefs or relationships. In other words, this approach can be very useful to compare an IS practitioner’s (user’s or manager’s) mental models from what things mean from an analytical, anthropological perspective.

It would be pertinent to point out at this stage that ISM, as a research approach, may have appeared to be overkill when dealing with four variables. We need to keep in mind that our purpose was to demonstrate the efficacy of ISM. Needless to say, ISM is far more effective when a researcher is confronted with a large number of variables (maybe 10 or more) and where causal ambiguities are a result of the novelty of the phenomenon or the inherent complexities.

Finally, the method, as explained in this study, may not appear as inductive as suggested in the introduction. This is because, it may seem that ISM can only be used to generate models of which the elements are already known. ISM, as shown here, is capable of generating the relationships between the elements. In order to elicit a shared mental model, a variant of the approach presented here, would work better. This would involve a two-step approach in which first the elements are collected followed by the relationships. This is the suggested approach for researchers planning to adopt this approach.

### 6. CONCLUSIONS

In this study, we were able to show that a qualitative, open-ended and respondent-driven approach successfully generated a well-accepted model (theory). The main implication for researchers is that the use of such approaches can form an extremely effective and efficient method for capturing the shared mental models of IS practitioners. It allows IS researchers to perform research that is interpretive and grounded in data efficiently. This is important because, in many instances, IS phenomena are so dynamic and changes occur so fast in the IS domain that it is unreasonable to expect researchers to study stable phenomena and replicate or disconfirm results obtained by other researchers. Moreover, what happens in the field, more often than not, drives academic IS research – and not the other way round. Hence, it is important to employ methodologies like ISM that efficiently allow the capture and synthesis of practitioners’ viewpoints. From a practical perspective this approach is even more valuable because this approach is context sensitive and can be replicated effectively by researchers across contexts. By using ISM we have been able to “focus on the concerns of practice, provide real value to [IS] professionals Benbasat and Zmud (1999, p. 5)” and apply a balance of pragmatic and academic tone.

**References**


Warfield, J. N. Structuring complex systems, Battelle Monograph Number 4, Battelle Memorial Institute, Columbus, OH, 1974.


Warfield, J. N. Patterns and Behavior, Institute for Advanced Study in the Integrative Sciences, George Mason University, January, 1998.


Appendix A

<table>
<thead>
<tr>
<th>Performance expectancy</th>
<th>Effort expectancy</th>
<th>Social influence</th>
<th>Facilitating conditions</th>
<th>Behavioral intention</th>
<th>Use behavior</th>
<th>Self-efficacy</th>
<th>Anxiety</th>
<th>IT-enabled productivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance expectancy</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Effort expectancy</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Social influence</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facilitating conditions</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavioral intention</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use behavior</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-efficacy</td>
<td></td>
<td></td>
<td>34</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT-enabled productivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table shows the overall data collection framework. Users were requested to fill in this table with \( \Phi \), \( \vec{\Phi} \) or O based on the protocol explained in the body of the paper. Cells that are shaded darker (cell numbers 1, 2, 4, 5, 9, 11, 12 and 27) have been used for analysis in this paper. The cell numbers were used by the respondents and interviewers to link their responses with justifications for those responses.

Appendix B

The following steps outline how the interpretive structural modeling methodology is implemented:

i) **Identify elements**: The elements of the system are identified and listed. This may be achieved through past research, brainstorming, or using the nominal group technique.

ii) **Establish a contextual relationship**: A contextual relationship between elements is established, depending upon the objective of the modeling exercise. This is a verb or verb phrase like “increases” or “is more important than” or “leads to.”
**iii) Prepare a reachability Matrix:** For the contextual relation from element $E_i$ to $E_j$, but not in the reverse direction, then element $E_{ij} = 1$ and $E_{ji} = 0$ in RM. For the contextual relation from $E_j$ to $E_i$, but not in the reverse direction, then element $E_{ij} = 0$ and $E_{ji} = 1$ in RM. For an interrelation between $E_i$ and $E_j$ (both directions), then element $E_{ij} = 1$ and $E_{ji} = 1$ in RM.

To represent that $E_i$ and $E_j$ are unrelated, then element $E_{ij} = 0$ and $E_{ji} = 0$ in RM.

**iv) Perform level partitioning:** Level partitioning is done in order to classify the elements into different levels of the ISM structure. For this purpose, two sets are associated with each element $E_i$ of the system - A *Reachability Set* ($R_i$) that is a set of all elements that can be reached from the element $E_i$, and an *Antecedent Set* ($A_i$), that is a set of all elements that element $E_i$ can be reached by. In the first iteration, all elements, for which $R_i = R_i \cap A_i$, are Level I Elements. In successive iterations, the elements identified as level elements in the previous iterations are deleted, and new elements are selected for successive levels using the same rule. Accordingly, all the elements of the system are grouped into different levels.

**v) Develop canonical matrix:** grouping together elements in the same level develops this matrix. The resultant matrix has most of its upper triangular elements as 0, and lower triangular elements as 1. This matrix is then used to prepare a Digraph.

**vi) Draw the digraph:** Digraph is a term derived from **Directional Graph**, and as the name suggests, is a graphical representation of the elements, their directed relationships, and hierarchical levels. The initial digraph is prepared on the basis of the canonical matrix. This is then pruned by removing all transitivities, to form a final digraph.

**vii) Create the interpretive structural model:** The ISM is generated by replacing all element numbers with the actual element description. The ISM therefore, gives a very clear picture of the system of elements, and their flow of relationships.
**THE SIGNIFICANT OTHERS OF SUBJECTIVE NORM - A SCIENTOMETRIC STUDY OF SUBJECTIVE NORM IN IS TOP-JOURNALS OVER TWO DECADES**

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0743.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Technology adoption, Behavioural theories, Adoption, Social inclusion</td>
</tr>
</tbody>
</table>
THE SIGNIFICANT OTHERS OF SUBJECTIVE NORM -

A SCIENTOMETRIC STUDY OF SUBJECTIVE NORM IN IS TOP-JOURNALS OVER TWO DECADES

Eckhardt, Andreas, Goethe-University Frankfurt, Centre of Human Resources Information Systems, Grueneburgplatz 1, 60323 Frankfurt a. Main, Germany, eckhardt@is-frankfurt.de

Abstract

Undoubtedly social influence in IS research is an issue that needs to be much more clearly examined. In order to assuage the wants for more research in this field we conducted a scientometric analysis of subjective norm in the IS top-journals of the JAIS ranking for the last two decades. In total 113 empirical and conceptual research models predominately in adoption research contained the factor subjective norm. The results revealed that subjective norm is just in seven of ten models a significant antecedent. To gain more knowledge about this problem we followed the ideas of social psychology and marketing researchers and correlated the strength of the impact of subjective norm with its measurement as well as the impact of intention with the impact on other endogenous variables. The results show a significant negative correlation between the significant antecedent subjective norm and its original measurement, the perceived opinion of important others. Furthermore it revealed a significant negative correlation between the significant impact of subjective norm on intention and the significant impact on other endogenous variables.

Keywords: Subjective Norm, Social Influence, Scientometric Study, Meta-Analysis
1 INTRODUCTION

“There’s one advantage to being 102. There’s no peer pressure.” (Wolfberg 2008)

In older ages the social influence of referent groups in private, public or workplace contexts might be not that important anymore as Dennis Wolfberg - a US comedian - assumed but it is exceptionally important in the whole life before. As we are confronted almost throughout our daily complete behavior by the opinions, actions and advices of important others (Fishbein and Ajzen 1975; Triandis 1971) it is not surprising that subjective norm or social factors are an important determinant for an individual’s intention and behavior respectively to use or adopt a specific technology or information system (e.g. Thompson et al. 1991; Lewis et al. 2003; Lee et al. 2006). However two decades after the introduction of the Technology Acceptance Model (TAM) (Davis et al. 1989) IS research still struggles to define and measure social influence appropriately. Various researchers have claimed for "more sophisticated methods for assessing the specific types of social influence" (Davis et al. 1989), “additional research that clarifies the precise role of social pressure in technology acceptance” (Agarwal 2000) or “further research that should study if this factor generate any direct impact on intention to adopt” (Lu et al. 2005). Compared to other classic technology adoption parameters subjective norm or social influence never reached their path significance for the impact on an individual’s behavioral intention or other endogenous variables (Schepers and Wetzels 2007). Therefore other researchers suggest “that social norms need to be conceptualized in a more distinguishing manner to capture the nuances of the social environment” (Srite and Karahanna 2006).

In the JAIS special issue on TAM in 2007 Benbasat and Barki recommended for further research in this field to get back to the underlying Theories of Reasoned Action (Fishbein and Ajzen 1975) and Planned Behavior (Ajzen 1985) (Benbasat and Barki 2007). So we focus our research on two specific aspects of social influence in the underlying theories: its concrete measurement and its general role in dependence with other parameters.

Surprisingly besides being insignificant (e.g. Mathieson 1991; Hsieh et al. 2008) or excluded due to insignificance (e.g. Davis et al. 1989; Ha et al. 2007) from numerous empirical models, the parameter subjective norm or social factors was never critically questioned in IS research like in other fields as social psychology or marketing. Researchers with these backgrounds particularly criticized the basic measurement of subjective norm in the underlying theories (Fishbein and Ajzen 1975; Triandis 1971) as not distinguishing enough to analyze social influence appropriately (e.g. Ahtola 1976; Miniard and Cohen 1983; Liska 1984). In their opinion is “the person’s perception that most people who are important to him think he should or should not perform the behavior in question” (Fishbein and Ajzen, 1975) not adequately as an individual does not feel the pressure of collective important others, but of individual groups as peers or superiors in a working environment respectively friends or parents in a private environment (Eckhardt et al. 2008). Within this approach we aim to reassess these points of critique for IS research by conducting a scientometric study of all IS top-journals of the JAIS ranking (Lowry et al. 2004) for the last twenty years. With the results of this study and the help of a correlation analysis we would like to answer the following research question:

RQ1: Is there a correlation between how Subjective Norm is measured and the significance of the impact?

Furthermore another point of critique regarded the relationship between attitudinal and normative beliefs (Liska 1984; Miniard and Cohen 1983). Social psychology and marketing researchers discussed that a further weakness of the TRA (Fishbein and Ajzen 1975) is its complete ignorance of a causal relationship between attitudes and social norms (Liska 1984). Fishbein and Ajzen regarded these effects as totally independent. It was mentioned that these effects might be independent by definition but not independent by reason as well as statistically interactive (Liska, 1984). Despite the inclusion of potential interaction and dependency effects in the Theory of Planned Behavior (Ajzen 1985) and the confirmed importance of a further examination (Yang and Yoo 2004) this fact was barely observed in further research approaches. By analyzing the data of 113 empirical and conceptual
models containing the factor subjective norm or social influence found through our scientometric study we address the following second research question:

RQ2: Are there any correlations between the influence of subjective norm on behavioral intention and the influence on other endogenous variables?

We answer both research questions with a correlation analysis in SPSS v.16. Beforehand we describe the theoretical derivation of the observed factor subjective norm or social influence and its function in the underlying Theories of Reasoned Action (TRA) and Planned Behavior (TPB) (Fishbein and Ajzen 1975; Ajzen 1985) as well as Interpersonal Behavior (TIPB) (Triandis 1971; 1980) in section 2. The following section includes the introduction of our research method (scientometric study) and the description of our data pool plus details regarding the search process and the database access. After the descriptive statistics and the correlation analyses in section 4 we discuss the results of our approach. By concluding our scientometric study about subjective norm we provide implications for further research in section 6.

2 THEORETICAL BACKGROUND

After the detailed description of our research objectives and its related research question in the introduction we will use this section to briefly describe the theoretical background of subjective norm respectively social influence in IS research especially in the field of technology adoption, acceptance and usage. Furthermore we introduce the underlying theories and models which include the parameter subjective norm.

2.1 The Term Subjective Norm

The origin of the term subjective norm or social influence lies in the cradle of social psychology research back in the early fifties of the 20th century. Luminaries in social psychology research as Solomon Asch, Kurt Lewin or Leon Festinger introduced and experimentally proofed the concept of social influence as a pressure of conformity on an individual human being to act conform to the behavior of a distinct group or person (Asch 1951; Lewin 1952). This also includes a continuous comparison from an individual’s point of view with the behaviors, opinions, actions of peers (Festinger 1954). Social influence expresses itself in an individual person as a change of thinking or feeling concerning a specific behavior due to communication with another individual or a person. In 1962 the knowledge about this topic was further deepened by the work of Everett M. Rogers on the diffusion of innovations. Within his approach Rogers describes social influence as norms or the roles of opinion leaders and change agents in innovation diffusion in a distinct social system which is defined as “a set of interrelated units that are engaged in joint problem solving to accomplish a common goal” (Rogers 2003)

Altogether the term social influence in IS research could be predominately related to the concepts of peer pressure and social comparison in a distinct social system (Eckhardt et al. 2009). The construct subjective norm as well as social factors was than founded about 10 years later, introduced within the competing Theories of Reasoned Action (Fishbein and Ajzen 1975) and Interpersonal Behavior (Triandis 1971). A first approach adapting the factor social influence for management literature and forming a factor of subjective norm was made by Gerald Salancik and Jeffrey Pfeffer who built a conceptual framework to describe social information processes on an individual’s job attitude (Salancik and Pfeffer 1978). The introduction for subjective norm in IS research started with an exclusion from the Technology Acceptance Model (Davis et al. 1989) due to insignificance. Subsequent approaches in technology adoption including subjective norm kept the balance between significant (Thompson et al. 1991; Taylor and Todd 1995) and insignificant results for the factor of subjective norm (Mathieson 1991; Dishaw and Strong 1999). Beside these mixed results one thing remained over the years, a large number of IS researchers requested a further investigation of this “unloved child” of technology adoption research (Davis et al. 1989; Agarwal 2000; Lu et al. 2005).
Beyond technology adoption research subjective norm or social influence was comparatively seldom used in IS research. But there are some examples as Austin et al. (1993) who evaluated the factor social influence in the context of who controls the technology in group support systems (Austin et al. 1993). Other researchers included social influence in the context of media richness concerning electronic communication (Schmitz and Fulk 1991). For this purpose Janet Fulk and her co-authors Schmitz and Steinfield developed a social influence model of technology use to incorporate the influence of workplace referents such as superiors or co-workers on the use of electronic communication (Fulk et al. 1990). Also in the field of IT turnover first approaches were made to observe the impact of social influence in the form of social support on an IT professional’s turnover intention (Lee 2002).

Overall it can be concluded that a majority of empirical and conceptual models including any term of subjective norm, social influence, normative beliefs, social norms, etc. in IS research is related to the field of technology adoption, nevertheless social influence appears in a great variety of synonyms as social support in IT turnover (Lee 2002), as social presence in telecommunication (Short et al. 1976) and in online auctions bidding behavior (Rafaeli and Noy 2005) or as social status in virtual team management (Austin et al. 1993) in other IS domains as well.

2.2 Basic Theories Including the Term Subjective Norm

As concluded in the previous subsection a majority of empirical and conceptual models including a factor for social influence in IS research could be related to technology adoption research. The list of popular and frequently cited adoption models is countless, starting with the most famous Technology Acceptance Model (TAM) (Davis et al. 1989), the Model for PC Utilization (Thompson et al. 1991), the Model of Adoption of Technology in Households (Venkatesh and Brown 2001) up to the UTAUT Model (Venkatesh et al. 2003), just to name a few. All of these adoption models and almost all remaining of the not mentioned ones base upon two competing behavioral theories from social psychology research: The Theory of Reasoned Action (Fishbein and Ajzen 1975 and its extension the Theory of Planned Behavior (Ajzen 1985) as well as the Theory of Interpersonal Behavior (Triandis 1971; Triandis 1980).

A lot more recently published research models have their seeds in the Theory of Reasoned Action (TRA) and the Theory of Planned Behavior (TPB) by Icek Ajzen and Martin Fishbein (Fishbein and Ajzen 1975; Ajzen 1985). The objective of their theories is to explain the determinants that predict an individual’s specific behavior and the behavioral intention to adopt the respective behavior. The antecedents for an individual’s behavioral intention are an individual’s attitude toward the behavior, perceived behavioral control and subjective norm. Subjective norm is referred to as “the perceived social pressure to perform or not to perform the behavior as a sum of the perceived expectations of specific referent individuals and/or groups weighted by the individual's "motivation to comply" with those expectations” (Fishbein and Ajzen 1975).

Interestingly although accepted within the psychological literature and founded before the TRA the Theory of Interpersonal Behavior by Harry Triandis (1971) has not been used frequently within the IS context. Just a few but highly significant approaches include this theory as underlying (Thompson 1991; Lee et al. 2001). Already in 1971 Triandis argued “that an individual’s behavior is influenced by social norms, which depend on messages received from others and reflect what individuals think they should do” (Triandis 1971). Nine years later, Triandis expanded this term and called it social factors, that is, "the individual's internalization of the reference groups' subjective culture, and specific interpersonal agreements that the individual has made with others, in specific social situations" (p. 210). The reference groups’ subjective culture includes norms (self-instructions to do what is perceived to be appropriate and by members of a distinct culture in a certain environment and situation); roles (which are related with behaviors that are considered appropriate but concerned to persons holding a particular position in a social system, society or group); and self concept (abstract
categories with highly emotional components) (Triandis 1980). Both extended theories (Ajzen 1985; Triandis 1980) are shown in the following Figure 1.

![Figure 1. The Theory of Planned Behavior (Fishbein and Ajzen 1975; Ajzen 1985) and the Theory of Interpersonal Behavior (Triandis 1971; Triandis 1980)](image)

To what extent these theories are related to the criticism for measurement and role definition of subjective norm as described in the introduction will be analyzed with the help of our scientometric study in section 4.

3 RESEARCH METHODOLOGY

The objectives of this section are to introduce our research method, the scientometric study, to argue why we chose the JAIS ranking of 2004 (Lowry et al. 2004) as our data pool and to describe how we searched through this large amount of data to find valuable content for our scientometric study about subjective norm in IS research.

3.1 Research Method

We chose scientometrics as research method for our literature review about subjective norm. Reason for that choice was the adequacy of the scientometric analysis for our research approach as it answers particular questions about way and form IS researchers publish their contents. For example researchers defined scientometrics as the quantitative study of research (Davis 2001) or the scientific study of the process of science (Lowry et al. 2004). In this context it needs to be outlined how this research method differs from regular surveys. Hunter et al. (1982) highlighted the major difference between both forms. A survey is used to collect data about people’s behavior, opinion or background. On the contrary a scientometric analysis focuses on the article itself and not the observed people (Hunter et al. 1982). With employed tools as citation analysis or meta-analysis a scientometric analysis observes the affiliations of authors, paper abstracts and texts or references and appendices. Overall scientometric studies are considered to become very important and highly valued for scientific research in future (Straub 2006). Detailed information about the ranking we observed and the way we accessed the included journals is described in the following subsection.

3.2 Data Pool and Included Publications

If you conduct a scientometric analysis major questions a priori are: How far do you want to go and where you got to draw the line? Unlike other already conducted quantitative meta-analyses on subjective norm that included dissertation theses and conference proceedings (Schepers and Wetzels 2007) or journal articles from other research domains as marketing or social psychology (Eckhardt et al. 2008) we focus in our approach solely on IS top journals. In order to exactly define and concentrate our approach we decided to choose a journal ranking as boundaries for the scientometric study. There
is a broad range of IS journal rankings at present with different foci national and international so we had to analyze several rankings to find the perfect fit to our research objectives. Finally we decided to use the JAIS ranking by Lowry et al., published in 2004. It is one of the most cited IS journal rankings and part of the MIS journal rankings of the Association for Information Systems and widely known as an extremely comprehensive ranking of IS journals (Mbarika et al. 2005) with an outstanding reputation. In total we included all 48 journals of the five world rankings in our scientometric study. We started our approach two decades ago and searched through every single issue since 1988 of all 48 journals included in the JAIS ranking (Lowry et al. 2004), more than 20,000 articles overall. We accessed these journals for our scientometric study via literature online databases and electronic libraries. The actual search process with its related search criteria will be described in the following subsection.

3.3 Database Search Process and Criteria

With the objective to conduct our scientometric study we accessed the included IS journals via ten databases and electronic libraries. These sources were in alphabetical order: ACM Digital Library, AIS Library, EBSCOhost, Elsevier, Emerald, IEEE Xplore, JSTOR, Palgrave Macmillan, ScienceDirect and Wiley InterScience. IS journals which were not accessible via one of these sources were retrieved with the help of colleagues in other universities and IS research institutions.

Two main search techniques were mostly provided by the respective databases or e-libraries, the “General Search and the “Advanced Search”. Both search techniques include the Boolean operators (“AND”+”OR”) to facilitate the search with more search items. Like other literature research approaches (e.g. Dwivedi et al. 2008) we predominately used the “General Search” as it allowed us to repeat our searches frequently with consistent results and without any confusion. For the purpose of finding all relevant IS journal articles we started our search with the term “Subjective Norm” and several synonyms or related terms already found in literature (Karahanna et al. 1999; Lewis et al. 2003: Lee et al. 2006; Kim et al. 2007) and used in other meta-analyses (Schepers and Wetzels 2007) as “Social Influence”, “Normative Beliefs”, “Social Norm”, “Social Pressure”, “Social Exchange”, “Peer Group Influence”, “Peer Group Norm”, “Peer Group Pressure”, “Peer Pressure”, “Superiors Influence”, “Superiors Pressure”, “Friends Influence”, etc. Our scientometric search was restricted to incidences of any of these search terms appearing in the body of the text or the article title. In order to sharpen and improve our results we added further information as frequently underlying models or theories which imply subjective norm as well as the names of the related and frequently cited authors. Therefore in databases as EBSCOhost we used extended search terms as the following example for the search through the journal Information Systems Research:

JN "Information Systems Research" AND (TX "technology acceptance model" OR TX "Venkatesh" OR TX "Ajzen" OR TX "subjective norm" OR TX "TAM" OR "TAM"2 OR TX "theory of planned behavior" OR TX "theory of reasoned action" OR TX "social influence" OR TX (important N2 me) OR TX "normative beliefs")

This search style resulted in the extraction of 1,856 articles providing topics and content related to social influence in IS research. All 1,856 records were then analyzed manually to examine and crosscheck their relevance for the overall study. A number of further analyses were then performed afterwards to categorize the findings due to their title, author, year of publication, research subject as well as most important to the individual role of subjective norm its significance, measurement and impact as exogenous and endogenous variable on further endogenous variables. After the categorization process all results were stored and coded within a database. In total 113 articles (107 empirical models and 6 conceptual frameworks) were included for our scientometric study. The results were first coded in SPSS v.16 and then counted and percentage values generated. The complete descriptive statistics is shown in the following Table 1. The empirical evaluation of these data and the implying correlation analyses are presented in the following section 4.
<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>Brah &amp; Cone</td>
<td>Investigating the Impact of Personal Computer Use on the Work of Small Business Personnel</td>
</tr>
<tr>
<td>1990</td>
<td>Davis et al.</td>
<td>A Conceptual Model of Computer Acceptance and Usage</td>
</tr>
<tr>
<td>1991</td>
<td>Venkatesh et al.</td>
<td>The Technology Acceptance Model: Understanding Users' Adoption of Information Technology</td>
</tr>
<tr>
<td>1992</td>
<td>Davis et al.</td>
<td>Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology</td>
</tr>
<tr>
<td>1993</td>
<td>Venkatesh</td>
<td>User Acceptance of Computer Technology: A Theoretical Disposition Model</td>
</tr>
<tr>
<td>1994</td>
<td>Venkatesh</td>
<td>A Conceptual Model of Computer Acceptance and Usage</td>
</tr>
<tr>
<td>1995</td>
<td>Venkatesh</td>
<td>The Technology Acceptance Model: Understanding Users' Adoption of Information Technology</td>
</tr>
<tr>
<td>1996</td>
<td>Venkatesh</td>
<td>User Acceptance of Computer Technology: A Theoretical Disposition Model</td>
</tr>
<tr>
<td>1997</td>
<td>Venkatesh</td>
<td>A Conceptual Model of Computer Acceptance and Usage</td>
</tr>
<tr>
<td>1998</td>
<td>Venkatesh</td>
<td>The Technology Acceptance Model: Understanding Users' Adoption of Information Technology</td>
</tr>
<tr>
<td>1999</td>
<td>Venkatesh</td>
<td>User Acceptance of Computer Technology: A Theoretical Disposition Model</td>
</tr>
<tr>
<td>2000</td>
<td>Venkatesh</td>
<td>A Conceptual Model of Computer Acceptance and Usage</td>
</tr>
<tr>
<td>2001</td>
<td>Venkatesh</td>
<td>The Technology Acceptance Model: Understanding Users' Adoption of Information Technology</td>
</tr>
<tr>
<td>2002</td>
<td>Venkatesh</td>
<td>User Acceptance of Computer Technology: A Theoretical Disposition Model</td>
</tr>
<tr>
<td>2003</td>
<td>Venkatesh</td>
<td>A Conceptual Model of Computer Acceptance and Usage</td>
</tr>
<tr>
<td>2004</td>
<td>Venkatesh</td>
<td>The Technology Acceptance Model: Understanding Users' Adoption of Information Technology</td>
</tr>
<tr>
<td>2005</td>
<td>Venkatesh</td>
<td>User Acceptance of Computer Technology: A Theoretical Disposition Model</td>
</tr>
<tr>
<td>2006</td>
<td>Venkatesh</td>
<td>A Conceptual Model of Computer Acceptance and Usage</td>
</tr>
<tr>
<td>2007</td>
<td>Venkatesh</td>
<td>The Technology Acceptance Model: Understanding Users' Adoption of Information Technology</td>
</tr>
<tr>
<td>2008</td>
<td>Venkatesh</td>
<td>User Acceptance of Computer Technology: A Theoretical Disposition Model</td>
</tr>
</tbody>
</table>

Table 1: Findings of the Search Study

Proceedings ECIS 2009
4 RESULTS

After the theoretical derivation of subjective norm and the description of its explicit role and measurement in the underlying theories in section 2 and the introduction of our research method in section 3 we present in this section the descriptive statistics of our scientometric study as well as the evaluation of the collected data with the help of a correlation analysis. For information about these journals including the parameter subjective norm as title, authors, name of journal and year of the publication, see Table 1. To answer the research questions asked in the introduction we collected explicit data regarding the measurement and of subjective norm.

Generally spoken the degree of significance of subjective norm or social influence in the empirical models analyzed is low. We found the parameter in 31.7 percent of all empirical models observed as an insignificant antecedent for an individual’s behavioral intention or further respective parameters as the specific behavior, attitude, perceived usefulness, perceived ease of use, perceived enjoyment, trust, etc. In 68.3 percent of all empirical models subjective norm was found as a significant antecedent for the parameters mentioned before. Compared to other meta-analyses as one by Wetzels and Schepers (2007) this represents a slightly lower percentage for the significant impact of subjective norm. They found the parameter as a significant antecedent for intention in 86.4 percent of all articles observed, for perceived usefulness in 91.7 percent and for perceived ease of use in 66.7 percent of all articles. However compared to our approach they used a much smaller sample size with 51. Additionally they included conference proceedings as well as not ranked IS journals in their approach and did not limit their findings to IS top-journals as we did. Compared to other regular adoption parameters as perceived usefulness, perceived ease of use or attitude, who were found significant on average in more than 90 percent of all cases in other meta-analyses (Ma and Liu 2004; King and He 2006) subjective norm is substandard regarding its significance.

Following the descriptive statistics of our scientometric study we analyze the correlations within the dataset with regard to our research questions. Therefore we collected the individual measurement items of each model observed according the following classification (Eckhardt et al. 2008): (1) private environment (Private), (2) workplace environment (Job), (3) important others (Important) and (4) public environment (Government). Furthermore we analyzed if subjective norm significantly influences an individual’s behavioral intention (SigInt), not significantly influences intention (NoSigInt), significantly intention with determinants (SigIntDet), not significantly influences intention with determinants (NoSignIntDet), significantly influences other endogenous variables (SigOther) or not significantly influences other endogenous variables (NoSigOther).

The following Table 2 shows the correlation between the values and factors mentioned above. As one can see there are a few interesting outcomes. With respect to our first research question we can note a significant negative correlation for p < 0.01 between subjective norm as a significant antecedent for intention and its measurement with the items of important others. Additionally interesting in this context is the significant positive correlation between subjective norm as a significant antecedent for intention and its measurement with influence groups of working environment. Therewith we can conclude that if subjective norm has a significant impact on behavioral intention in most cases the measured items are influence groups from the working environment and not the originally measured important others of the underlying definition by Fishbein and Ajzen’s Theory of Reasoned Action (Fishbein and Ajzen 1975). Concerning our second research question we evaluated a significant negative correlation between the significant influence of subjective norm on intention and the significant influence of subjective norm on other endogenous variables. So we can summarize for this correlation that if subjective norm has a significant impact on intention a significant impact of subjective norm on further endogenous variables as attitude is uncommon.
<table>
<thead>
<tr>
<th>Correlations</th>
<th>SigInt</th>
<th>NoSigInt</th>
<th>SigIntDet</th>
<th>NoSignIntDet</th>
<th>SigOther</th>
<th>NoSighOther</th>
<th>Private</th>
<th>Job</th>
<th>Important</th>
<th>Government</th>
</tr>
</thead>
<tbody>
<tr>
<td>SigInt</td>
<td>Correlation by Pearson</td>
<td>-.314**</td>
<td>.025</td>
<td>-.209*</td>
<td>-.055</td>
<td>.000</td>
<td>-.001</td>
<td>.796</td>
<td>.000</td>
<td>.314**</td>
</tr>
<tr>
<td></td>
<td>Significance (2-sided)</td>
<td>N</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
</tr>
<tr>
<td>NoSigInt</td>
<td>Correlation by Pearson</td>
<td>.344**</td>
<td>.000</td>
<td>.000</td>
<td>-.004</td>
<td>.032</td>
<td>.016</td>
<td>.055</td>
<td>.027</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Significance (2-sided)</td>
<td>N</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
</tr>
<tr>
<td>SigIntDet</td>
<td>Correlation by Pearson</td>
<td>.337**</td>
<td>.000</td>
<td>.000</td>
<td>-.004</td>
<td>.032</td>
<td>.016</td>
<td>.055</td>
<td>.027</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Significance (2-sided)</td>
<td>N</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
</tr>
<tr>
<td>NoSignIntDet</td>
<td>Correlation by Pearson</td>
<td>-.232</td>
<td>-.035</td>
<td>-.238</td>
<td>-.055</td>
<td>.000</td>
<td>.000</td>
<td>.055</td>
<td>.027</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Significance (2-sided)</td>
<td>N</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
</tr>
</tbody>
</table>

**. The correlation is significant for p < 0.05 (2-sided).
**. The correlation is significant for p < 0.01 (2-sided).

Table 2. Correlation Analyses

4.1 Limitations

Due to the research design our approach is limited through some facts. We might have found not all articles including subjective norm in our scientometric study because of the search mode we conducted within the literature databases. Furthermore the results of the correlation analysis for our second research question must be a little toned down as the significant antecedent subjective norm as for intention is not only negatively correlated with the significant antecedent subjective norm on other endogenous variables but also with the non significant antecedent subjective norm.

5 DISCUSSION

In a current research article analyzing the social influence of workplace referents on an individual’s IT adoption and non-adoption the authors sum up that a single cumulative subjective norm measure might be too naïve (Eckhardt et al. 2009). With regard to our results (see Table 3) this statement can be supported as we found a significant negative relationship between a significant impact of subjective norm and the original measurement of important others in the Theory of Reasoned Action (Fishbein and Ajzen 1975). Hence researchers using this theory or one of its numerous advanced versions need
to take into account that individuals perceive the influences, behavioral patterns and actions of their social environment quite differently as opinions and suggestions of their life partner, parents, children, etc. could not be equally treated to these of superiors, peers or subordinates. Comparable to the differences between regular IT adoption and household adoption social influence of referent groups differs according to the respective environment. A first step could be a classification of this social environment in private, job and public environment like we did in our approach. Nevertheless a precise social environment analysis before empirical surveys will be a necessary precondition for future approaches in technology adoption containing a construct for social influence.

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Result</th>
<th>Implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is there a correlation between how Subjective Norm is measured and the significance of the impact?</td>
<td>Significant negative correlation between the original measure “important others” and a significant impact</td>
<td>Several individual social influences instead of one collective subjective norm implying the item of important others need to be measured</td>
</tr>
<tr>
<td>Are there any correlations between the influence of subjective norm on behavioral intention and the influence on other endogenous variables?</td>
<td>Significant impact of subjective norm on behavioral intention is significantly negative correlated with the significant impact of subjective norm on other endogenous variables</td>
<td>Rethinking the general role of subjective norm in IT adoption its interdependencies and causal relationships</td>
</tr>
</tbody>
</table>

Table 3. Research Results and Implications

6 IMPLICATIONS FOR FUTURE RESEARCH

We motivated our approach in the beginning with the results and implications for further research by other researchers and can conclude in the end that there is still a distinct need to measure, define and use the parameter in IS research in the appropriate way. Our scientometric study of subjective norm in IS top-journals over the last two decades revealed that this factor is just in seven of ten cases a significant antecedent for behavioral intention or other endogenous variables. Compared to other regular parameters of predominately empirical adoption models as attitude, perceived usefulness or perceived ease this represents a comparatively low percentage (Wetzels and Schepers 2007; Ma and Liu 2004; King and He 2006). In order to clarify the understanding of social influence in IS research and to solve the problem of a frequent insignificance of subjective norm we adapted ideas and suggestions of social psychology and marketing researchers and analyzed the measurement and dependence of this parameter. Based on the data of 113 empirical and conceptual research models including subjective norm we conducted a correlation analysis to answer our research questions regarding measurement and dependence of subjective norm. The results showed a significant negative correlation between a significant impact of subjective norm on behavioral intention and the original measurement of important others by Fishbein and Ajzen (1975). Although we have not found a significant correlation between insignificant results for subjective norm and the measurement of important others, the results reveal that the general idea of the basic measurement detractors (e.g. Ahtola 1976; Miniard and Cohen 1983; Liska 1984) seem to constitute a reasonable start for further research. A further result based on the thoughts of social psychology and marketing researchers concerning our second research question reveals that a significant impact of subjective norm on behavioral intention is significantly negative correlated with the significant impact of subjective norm on other endogenous variables. This implies that there might be interdependencies between subjective norm and other parameters as attitude (Liska 1984; Miniard and Cohen 1983). Further research needs to clearly examine in this point how subjective norm or social influence could be formed without interdependencies to explain and affect an individual’s intention and within his actual behavior.
References


Asch, S.E. (1951). Effects of group pressure on the modification and distortion of judgments, Groups, leadership, men.


A SAD STORY: THE CASE OF CONSTRAINED INFRASTRUCTURES CAUSED BY IT

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0199.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>E-government, Case Study, Inter-organisational change, Global IT management</td>
</tr>
</tbody>
</table>
Abstract

Customs is one domain which has faced the need for information systems in a globalizing world. In customs there is a need for close collaboration between business and government in order to build robust technological infrastructures supporting processes related to export and import of goods. This has led to a move towards eCustoms. This case study reports from a longitudinal study of a Danish exporter which has faced the challenges of implementing eCustoms in its daily operations. The introduction of eCustoms to replace the existing customs infrastructure based on the Single Administrative Document (SAD) has led to new challenges for the Danish exporter such as de-standardized procedures and increased reporting. The particular contribution of the study is the observed longing to the “good old paper-based days” where things were more standardized than after the digitalization of the processes. The study is interpreted through the lens of “double infrastructures”. The underlying assumption guiding the study is that eCustoms represents one infrastructure on top of another well established infrastructure of global trade.

Keywords: eGovernment, interorganizational systems, global information systems, cross national studies.
1 INTRODUCTION

One effect of a globalized world is the increased traffic of goods from one country to another. That calls for new customs procedures. The phenomenon customs has been known for centuries and the privilege of duty leverage has been executed by local rulers for ages. The rulers’ power scope was often based on features of the landscape such as mountain passes or straits such as the Strait of Gibraltar or the entrance to the Baltic Sea, which could be monitored and governed with relatively few resources such as groups of scouts overlooking the landscape. The scouts could then target caravans or ships with goods via horseback or small boats. With the emergence of national states the customs procedures have developed to more comprehensive and predictable procedures. Furthermore, has the development of technologies led to much more robust monitoring techniques and control mechanisms which have led to an extensive administrative burden for businesses. The domain of eGovernment, of which eCustoms can be seen as a subset, has recognized that IS holds the capacity of reduction of the administrative burden (Lenk 2002) and any reduction in administrative tasks is welcomed by businesses which often consider their interaction with the public sector to be a burdensome task (Fountain & Osorio-Urzua 2001).

What has brought the issue of customs on the information systems (IS) agenda is an increased call for efficiency, security and streamlining of procedures (Tan, Klein, Rukanova, Higgins & Baida 2006). Somehow the current standardized customs infrastructure based on the Single Administrative Document (SAD) has to be improved. It is a general belief that IS hold the capacity of strengthening efficiency, providing tools for security and furthermore being an instrument for streamlining of procedures. This has led to a keen interest for eCustoms from national and supra-national bodies such as the EU and WTO. eCustoms represent a complex mix of stakeholders from national governments, international trade, regulatory bodies and businesses. To increase complexity on top there is the technology dimension which has to accommodate the needs for interoperability, transparency, security and so on. The interorganizational information systems involving public and private parties require robust standards and infrastructures which can support the mutual interest in seamless administration of export and import.

The impact of implementing administrative IS in organizations is well studied and studies that report challenges and problems are numerous (McKeen & Smith 2002). Explanations to the problems organizations face include resistance to change and redistribution of power, limited understanding of systems and poor integration of systems (Markus 1987). At the inter-organizational level one well reported obstacle is standards (Damsgaard & Truex 2000). Standards are crucial for supporting interoperability and ultimately necessary for building robust infrastructures.

In this paper we report on a case study of eCustoms implementation in Denmark. The general purpose of the study was to further the understanding of IT’s role in IT-enabled infrastructures. The research was triggered by the discovery that the European move from paper-based administration of customs to digital means of data transfer has created a vacuum of standards. This has led to derogation of an infrastructure built on top a well functioning infrastructure. Contributing to the purpose above, we will in this paper seek to explain how digitalization has generated constraining conditions for an information infrastructure.

Section 2 presents literature dealing with eCustoms and standardization and introduces our analytical tool for the case, the concept of infrastructure. Section 3 presents the applied research approach. Section 4 presents the case of European eCustoms, a move from highly standardized paper-based customs practices to de-standardized eCustoms practices. In Section 5 the case of eCustoms is analyzed in the light of infrastructures. Finally, Section 6 provides some reflection on the case and it delivers our conclusions of the SAD story of eCustoms.
2 THE SIGNIFICANCE OF INFRASTRUCTURES IN THE DOMAIN OF E-CUSTOMS

“International trade is characterized not only by the physical movement of goods across national boundaries but by voluminous paperwork that captures information pertinent to identification, delivery, and government control of transported goods” (Teo, Tan & Wei 1997).

This section will present international trade as consisting of two parallel flows. One is the physical flow of goods and the other is the information flow. We approach the systemic view of these flows as infrastructures. Infrastructures that since long doubtlessly have been bulky and awkward, but still functioned in a harmonized manor.

2.1 The double infrastructures of international trade

Since historic time routes of trade has been utilized to transport salt, glass, gold, spices etc. from one continent to another. Examples include the Silk route to and from the Far East and the Amber road covering Europe from North to South. With technological developments camels and sail-boats have been replaced by planes and cargo-ships which today move goods round the globe basically following the same routes. It is our claim that international trade is an infrastructure, a physical infrastructure. It is “something upon which something else operates” (Star & Ruhleder, 1996) namely the movement of goods. The robustness of the shipping industry infrastructure is well documented (see for example Stopford 1997). It is beyond the scope of this paper to elaborate further on that issue and it is furthermore less relevant in this context where focus is on the administrative procedures related to export rather than on the physical movement of the goods.

It is recognized that the two layers of infrastructure can be difficult to separate. An example is the introduction of a smart seal, the TREC device, which is attached to a container (Baida, Rukanova, Liu, & Tan 2007). The TREC is a container-mounted device which has a mobile receiver tracking the container’s precise location with sensors monitoring environmental parameters in the container, the physical state of the container, and communication modules for exchanging data (e.g., via handheld devices, via satellite, GSM/GPRS or short range wireless). Data which for example can be transferred to customs offices. However, for the purpose of simplicity the two layers are separated and in the remainder of the paper the term infrastructure refers to an administrative customs infrastructure.

On top of the physical infrastructure which secures the movement of goods there is another infrastructure, the administrative infrastructure of customs. Customs deals with the administrative procedures related to import and export. It registers what goes where to who, making sure that tax and excise is paid in the right amount to the right authorities. Customs is regulated thoroughly by national governments and international agencies such as the EU and the World Trade Organization (WTO). The regulation serves as a strengthening of the infrastructure making it more and more fine-grained with respect to security and control. The administrative burden of customs is at large delegated to the businesses and exporting businesses have the responsibility to report the flow of goods but national customs authorities have the right (and obligation) to perform physical control at the site of the exporter.

2.2 Information infrastructures

In IS research infrastructure has often been discussed within the domain of interorganizational information systems (IOS) (see for example Kurnia & Johnston 2000). Most of the literature on IOS appeared in the mid 1990s where there was attention on the benefits of IOS as a tool for improving business performance. By using IOS as point of reference for infrastructure it is taken for given that organizations are able to exchange data electronically via computer based networks using standardized messages (Cash & Konsynski 1985). The Cash and Konsynski article was one of the first academic
works where IOS was defined: “IOS consists of a computer and communication infrastructure that permits sharing of information systems (IS) applications across organizational boundaries” (p.135).

IOS is thus a combination of IS and an interorganizational environment and has as a consequence essentially both technological and organizational characteristics (Johnston & Vitale 1988). The interorganizational environment consists of supply chains, trading partners, standards organizations, industry bodies, transport companies, trade associations, software providers etc. (Kurnia & Johnston 2000). The scope of the systems, the involvement of different organizations with differing goals and the range and nature of possible relationships between the parties involved makes the situation one of extreme complexity (Gregor & Johnston 2001). That is in particular the case in relation to eCustoms which involves public and private organizations and different nations, supranational entities (e.g. EU and WTO), and finally different economic zones (e.g. EU and the US) which have to agree on a standard.

The topic on IT standardization represents a long term issue which has remained unresolved for decades. It has been argued that “Standardization forms a key feature of modernization” (Hanseth, Jacucci, Grisot, & Aanestad 2006). eCustoms is by the EU seen as an modernization project and standardization should accordingly be addressed. One of the well described chapters in that ode relates to EDI, where no standard has been agreed upon so far, and it is highly unlikely that there would be one in the future (Damsgaard & Truex 2000). Although it is recognized that standards and standard-setting agencies are prerequisites for successful implementation of eCustoms, agreeing on international standards and achieving interoperability between different standard-based systems across the EU member states and at an international level remains still a great challenge (Henriksen, Rukanova and Tan 2008). The challenge is among other factors driven by socio-political traits, which influence the collaboration across organizations (Damsgaard & Lytyinen 1998, 2001).

These traits are among others related to the interorganizational and networked aspects which are not centrally controlled by any authority. It involves the linking of organizations which imply that organizational boundaries are lowered leading to increased transparency, the dependency on critical mass and, the dependency of infrastructures to secure efficient and reliable transactions between trading partners (Damsgaard & Lytyinen 1998, 2001). Those aspects have to be considered when studying the drivers and barriers for adoption and diffusion of IOS including eCustoms. It may be technically possible to fulfil the requirements for an IOS outlined by Johnston and Vitale (1988) twenty years ago but the socio-political challenges constraint the building of an administrative infrastructure. This aspect is eminently demonstrated in the case of customs where different stakeholders have different vested interests which are necessarily not compatible with a global standard.

In their study of the tension of flexibility versus standardization of information infrastructure Hanseth, Monteiro and Hatling (1996) discuss the need for shared international standards from a socio-political perspective. They argue that it is possible to have bilateral agreements among a limited number of communication partners. It is however not cost-effective or even possible in the long run if communities are to share an infrastructure based on bilateral agreements. An information infrastructure holds according to Hanseth et al. (1996) seven characteristics: shareable, common, enabling, physical embodiment of an architecture, enduring, scale, and economically sustainable.

Star and Ruhleder (1996) argue that there is a common misconception of infrastructures where they are seen as something which are built and maintained and which sinks into an invisible background but which is still there at hand and completely transparent. In their article they emphasize that an infrastructure is something upon which something else “runs” or “operates” but it is not completely transparent and also often not invisible. Star and Ruhleder (1996) define eight dimensions of infrastructure (see Table 1). Star and Ruhleder (1996) stress that the eight dimensions of infrastructure are without an absolute boundary and furthermore that infrastructure is something which is not stripped of use. The eight dimensions will be applied in the analysis of the eCustoms case.
Table 1 Eight dimensions of infrastructure. Adapted from Star and Ruhleder (1996)

<table>
<thead>
<tr>
<th>Infrastructure dimension</th>
<th>Characteristics of dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embeddedness</td>
<td>Infrastructure is sunk into other structures, social arrangements and technologies</td>
</tr>
<tr>
<td>Transparency</td>
<td>Infrastructure is transparent to use and invisibility supports using it. It does not need to be invented each time it is used</td>
</tr>
<tr>
<td>Reach or scope</td>
<td>Infrastructure has reach beyond a single event. Its reach or scope is either spatial or temporal</td>
</tr>
<tr>
<td>Learned as part of membership</td>
<td>Outsiders of the infrastructure encounter the infrastructure as a target object to be learned about</td>
</tr>
<tr>
<td>Links with conventions of practice</td>
<td>Infrastructure both shapes and is shaped by the conventions of a community of practice</td>
</tr>
<tr>
<td>Embodiment of standards</td>
<td>Infrastructure takes on transparency by plugging into other infrastructures and tools in a standardized fashion</td>
</tr>
<tr>
<td>Built on an installed base</td>
<td>Infrastructures develops based other bases for example do optical fibres run along railways or roads which represent existing bases or infrastructures</td>
</tr>
<tr>
<td>Becomes visible upon breakdown</td>
<td>The normally invisible quality of working infrastructure becomes visible with it breaks.</td>
</tr>
</tbody>
</table>

Our presumption is thus that customs can be analyzed as double infrastructures of which the physical flow of goods is one part and the information flow is the other. The information flow we approach by the eight dimensions of infrastructure by Star and Ruhleder (1996). The methodological section below describes our work of collecting and analyzing data to fulfil this setup.

3 METHOD

The study of eCustoms is based on a longitudinal project on eCustoms implementation throughout Europe involving multiple stakeholders from government, businesses, a standardization body (UN), and academia. The project started January 2006. The project is built around the concept of Living-Labs (LLs). Based on collaboration in LL issues are identified and further explored among the involved partners. The present study reports from the specific experiences gained in the third LL of the project which is situated in Denmark. The Danish LL focuses on export of food and involves a major dairy producer (DairyComp), Danish customs (Customs), a Danish software producer (SoftComp), and finally researchers from national and international institutions.

The significance of the two previous LL in the project should not be underestimated. The previous LLs have provided a thorough understanding of eCustoms especially among the practitioners who have achieved an opportunity to share experiences across countries on a more informal basis than normal. The LLs should therefore be seen a cumulative process where all participants learn from each other and bring in the knowledge from earlier LLs and work practices in general.

The data which is used for this paper has been collected via different means including company visits, 9 workshops involving all partners of the project, 5 general project meetings, 4 specific Food LL meetings with stakeholders involved in the Food LL, email exchange after meetings, study of minutes and other written documentation provided by the food producer and the national customs authorities. Minutes of meetings taken by researches have been distributed to all participants for approval and commenting. During the project extensive written reporting in the form of deliverables has been produced due to the character of the project, which is EU-funded. These sources of data provide a robust material for documenting data which in nature is unstructured and qualitative.

The data collected in the project covered many different theoretical perspectives and areas of interest as the project objective is to find a solution to a real world problem. The idea of double infrastructure as theoretical lens was triggered in workshops with DairyComp and Customs. Following the idea 10 interviews were made with DairyCorp (4), Customs (3), the developer of DairyComp’s eCustoms...
interface, DairyComp’s logistics service provider, and one other national customs organization (Swedish customs) to explore the explanatory potential of the infrastructure concept. Interviews were recorded and documented by notes. The interviews were guided by the dimensions of infrastructure and were conducted as iterative process towards saturation – meaning a search for being able to describe European eCustoms as an information infrastructure.

4 FROM PAPER BASED TO ELECTRONIC CUSTOMS PROCESSES IN EUROPE

The inspiration to this paper is based on a dialogue between an exporter and a customs officer which took place during a project meeting in summer 2008:

Exporter: The dream scenario [of eCustoms] would be a return to the SAD [single administrative document].

Customs officer: Yes, that is the ultimate example of standardization.

The small dialogue illustrates the problems of implementing eCustoms with respect to finding a suitable and universal standard for exchange of data.

4.1 The paper-based customs infrastructure

In the 1980’s, a company that was an active exporter or importer on the European market had to deal with a set of more than 200 different forms to cover different customs procedures, in different countries and for different modes of transport. As a response to the call for a reduction of the administrative burden for economic operators in an all the time more and more globalized arena for businesses the single administrative document (SAD) and related legislation was introduced in the EU. SAD, specified by the EC in Regulation No 1875/2006 and No 648/2005, presented a general form for all types of customs declarations that was applicable to 18 European countries.

The original SAD was introduced January 1st, 1988. Originally 18 countries did employ the document as standard for their customs declaration. In 1993 the SAD disappeared from trade within the EU as a result of the creation of a common internal market. Ten years later, in 2003, the SAD was reintroduced. The plan was to adapt to the evolution that had occurred since its first creation. The required data to submit was reduced by about a quarter, but perhaps more important was the further diffusion of the SAD as a standard for customs declarations. When the reformed SAD and attached legislation was taken into practice by 2007, the format was adopted by more than 30 countries, including for example Russia.

The SAD was originally a set of 8 copies of the same document printed on carbon paper which was handed out to the different authorities during the journey of the goods. All copies had a specific purpose. The first copy was for the country were export formalities where carried out, the second was used for statistics in the export country, the third copy was returned to the exporter, and so on. Some fields where mandatory to collect for all countries, while some fields were optional to collect. The SAD standardized data was submitted to national customs and with one and the same operation a company could do export clearance, issue the transport document (customs), and prepare the clearance in the country of destination.

4.2 Introducing eCustoms

DairyComp is one of Europe’s largest dairy companies with export from several European countries. Major export destinations include the usual suspects Russia, USA, and China but also large customers of milk powder such as Oman and the Dominican Republic. Since a few years, DairyComp submits data needed for customs procedures when exporting from Denmark electronically through the Danish
“e-Export system” which has been massively diffused among Danish exporters (Bjørn-Andersen, Razmerita and Henriksen 2007).

Customs messages can still be sent through a paper based equivalent which it is possible to use for companies who by some reason do not want to or is not able to submit export data electronically. The paper-equivalent is still based on the SAD. The data model is a field by field translation of the SAD-fields that the Danish customs were using prior to the computerization. The e-Export system can be reached through either UN/EDIFACT messages or XML messages. At present there are no international standards for electronic customs declaration messages through XML. Regarding UN/EDIFACT directory D96.B is used. The sent messages are based on a range of UN/EDIFACT documents.

Regarding data meaning, the e-Export system benefits from being a direct derivation from the SAD, since the need for specifying data meaning was highlighted already when the SAD was introduced. A substantial amount of work was already done by UN/CEFACT, ISO, and the EC/TAXAUD to specify the meaning of data. Also for such fairly interpretable fields such as “product description” there are appropriate guidelines on how to provide and interpret data.

According to Regulation (EC) No 648/2005, member countries of the European Union are expected to implement systems that can handle automated risk analysis of exported and imported goods by June 2009. Simultaneously the European customs organizations are implementing or revising eCustoms.

Besides the mandatory fields the SAD presents a number of optional fields that national customs can demand if they want to. In reality, this have in led to that different XML schemas are used for the data transfer related to export declarations. Although not investigated, it is very unlikely that any two European countries would have implemented exactly the same XML schema without purposely collaborating on the matter. Consequently, although not deviating from the stipulated data model the transferred data differs from country to country. Besides the differences in data model, eCustoms also differs in the mode of transfer between company and customs, as well as in timing of implementation. New regulation is implemented at different dates in each country.

The DairyCorp delivers a very illustrative example of the de-standardized processes within a small geographical scope. DairyCorp also has business activities in Sweden an it exports from Sweden too. The two branches of DairyCorp have implemented different versions of eCustoms solutions in order to live up to national standards and regulation. Both the Danish and the Swedish eCustoms solution build on the SAD, but the implementation still differs from each other.

The SAD reform was introduced as a step towards digitalizing customs operations into a process that by the EU is named eCustoms. The European eCustoms project has a span that covers development of customs processes until 2013. In a first phase of the eCustoms project, submissions to national customs were to be possible via electronic data transfer. The outcome of that development is that almost every country that before applied the very same, standardized SAD for customs clearance now have developed individual eCustoms solutions that enforce companies to set up individual export processes for each country where the company is exporting from. In short, the introduction of eCustoms has led to a de-standardization of European export processes. Ultimately leading to a breaking of an infrastructure which had demonstrated its robustness during the paper-based days but which had been lost in the process of digitalization because each and every country has developed its own system based on standards which are not interoperable beyond its borders.

5 FROM PAPER BASED CUSTOMS TO ECUSTOMS: LOSING THE INFRASTRUCTURE IN THE DIGITALIZATION

As our field data suggests while one infrastructure, the administrative customs infrastructure, goes through radical and extensive changes the corresponding, physical trade infrastructure tries to keep on as usual. International trade is steadily increasing due to the general demands in the markets. But as
Table 2  
Eight dimensions of infrastructure applied on the eCustoms infrastructure

<table>
<thead>
<tr>
<th>Infrastructure dimension</th>
<th>Characteristics of dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embeddedness</td>
<td>Infrastructure is sunk into other structures, social arrangements and technologies. The dimension of embeddedness is lost because the SAD is not supported by the technologies adopted by the different exporting countries. Business is going global with international trade networks, delocalized operations but customs organizations are still to some extent working in national silos.</td>
</tr>
<tr>
<td>Transparency</td>
<td>Infrastructure is transparent to use and invisibility supports the use of it. It does not need to be invented each time it is used. Infrastructure differs from country to country. For a company used to export to one European country has to relearn when exporting to another European country. The eCustoms infrastructure of today is not invisible but has to be manipulated and maintained by humans in order to fulfil its task.</td>
</tr>
<tr>
<td>Reach or scope</td>
<td>Infrastructure has reach beyond a single event. Its reach or scope is either spatial or temporal. The eCustoms infrastructure is intended to be a backbone for all trade processes in and out of Europe. By interfacing with exporting companies’ systems, data from internal systems and export data should flow seamlessly and without human intervention.</td>
</tr>
<tr>
<td>Learned as part of membership</td>
<td>Outsiders of the infrastructure encounter the infrastructure as a target object to be learned about. Russian customs adopted SAD but will not adopt eCustoms that match each European country. For transit goods, goods that only passes through the EU on its way to a non-EU destination, the EU has managed to develop one single system. This system and its output is recognized by Russia as valid proof for the control of goods. Thus, a well-functioning eCustoms infrastructure for Europe has potential of being recognized.</td>
</tr>
<tr>
<td>Links with conventions of practice</td>
<td>Infrastructure both shapes and is shaped by the conventions of a community of practice. Customs builds on an extensive tradition of trade processes. In regions (i.e. Scandinavian countries) they resemble to each other. eCustoms will further the tradition. As eCustoms systems are implemented and have to be used by exporting companies they stipulate the conventions of practice.</td>
</tr>
<tr>
<td>Embodiment of standards</td>
<td>Infrastructure takes on transparency by plugging into other infrastructures and tools in a standardized fashion. The different national eCustoms systems have different interfaces to exporting companies. Exporting from all European countries would require 27 different interfaces. The eCustoms systems follows the same process flow but have differences in the data model, the way of transferring data, and the timing of implementation. The consequence is that all systems that basically performs the same task and fulfils the same purpose needs to be treated as completely different systems.</td>
</tr>
<tr>
<td>Built on an installed base</td>
<td>Infrastructures develops based other bases for example do optical fibres run along railways or roads which represent existing bases or infrastructures. eCustoms builds on Internet, XML and web-technologies. Both companies and customs organizations already use Internet for various purposes. In Denmark it is possible to use Internet and the FTP protocol to upload messages for keeping control of what is exported and when.</td>
</tr>
<tr>
<td>Becomes visible upon breakdown</td>
<td>The normally invisible quality of working infrastructure becomes visible with it breaks. The differences have become visible after the migration to eCustoms. DairyComp has to implement different solutions for eCustoms in each country where they export to. When the implementations are working at both DairyComp’s site and at the national customs organization data is available when needed without any human intervention. When any of national eCustoms systems breaks down DairyComp has to put their export processes in that country on hold and fall back on paper-based procedures.</td>
</tr>
</tbody>
</table>
captured by Star and Ruhleder (1996) the double infrastructures are interrelated and a rupture in one of the structures causes problems in the dependent infrastructure. Table 2 presents an overview of eCustoms according to the eight dimensions of infrastructures by Star and Ruhleder.

5.1 Consequence of the constraining infrastructure

eCustoms in the EU is being introduced due to several different reasons. One of them is to reduce the administrative burden for companies who have business throughout Europe (TAXUD, 2007). By harmonizing and digitizing customs processes increased security and control should lead to a decrease in the administrative burden for pan-European companies – not the opposite. In the light of this objective we analyzed the consequences of national differences in regulation implementation.

The consequence of eCustoms implementation for DairyComp can be summarized as follows:
- A need for several export departments – one for each country.
- Different system modules reporting to each customs respectively.
- Additional system development costs of >100,000 Euro per new export country.
- Additional maintenance costs, which normally are substantially larger than the development cost for IT systems.

Today not all national customs require DairyComp to do electronic reporting in their export declarations, therefore DairyComp has chosen not to if not required. The reason is that the development and future maintenance cost of extending the existing system with a module for handling electronic export declarations would far overshadow the efficiency gains. DairyComp investigated the cost of developing such a module and had an offer of about 100,000 Euro. Costs which are not directly serving as a means for improving production process but rather just contributing to the improvements of the infrastructural technologies (Carr 2004).

This represents a paradox in relation to the overall eGovernment discourse where the belief is that implementing of initiatives such as eCustoms is a means if streamlining processes (Andersen 2004). The implementation of eCustoms only on a partly basis leads to reduction in work because the procedures of traditional paper handling has to be maintained along with the new electronic reporting procedures which are not standardized.

5.2 Revisiting the standardizing effect of IT

As mentioned above do Hanseth et al. (2006) state that “Standardization forms a key feature of modernization”. They do however similar to the eCustoms study observe that efforts towards standardization led to de-standardization. They state: “The end effect was the undermining of the order-making and increased disorder …. where efforts and actions taken toward standardization and stability lead to an opposite result.” (Hanseth et al. 2006, p. 567).

The idea of the European eCustoms initiative is to have a harmonized set of export processes in all European countries. Assuming that the outset is successful by 2013 which is the currently targeted date, the level of standardization should have reached the levels of the highly structured procedures driven by the paper-based SAD-form. Figure 1 highlights that the constraining of an existing, well-functioning infrastructure is followed by a transition phase in which the degree of standardization drops significantly. With the perspective of double infrastructures it becomes apparent that any dependent infrastructure suffers from the obstacles created by the move towards the implementation of de-standardizing practices such as eCustoms.
Discussion and Conclusions

Based on longitudinal interaction with different stakeholders in eCustoms we make the following contributions. We introduce the concept of double infrastructure and show its relevance for understanding how the transition to eCustoms affects international trade. Using the infrastructure concept, we have also been able to explain how one specific eCustoms project, the European eCustoms project, disrupts an existing infrastructure and constrains the corresponding physical infrastructure. Based on our findings, we question the well-established assumption that introduction of IT per se leads to higher levels of standardization. Our research suggests the presence of a transition state in which current infrastructures may constrain the corresponding physical infrastructure due to the introduction of IT. The finding extends the current understanding of the interplay between information systems and physical infrastructures in the concept double infrastructures.

Interactions between government and business differ in nature and often have conflicting goals (Thompson, Rust & Rhoda 2005). Businesses want to reduce the administrative burden and government wants to increase control and security. IT has been seen as the panacea to solve this conflict. In the eCustoms case IT can be the means which breaks the vital infrastructure which is prerequisite for lifting the administrative burden thereby solving the burdensome interaction between businesses and government. The administrative infrastructure is broken but exporting businesses are still carrying out the activity. The reason is simple: export of goods is the foundation for many businesses and the interaction with public sector is just seen as a burdensome relationship (Fountain & Osorio-Urzua 2001). Furthermore, the introduction of the concept of eCustoms is just another burden which businesses have to deal with in their interaction with public sector.

Throughout the world governments, national customs organizations, international collaboration organizations and interest organizations are approaching digitization of customs processes. Some parts of the world have come further than others (e.g. Singapore, ASEAN) while others are only just started to approach the idea. Devolving an efficient infrastructure for customs processes is becoming a key challenge for competitiveness of nations and regions. For example, the EU has noted that its food industry is poorly equipped to tackle the increased globalization (Wijnands, Van Der Meulen, & Poppe 2007). The move towards eCustoms is one constituent action in addressing this issue. Thus, the presence of efficient eCustoms infrastructure becomes a democratic principle and a prerequisite for developing countries to compete in a globalized world. As the eCustoms being developed worldwide are based on IS to increase security, traceability and control the absence of such systems will effectively exclude regions without eCustoms from the international trade scene.
With regard to the magnitude of consequences that the move towards eCustoms has for governments and businesses, we encourage more research into the subject matter. One specific contribution that could substantially improve the work with eCustoms (and other similar eGovernment initiatives) is an increased understanding on how effective Regulations are when it comes to IS and where their limitations lay. The current understanding on what can be achieved by regulations is still very limited. Additional perspectives on the transition and its effects are needed. Finally, we also see a need to further the understanding of how the transition can be achieved. Logically applied IS research should be able to play a key role in the transition, but if so there is a need to develop strong relationships with exporters, logistics service providers, customs organizations and other national and international stakeholders. Here the growing interest in IS research as a design science is a prospectus development that we see could be beneficial for improving international trade.

7 REFERENCES


**Contribution of Internet to a Democratic Society**

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0436.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Information, Information society, E-democracy, Social inclusion</td>
</tr>
</tbody>
</table>
CONTRIBUTION OF INTERNET TO A DEMOCRATIC SOCIETY

Abrar Haider, School of Computer and Information Science, University of South Australia, Mawson Lakes Campus, SA 5095, Australia, abrar.haider@unisa.edu.au

Abstract

Equitable information access and freedom of expression are viewed as essential aspects of a truly democratic society, whereby all citizens are kept informed and enlightened. A natural extension of this ideal is the human need and desire to communicate and exchange information with others. With the advent of the internet, and extraordinary growth in information and communication technologies in recent years, more information than ever before is made freely available and easily shared. Information is now available in a plethora of digital formats and can be exchanged across time zones, countries and groups in seconds, and this makes communicating and connecting easier and more efficient. This paper discusses the role of free flow of information through internet in shaping democratic values. This paper discusses the role of internet as a democratic tool that allows significant benefits for society at large in a dynamic global environment. It concludes that internet’s contribution to democracy has not fully matured and its potential to revitalize democracy and outweigh potential for oppression and control is a dynamic and multifaceted issue in the global environment. Internet, however, has a passive rather than active role in a democratic society. It, thus, influences only those who interact with it.

Keywords: Internet, democracy, information, democratic society.
1 INTRODUCTION

The contribution of the internet to a democratic society is a complex question, as the terms are quite subjective. The term ‘democratic society’ is difficult to define. The definition of democracy is not the purpose of this paper, however the concept needs to be explored to determine the level of influence that the internet has had. Additionally, in a world of increasing ‘globalisation’, the term ‘democratic society’ can mean a global society as well as being defined by geo-political borders. The internet has the potential to be a very powerful tool in the evolution of the democratic society, especially in regard to the accessibility of information; however it also has the potential to undermine the very fundamental concepts of what a democratic society is all about. This paper discusses the role of internet in a democracy, and highlights the power of internet as an information enabler. It starts with a discussion of democracy and the role of information in a democratic society. This is followed by a discussion on the influence of internet on democracy and vice versa. The paper concludes with drawing some conclusions from real world.

2 DEMOCRATIC SOCIETY

The term ‘democracy’ has its origins in Greece, meaning “people rule”, but the concept of democracy and what constitutes a democratic society is a rather personal view. A purest would determine that true democracy would mean equality, that all people would have a voice and participate in the governing of a society. Held (2006, p1) describes democracy as “having some form of political equality”. The author continues to argue that “most recent definitions refer to democracy in terms of a liberal democracy, where the term can only be applied to ‘governmental affairs’ as opposed to economic and social spheres, and that the concept of democracy belongs to the nation state” (Held 2006, p311). Tuzzi et al. (2007, p.31) view democracy as a political concept, concerning the collectively binding decisions about the rules and policies of a group, association or society. Such decision-making can be said to be democratic to the extent that it is subject to the controlling influence of all members of the collectivity (eg. Citizens) considered as equals. The starting point of democracy thus, is effectively ‘popular control’ and ‘political equality.’ Tuzzi et al. (2007, p. 36-37) take it further by suggesting several ‘evaluation’ aspects that define democracy,

a. Rule of law, that the decisional output and implementation of the law are supreme, with independent judiciary and fair resolution of lawsuits required;

b. Accountability, that the elected political leaders are obligated to answer for their political decisions when queried by citizens, opponents or constitutional bodies; this implies an informed public that actively participates in the political process;

c. Respect for rights and achievement for freedoms, that implementation of social, political and economic equality is progressive; and

d. Responsiveness; that those that are governed are able to be satisfied that their demands are met through policies, services and distribution; also requires an informed and engaged public/society.

The concept of representative democracy, such as that in Australia, relies on the view that informed citizens make informed decisions to determine those people who are elected to make decisions on their behalf. The role of the internet in assisting with this concept is to provide access to information to inform citizens to make those informed decisions. With the advances in technology, where information can be transmitted quickly, the notion of global democracy has been envisioned. This is supported by eCommerce, where multinational corporations and mass media have helped to provide a
view of a global society. Klein (2001) suggests there are four preconditions of a democracy, whether the democratic society is a global one or determined by the nations. These preconditions are:

a. Membership – who is allowed to vote?

b. Communication capability - a common language, the ability to transmit and receive information

c. A system of interest aggregation - such as political parties and lobby groups

d. A democratic culture - where people have faith in the ‘system’, that although their view may not have won out, they accept the decision as part of the democratic process

Scholars argue about the role of these preconditions in a global democracy, and suggest that democracy can only remain in the nation state (Held, 2006). Klein (2001) argues that scholars may underestimate the impact of the internet. He implies that the internet reduces many barriers to global democracy. The internet can certainly increase the communication capability, can facilitate the voting of the Membership and possibly aid in promoting interest, but the internet in not able to change or impart a ‘democratic culture’. This feature of Klein’s (2001) democratic society is essential for a democracy, as people need to have faith that the ‘system’ works. If one loses faith in the system, and there is a collective loss of faith, it results in an unstable government. This view is supported by Zinnbauer (2001) who suggests that whilst the ‘internet can act as an independent stimulant for political engagement’ it is unlikely to “install a deliberative democratic culture amongst users”. Samoriski (2002, pp.42-45) believes that the internet has several implications for the public, communities and government bodies. Firstly, because democracy can be harmed by the concentration of power toward economic interests, communications channels (such as the internet) can be tools through which citizens can discover shared interests and ‘demand a more representative government’. Secondly, due to the information that is available on it and accessibility improvements, more citizens should be able to stay informed and thus make higher-quality decisions for democratic self-government (or when choosing representatives). Thirdly, anyone is allowed to have a voice and even antagonistic ideas can be disseminated; and fourthly, due to the internet’s anonymity factor, people are more likely to speak their mind and protest about injustices without fear of retaliation or prejudice.

3 ACCESS TO INFORMATION – A FUNDAMENTAL PRINCIPLE

The fundamental principles of democracy are freedom and equality. For citizens to be active participants in a democracy, they need to be able to make informed decisions that impact on society. This requires people to have the freedom to access information and to have freedom of expression (Hamilton & Ole Pors, 2003). Habermas (1989, cited in Hamilton & Ole Pors, 2003 p 411) suggests that with greater access to information comes greater participation in democratic process. Therefore, the internet as a tool for the access and transfer of information, must certainly contribute to a democratic society, if the information itself is released on the internet.

In Australia we have the Freedom of Information (FOI) legislation, a legally enforceable right to access public information. Whilst it varies from state to state, and federally, in South Australia it aims to promote the transparency of the government. In Queensland there has recently been a review into the FOI legislation where recommendations have been for a more proactive approach to releasing of information, and to change the culture from one of secrecy, to one of openness (Field, 2008). The release of the Queensland review has obviously revealed that there are significant improvements that can be made to their FOI legislation, and the role of the internet in accessing that information will be taken into account.
The proliferation of the internet could be compared to the rise in publications in eighteenth century Europe. Travelling libraries originated in England around 1740 (Chambers et. al., 1995), which opened up a range of reading material to the masses. Whilst these travelling libraries contributed to the spread of information and ideas, there were limitations. Many people in rural Europe were illiterate, which parallels the lack of eLiteracy in developing nations and even in Australia. However, where the revolutionaries and intellectuals in Europe used pictures and the spoken word to ‘drum up support’ and spread ideas, people still need a reasonable level of eLiteracy to access the pictures or the spoken word on the internet.

To overcome the levels of eLiteracy and other socio-economic reasons for lack of access to the internet, the Public Library has extended it’s historical role, to provide access to the information using the internet. As Byrne (2003) suggests “the aim of the Public Library was to offer information to all, no matter what your social or economic status, it is “an institution to educate for democratic living”. Libraries are however public institutions and subject to public funding, which means there is the potential for information to be ‘controlled’ by governments by being tied to funding. In addition, Libraries are have the restrictions of budgets which in turn can restrict the amount of information a library can purchase. The internet is a relatively inexpensive ‘virtual library’, still with restrictions (that will be discussed later), but a wealth of information nonetheless. It is up to Librarians to help people to find the information they require – to turn information into knowledge. When it comes to accessing information, Libraries are paramount and as Byrne (2003) says, “a healthy library supports a healthy democracy, and a healthy democracy supports a healthy library”.

4 BARRIERS TO INFORMATION

Byrne (2003) believes that the internet has “reinforced the democracy paradigm” as there is universal access to information. The statement should be changed to there is the potential for universal access to information. There are a number or barriers to information on the internet. Some of these barriers are being overcome by adding more information to the internet (such as scanning older documents); other barriers will not be so easy to overcome. In fact some barriers are being constructed.

As mentioned previously, two important facets of a democracy are freedom of expression and freedom of access to information however as Hamilton and Ole Pors (2003) discuss, these come with limitations, which vary from country to country, and these limitations have increased since September 11, 2001. Unfortunately, as in society, there are ‘undesirable’ groups in cyber space. Where the internet has the potential for rallying or uniting people with common interests, to participate in the democratic process, so too can the internet be used to unite people of questionable morals. The internet provides a perceived ‘safe space’ for some people to gather together to perpetuate their warped viewpoints, such as Holocaust deniers, paedophiles and terrorists (Joint, 2005). The limitation to freedom of access to information is a contentious issue. Where does protection become censorship? Some scholars (Nijboer, 2004) argue that greater regulation of the internet is not necessary, this is despite the substantial increase in reports of internet violations (hackers, virus’ etc) since 2000 (CERT, 2004 cited in Nijboer, 2004). Concepts such as anonymity on the internet, another core feature of freedom of expression, are no longer preserved in many countries. In Italy, journalists have to register to publish work on the internet, and the Patriot Act in the United States of America, allows for the monitoring of individuals’ internet usage (Nijboer, 2004). These barriers are continuing to be built, obviously having an impact on democracy.
With the improvements in technology, there are also more sophisticated forms of crime and terrorism. With the culture of fear that seems apparent since September 11, does it create a more insular society? Erikkson (1999) outlines the “Network Revolution” and describes a shift in the method of warfare from Radio, mechanised forces and weapons of mass destruction to cyber weapons, precision munitions and high performance special operations. He goes on to write that infrastructure is no longer the vulnerable target, but knowledge and information assets are. As information is more of a commodity, therefore it is harder to acquire, which conflicts with the fundamentals of democracy. Therefore the internet, in this case, has become a barrier to a democratic society.

Censorship and monitoring of Information on the internet is not only conducted by the Government, many major websites are still controlled by rich multinationals or media moguls. In particular, the major ‘traditional’ media has a presence in cyber space, whether it is the electronic versions of a newspaper, a website of their television station or a presence that is not so obvious. In addition to these major websites there is the associated advertising of products or sites. With advertising, people are drawn to specific websites, and when people are drawn to specific websites, there is an element of control. McChesney (cited in Skinner et al. 2005) believes that the way the media is controlled, funded and structured should be part of democratic debate. He argues that the mass media undermines democracy. He also argues that the policies in the USA do not support democratic principles, but instead make it easier for mass media corporations to strike deals and alliances, thus controlling, to a certain extent, the information we, as consumers, receive. Major search engines also support this structure of corporate control where major companies can pay for priority retrieval, so their website is listed first (Novek, 2000). This view is contradicted by Castells (1998, cited in Kellerman, 2000) who suggests that whilst many cultural battles (the battle of the power to impose behaviour) are fought in the media, the media are not the power holders. He argues that the power lies in the networks of information exchange and the use of symbols to relate cultural movements through icons. This view seems to be a little old and does not take into account the level of sophistication of the mass popular media and their ability to combine radio, television and the internet to reinforce their views of the world.

Other barriers to the internet include the socio-economic inequities (Zinnbauer, 2001). Geldof (2005) discusses the difficulties of using ICT’s in developing countries, such as African nations. Politically, the nation needs to be stable, with a clear ICT policy, where the use of ICT’s is linked to the reduction in poverty. There are many politically unstable countries in Africa so if Geldof’s view is correct then many African nations will not have access to the internet for many years. Scholars (Eriksson, 1999) recognise that some Governments/Leaders identify the internet as a threat to their regime, and therefore resist the influence of the internet. Additionally, in many of these developing countries, there needs to be a focus on developing the infrastructure to utilise the internet and provide access for the people. Access is not enough, there needs to be education to develop the skills so people can use the internet. Things we in the West take for granted such as basic literacy. After all it is informed citizens, making informed decisions that are at the cornerstone of a democratic society.

5 OTHER USES OF THE INTERNET IN A DEMOCRATIC SOCIETY

It can be argued that the internet’s biggest contribution to a democratic society is the acquisition of information. There are also other applications that the internet has been used for. Some countries have tried to adapt the internet to the political sphere. In Germany, the parliament has a website where different issues are discussed, facilitated by politicians and in Denmark there were online public
debates in 1999 (Ferdinand, 2000). The internet has made it quicker and easier to contact elected representatives. Once politicians were contactable by ‘snail mail’, now they have email, blogs and participate in online forums. Some politicians have their own website, and at the last Federal election in Australia, Kevin Rudd (current prime minister of Australia) had his own Facebook page. Hansard (traditional name for the printed transcripts of parliamentary debates in the Westminster system of government) is also available online and through ‘youtube’, you can watch policy speeches on demand. As a communication tool, rather than a research tool, the internet has the potential to revolutionise the way we communicate with politicians.

Another aspect of the internet’s contribution to democratic society is the ease in which people can start their own businesses. As one of the features of a democracy is equality, the internet makes it more affordable to start a business. This ‘economic democracy’ has provided many people with the option of beginning their own business, without having to incur the costs associated with leasing a shopfront. Artists, in particular musicians, have found the internet to be a ‘double edged sword’. Some musicians have used their democratic right to free speech and enjoyed the benefit of an economic democracy to launch their careers on the internet. The internet has made information easy to access, easy to download, copy and manipulate (Oddie, 1999). This concept is open to abuse, with many illegal downloading sites or file sharing sites (Napster used to be one, Lime wire is another). Like any right, there is a responsibility, and the responsibility in this case is to use the information we access from the internet in the correct way.

The biggest influence of the internet has been on economic freedoms rather than political; with marketing, procurement and recruiting completed over the internet. Internal efficiency and communications have been boosted since its advent. Although not a democratic ideal that we are looking at specifically, economic freedom to sell, buy and market to a worldwide audience (thus earn an income) has been quite revolutionary (globalisation etc). There has been some impact on politics and subsequent democratic improvement, according to Ferdinand (2000, pp.5-8):

a. Some see the internet (communications and information available through it) as a way for increased transparency of government administration (to outside observers) and thus governments will be more ‘amenable to democratic pressures’;

b. Professional (political) party organizers (particularly in the US) have quickly utilised the internet in campaigns. Not only is the internet less expensive to campaign on (versus large TV campaigns), but it can be used to reach particular demographics and even those that would usually be disenchanted with politics (thus increasing the number of citizens who are interested and participating in the democratic process of elections); and

c. Many national parliaments are using internet to educate citizens about their activities and legislative procedures; who have at least broken down barriers between the political and public spheres.

Tuzzi et al. (2007, p.32) put forward their opinion that evolving information and communication technologies are becoming increasingly relevant in political systems. They also believe that there is a trend of public disengagement with traditional political life; with expressions of demand placed on direct civil engagement (from politicians) in such processes. Howard Dean (2007, p.94), chairman of the Democratic National Committee (ruling body of major US political party the Democrats), believes that the ‘internet is the most significant tool for building democracy since the invention of the printing press. People are now easily able to create, discover and connect with networks [of people] within hours, anywhere around the globe.’ He goes on to reflect over the resultant demands on leaders in terms of two-way communication and authenticity placed by diverse communities of university students, ethnic groups etc with shared interests. Dean (2007, pp.95-98) also believes that the internet:
a. Allows online communities to question political candidates (that may one day represent them) and through sheer people power, change the course of campaign outcomes;
b. Can be used by both party insiders and outsiders to distribute critical video or audio excerpts of slip-ups from opponents, thus gaining momentum;
c. Allows campaign professionals to manage events and support, address concerns of the public and conduct fundraising from large numbers of people;
d. Allows parties to develop social-networking software and tailor it to political activism, getting communities directly involved in campaigns and debate; and
e. Provides pro-democracy citizens of authoritarian nations like Iran or China with the means of evading government interference in exchanging ideas.

6 DOES DEMOCRACY CONTRIBUTE TO INTERNET OR THE INTERNET CONTRIBUTES TO DEMOCRACY

Gutstein (cited in Hrynysyn, 2001) states that a “democratic society rests on a bedrock of freely exchanged public information”. The internet has the potential to contribute to this freely exchanged public information, by providing access to information. Whist the internet contributes to democracy, it is our democratic society that provides the information in the first place. Zinnbauer (2001) takes the viewpoint that the internet can be a tool in the democratisation of a civil society, but it is up to governments to promote democracy through their own policies in relation to Freedom of Information, freedom of speech. In more controlled countries, such as China, there is still access to the internet; however the censorship is a lot greater. This censorship does not prevent the transmission of ideas or information, but it certainly makes it more difficult. So it is the lack of democracy that impacts on the contribution of the internet.

Internet access and its contribution to democratic society has the greatest potential in countries that have the policies to promote access to information. However democracy is facing a different challenge. Ferdinand (2000) discusses the increase in political apathy in the West. This is illustrated by the decrease in the membership of political parties, a decrease in voter turnout in countries that do not have compulsory voting and there is an increase in the public’s dissatisfaction in Parliament. The only exception to this seems to be single issue causes. Single issue causes seem to be enhanced by the internet through Social Networking sites.

The impact of social networking sites, such as Facebook, may open people’s eyes and minds to the political issues of a country at a particular time, and do it in a swift manner. One only has to look at the range of “groups” that occur on Facebook to see the power of conscious raising that can go on. Whether people act on their views though, is another matter. There are 209735 people that are members of the group “F$@! Off Japan, leave the whales alone” (Facebook, 2008 Appendix 1), but what else are these people doing to show their objection, and their active participation? Are they boycotting Japanese goods? Are they writing letters to the newspaper? Are they donating to GreenPeace? Even without these actions, these group members are, engaging in debate, increasing their knowledge, and therefore becoming more informed citizens.

The internet seems to be facilitated by a democratic society, and a democratic society seems to facilitate the internet. When societies provide policies that promote democracy, individuals can freely use the internet for their own purposes, such as promoting their own cultures and views, where these
views can be explored in a safe, not threatening manner (Eriksson, 1999). In countries that do not promote democracy, then this feature of the internet can be perceived as a threat, which leads to censorship, where individuals cannot discuss differing viewpoints (that contradict the view of the government) and an underground culture can develop.

7 CONCLUSION

The improvements in technology have made it easier to access the internet, made the internet faster, and has allowed for more complex information to be placed on the internet. Once a research tool, the internet can now launch careers by providing an outlet to perform music, it can conduct business in virtual shops, it connects people through social networking sites amongst many other features. It facilitates the flow of information - but only to people that seek it. These factors illustrate the contribution of the internet to our (western) democratic society. In terms of how the internet has impacted on democracy, the internet is simply a technology, which is neither good nor bad, but it depends on how it is used (Joint, 2005). In countries that already are a democracy, the internet is another tool to seek information. In countries that have a more controlling form of government, the internet is also controlled, undermining the very principles of democracy. The key word should be potential – the internet has the potential to contribute very positively to the notion of a democratic society, but it is not yet used to its full potential.

References


Kellerman, A. (2000). Inform.….phases in the rise of the information society. Info, 2(6), pp 537-541


Samoriski, J. (2002). Issues in Cyberspace. Allyn & Bacon, Boston USA


APPENDIX 1

![Facebook Group](http://www.facebook.com/group.php?gid=7251222651)

Barriers for Transformation: Impediments for Transforming the Public Sector through e-Government

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0399.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-government, Organisational Change, Governance, Service innovation</td>
</tr>
</tbody>
</table>
BARRIERS FOR TRANSFORMATION: IMPEDIMENTS FOR TRANSFORMING THE PUBLIC SECTOR THROUGH E-GOVERNMENT

Van Veenstra, Anne Fleur, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, the Netherlands, a.f.e.vanveenstra@tudelft.nl

Klievink, Bram, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, the Netherlands, a.j.klievink@tudelft.nl

Janssen, Marijn, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, the Netherlands, m.f.w.h.a.janssen@tudelft.nl

Abstract

In the past decades many e-government initiatives have been introduced. Currently, attention is shifting towards enhancing the effectiveness of these initiatives by improving service delivery through organizational transformation. This is called Transformational Government and is often regarded as the next step in the development of electronic government, in which unique cross-organizational processes for customer service request are created. However, the creation of such service processes is cumbersome and even those organizations that have realized successful online service delivery are often hindered in their efforts to achieve transformational aspects such as service orientation and a networked organizational structure. As there is little insight in the barriers that organizations encounter on their path to Transformational Government, this research investigates those barriers by looking at two case studies that have made efforts to transform. The barriers we found include a lack of knowledge about necessary changes, a lack of change in the organization structure and the absence of a transformational mindset.

Keywords: Transformational Government, e-government, barriers, transformation, service delivery.
1 INTRODUCTION

Over the last decade, a first generation of e-government initiatives has been implemented. Many of these projects focused on realizing electronic service delivery to citizens and businesses. Recently, more sophisticated e-government initiatives have been set up that are expected to create results of a more invasive nature: “E-government is being increasingly seen as an enabler for a longer-term transformation of government that goes far beyond online service delivery,” the OECD (2005, p 164) notes. ICT has the potential to improve the performance of government organizations as they become part of networks, but organizations need to understand this potential (OECD e-Government Studies 2005, p 164). In order to achieve these objectives of transformation, a new effectiveness enhancing mindset is deemed necessary: Transformational Government (Irani & Elliman & Jackson 2007).

Governments acknowledge that major changes are needed to achieve the impact on organizations and service delivery that Information and Communication Technology (ICT) within government promised to deliver. Considering the complex, evasive and long-term nature of transformation, much research focuses on stages of growth models that both describe and provide guidance for the development of e-government initiatives (e.g. Andersen & Henriksen 2006; Layne & Lee 2001; Chen 2003; Gupta & Jana 2003; West 2004; U.N. 2002). The idea behind these maturity models is that transformation is expected to happen stage-wise, starting off from developing an information system and, eventually leading to transformed government. However, although a lot of e-government systems have been set-up in a relatively short time-scale (Irani et al. 2007), the “e-government revolution” (West 2004, p 15) is – in reality – falling short of its potential to transform government service delivery (West 2004).

Research in the field of e-government has identified challenges and barriers to its implementation at government organizations (Liu & Hwang 2003; Gil-Garcia & Pardo 2005; Janssen & Cresswell 2005; Coursey & Norris 2008). We argue that literature on Transformational Government can also benefit from identifying barriers impeding the development of e-government towards transformation. Barriers are those factors that can be identified as blocking the adoption of a transformational mindset and, consequently, the goal of achieving a transformed public administration. This paper aims to contribute to research carried out on Transformational Government by investigating why government organizations fail to achieve transformation. This is done by identifying and classifying the major barriers that occur when organizations aim for transformation. The main research question of this paper is, therefore: What are the barriers impeding transformation of government organizations?

This paper is based on two case studies on organizations in the Dutch public sector that have made initial efforts to transform through the adoption of e-government, but encounter barriers to progressing further. This paper is structured as follows. First, a background of Transformational Government is presented in order to provide insight in steps undertaken by organizations aiming for transformation. Second, a framework for barrier classification is derived from literature. Then, the case studies are investigated and the barriers we identify are categorized. After this analysis, the research findings are presented, followed by conclusions.

2 TRANSFORMATIONAL GOVERNMENT

In recent decades, government organizations on all levels have implemented e-government initiatives, such as realizing electronic front offices, creating authentication concepts, and setting up digital vital registries. As a result of legal obligations and widely available information technology, many government organizations have increased their online service delivery to citizens and decreased administrative costs for businesses. Both enhanced efficiency and effectiveness of government services are key objectives in these developments (Irani et al. 2007). Furthermore, some organizations have strived for organizational change to support these initiatives better. However, large-scale effects have not been achieved yet and little transformation of the public sector can be observed. In order for society as a whole to benefit from e-government initiatives, government organizations need to make
sure that they transform in order to realize demand-driven service delivery that truly meets citizens’ and business’ needs.

2.1 Characteristics of transformed government organizations

Transformation can be defined as a “complete change in character, condition” (West 2004, p 15), which refers to the invasiveness of the process. Transformational Government is the mindset that aims for political and organizational transformation of the entire public sector (West 2004; OECD e-Government Studies 2005). This transformation is spurred by the implementation of a large number of e-government initiatives that promise to change the way the public sector functions (Kim & Pan & Pan 2007), for example by encouraging cooperation between public officials and government organizations and the development of cross-agency ‘portals’ (West 2004, p 16). Irani & Elliman & Jackson (2007) extend this idea and define the rationale for Transformational Government as “the exploitation of e-government such that benefits can be realized” (Irani et al. 2007, p 327). From a business perspective, Transformational Government can thus be seen as the value added of e-government initiatives for citizens and businesses.

The realization of such value added requires transformation in multiple directions. Public administration is foremost expected to become more customer-oriented and act more pro-actively towards citizens and businesses in order to deliver these benefits (Peristeras & Tarabanis 2000). Government will then enable “fully integrated and fully executable online services”, as well as “options for website personalization […] and push technology” (West 2004, p 17). Examples include citizens and businesses only providing their information once to any government agency involved in a service, and single contact points in the form of designated websites that function as the unique information- and service access points for all government agencies. Realizing these initiatives requires change to occur not only in the service delivering front office, but also in the back office of organizations (Beynon-Davies 2007). This could lead to cooperation between multiple autonomous organizations and the transformation from a siloed structure to performing tasks as part of chains or networks (Castells 2000). This, in turn, requires that tasks should be assigned to the organization that is best-equipped to carry it out in order to achieve optimization of these processes. Shared service centers can then be formed, in which services from multiple organizations are concentrated in one joint centre (Janssen & Joha 2006). Following these elements, the transformation of public agencies includes some or all of the following dimensions: adoption of service orientation, formation of service chains, business process re-engineering of the back office, integration with other government agencies, formation of shared services, networked or modular organizational structure, and organizational and governance support for transformation.

A major challenge to achieve any of these dimensions of transformational is the structure of the public sector. The landscape of public administration is largely fragmented, as many different agencies exist that are responsible for their own specific tasks and that have a relatively large degree of autonomy. Therefore, transforming the processes of these organizations to form a service delivering chain is a difficult process. Although the introduction of e-government holds great promises for changing the public sector, many authors, therefore, claim that in reality very little of this transformation can be observed (Fountain 2001; West 2004; Coursey & Norris 2008). Furthermore, the practice of transformation is not considered to be one of radical re-design, but rather a process of ‘muddling through’ (Lindblom 1959), in which small, incremental steps are being made that might, in time, lead to change (Wildavsky 1984). These notions run squarely into the ideas promoted by Transformational Government as they claim that there are severe limitations to the capacity of the public sector to transform.

2.2 Classification of barriers to transformation

Research on the implementation of information systems in (government) organizations more often focuses on success factors (e.g. Akkermans & van Helden 2002; Somers & Nelson 2001; Poon &
Wagner 2001; Rosacker & Olson 2008) than on impeding factors or barriers (Liu & Hwang 2003; Gil-Garcia & Pardo 2005; Janssen & Cresswell 2005; Coursey & Norris 2008). Gil-Garcia & Pardo (2005) propose a comprehensive framework for identifying success strategies and challenges for e-government, comprising five categories: (1) Information and data, comprising success strategies such as having an overall plan, continual feedback, and training, (2) Information technology: ease of use, usefulness, and demonstrations and prototypes, (3) Organizational and managerial: project team skills and expertise, leadership, clear and realistic goals, identification of relevant stakeholders, planning, and communication, (4) Legal and regulatory: information technology policies and standards, and (5) Institutional and environmental: leadership, legislative support, and outsourcing and public-private partnerships.

In the field of business process re-engineering (BPR), Al-Mashari & Zairi (1999) identify barriers instead of success factors. Their approach enables us to identify barriers in the field of e-government as well. This fits the field of Transformational Government as there are little cases available in which transformation is an outright success. In the field of public sector change, Janssen & Cresswell (2005) identify barriers to re-engineer the public sector. In order to identify the barriers that government organizations encounter on their path to transformation, this research uses four of the abovementioned categories by Gil-Garcia & Pardo (2005), namely: information technology, organizational and managerial, legal and regulatory, and institutional and environmental. Following Liu & Hwang (2003), however, the two latter categories are grouped under the label of governance. Furthermore, following Hammer & Champy (1993), Al-Mashari & Zairi (1999) and Kim & Ramkaran (2004) we add the category of business processes to our framework. Accordingly, in the case studies that are analyzed in this paper, the factors are found in the following areas: technical, representing (partial) failure at the level of information systems; business process, which includes failure in re-engineering processes or the formation of service delivery chains; organizational and managerial, representing factors at the organizational level, and including cultural failure factors; and the governance category representing institutional factors impeding transformation of public administration.

![Figure 1: framework of categories of barriers impeding the transformation in public administration](image)

## 3 CASE STUDIES

### 3.1 Research problem, methodology and case selection

The focus on barriers impeding the transformation in public administration, instead of the more common success factors, is deliberate. A focus on success factors neglects factors blocking progress. Since there is a lack of cases that portray full transformation, it is not only necessary to know what factors contributed to the success of their partial transformation, but also the barriers that organizations encounter when trying to achieve a stage of transformed government.
Since the field of Transformational Government is still underdeveloped, there is a general lack of theory and research approaches. To research the main question of this paper, we use an interpretivist methodology for in-depth research of organizational case studies since this best fits the complexity of the matter (e.g. Klein & Myers, 1999). In the previous section, the concept of Transformational Government was defined and a number of efforts necessary for transformation are identified. Implementation of some or all of these efforts represents the theoretical ideal of a transformed government in which the effectiveness of electronic service delivery increases. The mismatch between this theoretical ideal and the current empirical state of transformation is used to identify barriers. For analyzing the cases, we use a triangulation of methods (Mingers, 2001). We base our case study research on a combination of research methods, including document analysis and interviews. Insight is needed in the complex organizational context in order to identify the relevant barriers. Single case studies are often seen as inferior to multiple case studies with respect to generalizability (Yin, 1989). Therefore we opted for two cases that are distinct in their size and transformation approach, but share that they are both seen as successful in improving service delivery. A limited number of case studies can be very successful in terms of theory formulation and theory testing (Yin, 1989). Both case studies provide complementary insights in the barriers that impede further improvement in providing the value added that goes with Transformational Government. This case selection method illustrates the goal of this paper, which is not to provide a full growth path to transformation, but to identify the barriers that organizations encounter on their path to transformation.

The cases were analyzed by comparing the vision to the accomplished transformation and determining the factors impeding the development towards the desired situation. The first case study was investigated using seven interviews with management, administrative staff, IT staff and IT architects. The interview results were compared with documentation including press releases and internal documents such as memos, reports, and presentations. This resulted in a list of factors distributed over the four categories described in figure 1. The second case study was investigated in a similar manner, with multiple (semi structured) interviews with people from various disciplines and document analysis being the second method of research.

3.2 Case studies

In order to support government organizations in their efforts to transform, several support programs have been set up in the Netherlands. Benchmarking studies are carried out and best practices are disseminated. Furthermore, local government agencies can receive support from a professional funded by the central government who gives advice tailored to the specific needs of the organization and culture.

During the last decade, the Dutch government actively stimulated the adoption of e-government initiatives, for example by setting a target that 65% of all government service delivery should be available to citizens and businesses online. In 2007, it was concluded that this goal was achieved by all agencies and the directive was changed into the aim that all government websites be awarded a grade seven out of ten. In order to define how this aim should be achieved, a national action program for better service delivery was set up. The main goals of this government-wide action plan are to identify generic e-government components with a focus on architecture and to increase the visibility of the effects on public administration functioning.

The program aims for a comprehensive outlook on electronic government as it not only defines what public electronic service delivery should look like, through defining service delivery criteria and vital infrastructure components, but it also identifies how public administration could enhance its effectiveness. Examples are given of networks of organizations performing services. Although transformation is, thus, an issue of concern for the public sector as a whole, in this study we look primarily at individual government agencies and at transformation on the organizational level.

Within this context, two Dutch cases are investigated. The first case represents a case of a transformed organization, while the other represents transformed service delivery. First, we will discuss a local
government that transformed into a service-oriented organization. This is a front-runner municipality that deployed service architecture. Second, we look at a semi-autonomous public agency that can be considered a good practice of multichannel service delivery because of their radical approach to both organizational structure and electronic service delivery.

3.2.1 Case study 1: organizational transformation in a front-runner municipality

The municipality under study is often viewed as a good practice in the Netherlands for its effort to transform into a demand-driven organization. Early 2000, this municipality’s city council decided that it wanted to become a front-runner in the field of e-government. A four-year action plan was developed which resulted in a high ranking on a government website comparison list. This municipality only succeeded in partial transformation as it transformed its organization, put in place the right technology, but failed to align it with its business processes. The main drivers of the transformation are a focus on service orientation which was implemented by introducing an integration broker, intra-organizational transformation and an ICT-driven organizational re-engineering.

During the implementation of this (first) action plan, the municipality experienced problems in connecting their applications and more and more money was spent on the control and maintenance activities. In addition, they found that the organization was not suitable for demand-driven service provisioning yet. In this project, the organizational structures were not altered and technology actually reinforced the existing structure. To tackle these problems, a new action plan was developed which primarily focused on shaping the organizational conditions and the technology needed for demand-driven service delivery.

The second action plan started with transforming the organization from a siloed structure to a service-oriented organization. Banking and insurance companies were used to serve as an example for this transformation. This was an invasive process which was managed separately from the implementation of ICT and resulted in separate departments for interaction with citizens and businesses. Furthermore, a number of legacy systems was opened up by adding web services. Thereafter, these systems were connected to the website to enable real-time service delivery. After almost two years, it was found that the idea of leveraging legacy systems by opening them up using web services was fruitful, even though it requires standardization and a sound architecture for integrating the various parts. The municipality bought an integration broker architecture that should contain the logic of connecting the website to the applications in the back office.

Later on, the municipality became aware that just having a broker was not enough. It was unclear which process steps should be performed in the front office or back office and what the broker should do. This requires extensive analysis of business processes, a comparison of similarities between these processes and the allocation of processes to the front and back office. In addition, this helped to map process steps to the new organizational structure, clarifying responsibilities and improving processes.

The municipality in this case study transformed its organization, acquired the right technology, but failed to align these systems with its business processes, which proved to be the bottleneck for transformation. The lack of knowledge on how these changes impacted the organization, its processes and the technology, can thus be seen as the main barrier for transformation. Currently, a number of business processes is analyzed and mapped to the broker architecture and website. This has resulted in the creation of several services that can be requested in an integral way and reuse of information stored in the back office. The automation of all business processes is a long-lasting process, as it requires extensive analysis of current processes, its corresponding structures and future possibilities. Furthermore, connections with other organizations are limited; only connections based on well-standardized formats are realized.

3.2.2 Case study 2: front office transformation in an executive agency

The second case studied was a large semi-autonomous executive agency, which is seen as a good practice of multiple-channel service delivery in the Netherlands. The organization is responsible for
millions of interactions with citizens each year. The agency addressed its fragmentation problem by concentrating the coordination of its multiple service channels (and the corresponding departments), into a single department. Although this structure proved very successful for managing the various streams of customer contacts, it did not succeed to transform the organization as a whole. Creating a department specialized in multi-channel management ensured that a state-of-the-art front end was set up, but at the back end, processes and information systems could not keep up with the service provisioning. In other words, service delivery chains were not formed.

The service strategy deployed by the multi-channel department focuses on a number of drivers: efficiency, image, customer satisfaction, and effectiveness. The strategy differs per channel and per customer segment, and for this purpose different service concepts were implemented. For example, customers with complex problems or language difficulties are served at the counter whereas the young and educated are stimulated to use a web-based self-service channel. The organization transformed to a multi-channel service organization with successes both on efficiency and effectiveness. Higher customer satisfaction and lower costs mark the success of this transformation of the front office, but the formation of service chains or networks with other organizations remains a challenge.

The formation of chains was hindered by the legacy in the back office. A main barrier to transformation is, thus, that the back office lags behind the front end redesign. Furthermore, in some service delivery chains, improved cooperation with other organizations would have been desirable, including more interoperability of information systems and a clearer allocation of responsibilities. Moreover, most developments in this case were introduced on an ad-hoc basis, with trial-and-error as the main strategy. The result is an uneven development of a single coordinating department, with back office departments lagging behind.

A second barrier impeding transformation is the absence of a vision on how this agency could enhance its service delivery by becoming part of a network of organizations. Following the trial-and-error approach, its transformation efforts were based on current structures and the desire to improve efficiency and customer satisfaction. Therefore, the transformation efforts reinforced the existing role of the organization. Alternative strategies, for example one that would redefine the organization’s function from a service providing agency to a broker in a service network, were not considered since there was no vision on what a transformed organization should look like or what’s its role in the network could be.

### 3.3 Findings: barriers identified

The two case studies in this paper represent organizations that have transformed partially. Therefore, the transformational efforts observed in the cases are summarized in table 1 in order to demonstrate the barriers that can be encountered when transformational efforts are undertaken. The barriers derived from the two cases are categorized, reflecting the factors that impede the studied organizations’ efforts to transform.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Transformational efforts</th>
<th>Barriers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governance</td>
<td>National action plan for enhanced service delivery (cases 1 and 2)</td>
<td>Misunderstanding the impact on long term control and maintenance efforts (case 1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absence of implementation guidance and best practices (case 1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Insufficient understanding of clients (case 2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of transformational readiness (case 2)</td>
</tr>
<tr>
<td>Organizational and managerial</td>
<td>Demand-driven customer orientation (cases 1 and 2)</td>
<td>Lack of an organization-wide overview (cases 1 and 2)</td>
</tr>
<tr>
<td></td>
<td>Change from a siloed structure to a process-oriented organization (case 1)</td>
<td>Detached front and back office (cases 1 and 2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gap between IT and administrative departments (cases 1 and 2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unclear responsibilities for process ownership (cases 1 and 2)</td>
</tr>
</tbody>
</table>
Table 1:  barriers identified in the case studies

Table 1 shows that barriers were observed at all levels. Furthermore, the common characteristic of the barriers observed seem to be (1) a lack of knowledge of – or mindset for – transformation, despite the presence of a national action plan for improving (online) service delivery, (2) lack of knowledge of the changes necessary for transformation and its implications, and (3) a lack of organizational change that needs to accompany transformational efforts. Possibly, if additional efforts of transformation are attempted, other barriers might be observed.

4       DISCUSSION: BARRIERS IMPEDING TRANSFORMATION

Since we have looked at only two cases studies, the barriers we identified are not generalizable to all government organizations in the process of transformation. However, by selecting two good practices with a different approach to transformation, the barriers found here are expected to apply to most (Dutch) government organizations. If the elements of Transformational Government found in literature are not even realized in these cases, other organizations are expected to have trouble realizing them as well. Therefore, the theoretical underpinnings and the main barriers identified in the cases by use of the framework can help organizations anticipate on these barriers early on in their transformation process. Although not generalizable, this research on barriers for Transformational Government provides valuable insights for both the theoretical advance of the research field and for the challenges faced by practitioners.

The greatest transformational effort undertaken by the organizations under study is the adoption of service orientation, which was identified as one of the dimensions of Transformational Government. In both cases, this was partly realized. However, service orientation was introduced mainly by technical means (such as broker technology) and front end innovations (such as a multi-channel strategy), instead of undertaking a thorough redesign of the organization and its business processes. Re-engineering the processes in the back office could have led to the formation of service provisioning chains throughout the organization. Therefore, while these organizations acknowledge that their service levels could be improved, they fail to implement all organizational changes necessary, such as assigning responsibilities, aligning processes and bridging gaps between front and back office, IT- and administrative departments. Hence, they appear to lack the understanding of how organizational transformation impacts the re-design of their business processes.

Accordingly, not any effort was observed to reposition the organization within the wider context of the public sector by enhancing service delivery through the formation of chains with other organizations.
Often, processes run through multiple departments or organizations. To accommodate this, service chains must be formed, business processes must be re-engineered where necessary, and systems must be able to interact with those at other government agencies. For the municipality under study, this is further complicated by their dependency on third party software vendors. Despite the presence of a national action plan, the organizations under study seem to lack an overarching transformational mindset to connect all those dimensions. Therefore, they have taken small steps on various paths without a clear vision on where they are heading.

In accordance with the theories that look at transformation as being not just a technical issue, we found barriers also at the governance, organizational and process levels. This leads us to believe that Transformational Government is not just about adopting the right technologies, acknowledging client’s wishes or agreeing on standards, but that it is about the fundamental re-orientation of the role and function of an organization as part of a network that offers services to citizens and businesses. Furthermore, it requires suitable governance mechanisms supporting transformed organizations and business processes need to be installed. Focusing on one part only – whether improving systems or setting up a service oriented front office – inhibits the comprehensive organizational and process reform that is needed for transformation. Often, like in the case studies, organizations take too narrow a focus on transformation and neglect the invasive nature of the process on all of the categories we identified.

For government organizations to advance on their path to transformation, we found barriers that need to be overcome, most notably a lack of mindset, lack of process alignment, and a lack of knowledge on implications of the introduction of e-government facilities. Taking these into account requires the public sector as a whole to determine and define where it is heading and how it will attempt to get there; an integral vision on Transformational Government is necessary. As long as this transformational mindset is lacking and knowledge on these changes is still limited, Transformational Government, as foreseen by scholars in many stages-of-growth models, is not likely to be within reach.

5 CONCLUSIONS

This research aimed to contribute to the research field of Transformational Government. This research field studies the transformation effects of e-government supporting enhanced use of ICT for more effective service delivery. In this paper barriers impeding transformation in the Netherlands were identified and categorized. We believe that identifying these barriers complements the present focus on success factors. Using literature review, a classification of barriers was developed consisting of governance, organization, business process and technology. This classification was used to analyze two cases in-depth. Barriers observed in both cases are related to a lack of knowledge on transformation leading to very limited process re-engineering with a front office focus and a failure of using ICT to enhance the effectiveness of e-government initiatives. Furthermore, an urgent lack of a transformational mindset was observed. The core of transformation is in developing a transformational mindset and a vision to guide organizational and service process development, as well as a collaborative culture. Further research should, thus, focus on improving theoretical and practical support for Transformational Government, such as refining current stages-of-growth models to reflect the complex nature of transformation and the difficulties in achieving a stage of transformed government.

References


Proceedings ECIS 2009
Public Administration and Finance.
Public Administration Review 64 (1), 15-27.
California.
UNPACKING MULTICULTURALISM IN THE ICT WORKPLACE: DIFFERENCES IN RESPONSES TO WORKPLACE SITUATIONS FOR ENGLISH AND NON-ENGLISH SPEAKING BACKGROUNDS.

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0066.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Cross-cultural issues, Knowledge heterogeneity, Knowledge Worker Performance, knowledge transfer</td>
</tr>
</tbody>
</table>
UNPACKING MULTICULTURALISM: DIFFERENCES IN RESPONSES TO ICT WORKPLACE SITUATIONS FOR ENGLISH AND NON-ENGLISH SPEAKING BACKGROUNDS.

Richards, Debbie, Macquarie University, North Ryde, NSW, 2109, Australia
richards@science.mq.edu.au

Busch, Peter, Macquarie University, North Ryde, NSW, 2109, Australia
busch@science.mq.edu.au

Abstract

Multiculturalism is one of the many factors contributing to the growing diversity found within the workforce comprising today’s organizations. This is particularly true in the ICT workplace. Diversity is often investigated in the context of teams and global organizations seeking to improve knowledge management and/or innovation strategies and practices. Understanding the role that culture plays in a multicultural society is complex, understudied and not well understood. Past studies into identifying and comparing different national cultures are difficult to apply to the multicultural context. Even the operationalisation of the culture construct can be problematic or impractical when many ethnic minorities exist. One approach commonly used in Australia is the categorization of individuals and groups into English (or Anglo) background and Non-English Speaking Background (NESB). In this paper we use these categories to investigate if differences can be found in the way in which individuals in these two cohorts respond to workplace situations.

Keywords: Multiculturalism, knowledge usage, non-English speaking background (NESB), workplace based scenarios.

1 INTRODUCTION

Much culture-based research focuses on organizational culture, seeking to answer questions such as what is the culture in a particular or type of organization (Cooke and Rousseau 1988); whether and how the current culture has a positive or negative impact on the organization (Alavi, Kayworth and Leidner 2005); or how a culture can be changed to improve such things as knowledge sharing and reuse (DeLong and Fahey 2000) or to encourage innovation (Amar 2004). However, when it comes to national or ethnicity-based culture, cultural diversity is treated as one, albeit major, independent variable in understanding the attitudes and approaches to knowledge use and sharing (Holden 2002). We seek to break down the ethnicity-based culture variable based on an analysis of empirical data collected in two multicultural Australian organizations and synthesize our findings with the existing culture literature.

Much of the information systems (IS) research to date has relied on Hofstede’s work which oversimplifies culture viewing it primarily at the national level (Myers and Tan 2002). New research is needed to consider the modern multicultural organization due to the effect culture can have on information systems in general and knowledge sharing strategies, knowledge management system design and development decisions, in particular (Ardichvili et al. 2006). The focus of the scant IS literature touching on multiculturalism is more from the viewpoint of diversity, with culture being one of the possible points of differentiation, and how diversity can impact on knowledge management strategies such as how to manage global organizations and teams to more productively create and share knowledge (e.g. Ely and Thomas, 2001; Haas 2006) or how innovation can be driven by putting people from the right cultures together (e.g. Dombrowski et al. 2007). Our focus is instead on knowledge usage and whether people of different cultural backgrounds employ different knowledge
and decision making strategies in the way they respond to problem situations. We see an individual’s response as demonstrating knowledge-in-action or practical intelligence (Sternberg, Wagner and Okagaki 1993) requiring the application of a combination of tacit, explicit and practical knowledge gained overtime through various experiences. In this way we seek to identify the role of culture in knowledge usage, particularly seeking to ensure the approach includes tacit knowledge usage where tacit knowledge is defined as knowledge that one uses to manage ‘oneself, others, and one’s career’ (Sternberg, Wagner and Okagaki 1993). We achieve this by capturing and comparing the responses of employees to a number of workplace-based scenarios along the lines of psychology Professor Sternberg (Sternberg et al. 1995). To narrow the knowledge context, the employees and workplace scenarios are within the domain of Information and Communication Technology (ICT).

2 OPERATIONALISING THE CULTURE CONSTRUCT

Culture has been defined as “the learned ideas, values, knowledge, rules and customs shared by members of a collectivity (such as those based on ethnicity, gender, sexuality, indigeneity, age, disability (Holmes, Hughes and Julian 2003, p.157)). When we speak of multiculturalism the notions of ethnicity and indigeneity are most relevant. Ethnicity has been defined as “the cultural background of a group of people who share a belief in common ancestry. A resource that can be mobilised for identification purposes” (Holmes, Hughes and Julian 2003, p. 154). Within Australia, some 244 ethnic minorities are represented (ABS 2006). The dominant group representing roughly 70% of the population when questioning ancestry (ABS 2006) are the Anglo-Celtic Australians or those of British or Irish background (Holmes, Hughes and Julian, 2003). Indigenous Australians form another separate grouping. Given the stigma of belonging to a minority, the term Non-English Speaking Background (NESB) is a more popular and less offensive term in the Australian context. Another alternative sometimes used, is Language Other Than English (LOTE), but because the term is also used as the name of a program to teach foreign languages at school, we will use the term NESB which is commonly used by the Australian Government as it allows needs to be addressed, such as the need for an interpreter, for a very diverse population. Membership of the NESB cohort is typically identified through questions concerning language/s spoken at home. Thus in our study we ask “if a language other than English is spoken in your home environment, please select these from the lists below”. Up to 4 languages in order of frequency of usage could be selected. We acknowledge this will not necessarily identify each person’s ethnicity but that in a multicultural country with cross-cultural marriages, many second generation Australians (including both authors) and a melting-pot of cultures around you, asking people to identify their ethnicity(ies) may also have been difficult for respondents to articulate. Also, while people tend to identify with the heritage of their parents and the “mother” country, it is questionable whether an individual who does not speak the mother tongue at all has truly been moulded by that culture and to what extent, since language is often quite tightly bound up with culture. This can be seen when a literal translation from one language to another results in nonsense because the ideas, values, customs, etc., embodied in words/phrases from the two languages differ greatly. Our phrasing of the question is different to asking “what languages do you know?” which would include languages learnt at school which does not reveal ancestry or significant cultural influences.

Adopting the above concepts of culture, ethnicity, Anglo, English speaking majority, ethnic minorities and NESB as our starting point, we found that our dataset could be divided into two roughly equal groups where English was the predominant language spoken at home (Anglo group) or a language other than English was the predominant language spoken at home (NESB group).

To guide our discussion and comparison of the cohorts that emerged from our data we used various cultural classifications, dimensions and characteristics from the existing culture literature from sociology, management and organization theory and practice. These include: Hofstede’s (1980) dimensions of power-distance, uncertainty avoidance, individualism-collectivism and masculinity-femininity; Hofstede and Bond’s (1988) concept of Confucian dynamism; Kluckholn and
Strodtbeck’s (1961) time-based types of cultures; Trompenaars and Hampden-Turner (1997) five value orientations: Universalism versus particularism, Communitarianism [collectivism] versus individualism (same as Hofstede’s dimension), Neutral versus emotional [affective], Diffuse versus specific, Achievement versus ascription; Concern for face (Ho 1976); Hall’s (1990) notion of High-context and Low-context and a number of other factors relevant to a study of culture. The complexity and dynamism of culture necessitates that while these characterizations and past studies may provide some explanation new research is needed (Myers and Tan 2002). We have conducted empirical research to identify what behaviours continue to exist in today’s multicultural society. Just as culture is a social construct, multiculturalism is a social construct which has ramifications for organizations in areas such as human resource management, knowledge management, organisational structure, deployment and adoption of technology and so on. In seeking to unpack multiculturalism we pose the following research question.

Research Question: Can patterns be found in the responses of ICT workers on the basis of their cultural background identified via languages spoken at home?

3 METHODOLOGY AND DATA COLLECTION

In seeking to find differences and patterns in the way that ICT workers apply their knowledge, we have employed a technique developed by psychologists at Yale (Sternberg et al. 1995) involving the use of workplace scenarios to discover how someone believes they would act in a given situation. An example of a scenario is shown in Figure 1. The technique is specifically designed as an instrument to measure tacit knowledge (Wagner and Sternberg 1991). Responding to a scenario involves problem solving and decision making. In this way by capturing the response we capture knowledge-in-action which involves the application of practical know-how, tacit and explicit knowledge acquired through education, training and experience. Cultural background may affect our education and training opportunities and achievements; culture will also affect how we (choose to) perceive, remember, (re)apply, receive and transfer the lessons we learnt both formally and informally (Young 2006). Thus we find scenarios to be an appropriate means to tap into the underlying influences of culture on the usage of knowledge.

To provide more structure than free text, assist participant decision making and allow quantitative analysis, the scenarios have a number of possible predefined responses which are measured on two seven-point Likert scales, to capture how the individual believes they would (realistically) and should (ethically) respond, as shown in Figure 1. This design follows from our pilot study in which we found that participants often wished to respond differently depending on whether they thought the response was something they “should do” (ethical response) or “would do” (realistic response). This corresponds to the findings of Wagner and Sternberg (1991a, 1991b). Figure 1 shows the list of six answer options for scenario 5.

Similarly, Chow, Deng and Ho (2000) designed two (2) scenarios concerning the day-to-day operations of a fictitious firm to test a number of hypotheses regarding knowledge usage. For our study we wanted greater ecological validity and domain coverage and used a process starting with interviews with 12 practitioners leading to a pilot survey in order to develop 16 ICT scenarios each with 6 to 13 answer options. The technique required practitioners to elaborate their ‘war stories’ or critical incidents they needed to resolve. Participants were made available to us by the organizations. There was no attempt to capture scenarios related to culture or any of the biographical variables collected in our dataset which included age, gender, languages other than English spoken at home, employment history, professional membership and educational background, with the intention of seeing if any of these variables independently or in combination, resulted in patterns of knowledge usage behaviour. We thus claim not to have intentionally built cultural biases into our scenarios. Each participant was randomly assigned 4 of the 16 scenarios.
A systems analyst in your section has done a superb job in designing a new system for a client organization. The system has yet to pass the coding and post implementation review stage, nevertheless you feel the job that was conducted is worthy of praise.

Because of the demand that exists in the IT industry, and the fact that you know this systems analyst is thinking of ‘moving on’, you are hesitant to immediately congratulate the person for fear of precipitating the person’s departure, believing they may take this as a possibility of gaining promotion in a different firm, or perhaps even branching out into their own company, more than likely with an IS management consultancy bias (something the person had once alluded to).

You are not actually this person’s boss, however you have worked successfully together on a number of projects and realise that it is unlikely you will be working with someone as easy to get along with again.

Rate each of the following responses in relation to the given scenario. It is advisable to read all of the responses before replying.

**Figure 1. Sample Scenario with 7-point ethical and realistic Likert scales plus answer options (Busch, 2008)**

### ETICAL

<table>
<thead>
<tr>
<th>Ethical</th>
<th>Extremely Bad</th>
<th>Neither Good nor Bad</th>
<th>Extremely Good</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose 1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### REALISTIC

<table>
<thead>
<tr>
<th>Realistic</th>
<th>Extremely Bad</th>
<th>Neither Good nor Bad</th>
<th>Extremely Good</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Talk as soon as possible with the analyst and ‘come straight out with it’, telling the person they did a good job, making no mention of employment possibilities in the industry, simply ‘assuming’ they will stay on.

2. Admit to yourself that the workplace is a professional environment and that the IT industry is competitive, the analyst deserves every chance they can get. With that in mind you don’t hesitate approaching the person to congratulate them even if it looks as if you are trying to ‘sidle up’ to them.

3. Consider talking to the team leader and mentioning the hard work put into the exercise by the systems analyst. What the leader decides to do from there is out of your control.

4. Do nothing positive or negative, simply ‘act normal’, if and when the analyst comes seeking advice of any kind you can simply let it slip that you thought they did an okay job.

5. Leave an anonymous congratulatory card on the person’s desk, something which has been word-processed after hours so that the handwriting is not a giveaway.

6. Next time a social occasion arises, just happen to mention ‘off the cuff’ to the analyst’s colleagues that you thought the person did a great job.

### 4 SAMPLE POPULATION AND DATA ANALYSIS TECHNIQUES

The complete data set included data from within the ICT departments of three Australian organizations of size small, medium and large. For the purposes of this paper focusing on culture within a multicultural environment, we have not included the data from our small organization because it was predominately of one ethnic background (Anglo-Celtic). The other two organizations were highly multicultural. The medium-size organization had approximately 1700 employees, 16 of which were ICT workers which can be briefly described as either a machine organisation or professional bureaucracy (Mintzberg 1991) whose main business was furniture retailing, but for whom the section of the organisation under study was the IT branch of the company. The large company had over 10,000 employees with around 1,400 ICT workers. This company was essentially a larger version of the previously mentioned organisation. From the medium size company, 13 of the 16 ICT workers chose to participate. In the large organization, despite having prepared over 1200 personally named
letters, only 165 were given the letters by the organization for political reasons unclear to us; resulting in 108 choosing to participate. Despite our disappointment, these are the limitations of using ‘real’ subjects from industry who can not be ‘voluntold’. To increase the sample size we have combined the data from both organizations. While no two organizations are identical and thus there could be differences due to variations in organizational type or culture, an internal analysis of the organization’s culture would have been a separate study. Further, to consider the effect of the organizational culture on individual behaviour would need to take into account how long each employee had been with the company and what other (types) of companies the individual had already been exposed to. Finally, the data was collected as individual responses and thus concerned individual knowledge usage, not organizational knowledge usage. Our main concern was that each individual included in the study was working in a multicultural ICT workplace.

Using language/s other than English spoken at home as an indicator of cultural background, we found that 69 individuals (49 males and 20 females) from 20-62 years of age reported no language other than English and thus formed our Anglo cohort. The remaining 52 individuals (30 males, 21 females and 1 unknown) whose ages ranged from 20-59 formed our NESB cohort. The distinction into Anglo and NESB groups fits the Australian political view of multiculturalism where Anglo is the ethnic majority and NESB comprise the ethnic minorities. Indigenous Australian languages were not reported, however a considerable number of European and Asian languages were. We note the predominance of males in our sample population, but point out that this roughly 65/35 gender split is higher than the distribution found within the Australian ICT industry and thus the possible bias towards male behaviours is consistent with the ICT workplace domain. Our lack of control over whom and how many participants in each organization, due to political and managerial reasons, resulted in fewer participants and uneven distributions across cohorts and scenarios. This also meant that only statistical analysis of the data was unlikely to provide statistically significant results and that qualitative techniques would be needed to combine the small numbers in order to discover differences and find cultural patterns.

Figure 2: Graphs for Scenario 5 Answer Options 1&3 for English speaking (Anglo) & NESB cohorts
As a first step we conducted quantitative analysis of the Likert-scale responses to each answer-option to determine if differences and patterns between the responses for each cohort could be found. For example, for scenario 5 answer option 1 the mean was 5.625 and 4.25 for the Anglo and NESB cohorts respectively. Based on the numbers we can not claim statistical significance, but we can see looking at Figure 2 that 87.5% (14/16) of the Anglo cohort were positive about this option, 2 were undecided realistically and only one thought ethically that it was a bad choice. In contrast, 41.6% (5/12) of the NESB cohort were realistically negative and only 58.3% were positive. Answer option 3 also shown in Fig. 1, had a means of 4.5 and 5.3 for Anglo and NESB cohorts, respectively, showing a stronger tendency towards favouring the option for the NESB group.

Once a quantitative difference was indicated, we began the qualitative phase of interpretation to determine if the responses were meaningful from a gender or cultural point of view. To achieve this we used a modified version of the technique known as member checking (Lincoln & Guba, 1985). This technique involved testing the data, analytic categories, interpretations and conclusions with members of those groups from whom the data were originally obtained. Acquiring a Likert scale responses to precoded choices does not reveal why someone chose that response. Member checking allows the opportunity to ask why and to gather further clarification and data. However, since the survey participants were anonymous to us and we had no direct access to them, we needed to find others of similar backgrounds to validate our interpretations. However, limitations of member checking have been found (e.g. Morse (1994), Angen (2000) and Sandelowski (1993)) due to conflicts between the researchers and members interpretations, goals and views. Furthermore, it has been found that members may regret having made a certain statement or choice and wish to have their data changed or removed. For these many reasons, we chose then to identify other NESB individuals from the main language groups represented and to ask them if they could interpret why that choice might have been made. We selected four scenarios which showed quantitative differences. Figure 4 includes some extracts from what was provided to our “member group”.

<table>
<thead>
<tr>
<th>PART A: Bio and Introduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the following pages you will be provided with a scenario and a set of possible responses. You will be shown a total of four scenarios and their responses.</td>
</tr>
<tr>
<td>In a previous study we have asked others working in ICT to read each scenario, all the responses for that scenario and then rate each of the responses in relation to the given scenario.</td>
</tr>
</tbody>
</table>

PART B: For the scenario you just read, others in your cohort (gender and cultural background), have given certain responses; please indicate why you think they may have responded in this way? For the cultural explanation consider things such as the importance of individual achievement, age, seniority, hierarchy, consensus, working as/belonging to a group, decision making processes, methods of communication, role of family/relationships, etc.

1. Talk as soon as possible with the analyst and ‘come straight out with it’, telling the person that they did a good job, making no mention of employment possibilities in the industry; simply ‘assuming’ they will stay on.

   **Females** were mixed in their responses, tending slightly more to choosing bad **realistically**. Ethically most were undecided. Can you suggest why from a **gender** viewpoint?

   **Your language group ethically** thought it was good or were undecided. **Realistically** no one was undecided but spread across good and bad with slightly more tending to good. Can you suggest why from a **cultural** viewpoint?

*Figure 3: Extract from Member Group survey used for interpretation of Likert scale responses.*
Table 1: Biographical details of NESB member checkers of the scenario response results

<table>
<thead>
<tr>
<th>ID</th>
<th>Gender</th>
<th>Language</th>
<th>Position</th>
<th>ACS Level</th>
<th>Yrs exp.</th>
<th>Highest Qualification</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>F</td>
<td>Indonesian</td>
<td>User Interface Designer</td>
<td>1</td>
<td>9</td>
<td>Masters(Research)</td>
<td>31</td>
</tr>
<tr>
<td>F2</td>
<td>F</td>
<td>Tamil</td>
<td>Senior S/W Engineer</td>
<td>3</td>
<td>10</td>
<td>Bachelor</td>
<td>32</td>
</tr>
<tr>
<td>F3</td>
<td>F</td>
<td>Mandarin/Chinese</td>
<td>Software Engineer</td>
<td>2</td>
<td>12</td>
<td>Graduate Bachelor</td>
<td>38</td>
</tr>
<tr>
<td>F4</td>
<td>F</td>
<td>Bulgarian</td>
<td>Software Engineer</td>
<td>3</td>
<td>15</td>
<td>PhD</td>
<td>42</td>
</tr>
<tr>
<td>M5</td>
<td>M</td>
<td>Marathi, Hindi</td>
<td>Technical Project Mgr</td>
<td>4</td>
<td>11</td>
<td>Graduate Bachelor</td>
<td>46</td>
</tr>
<tr>
<td>M6</td>
<td>M</td>
<td>Vietnamese</td>
<td>Research Engineer</td>
<td>2</td>
<td>6</td>
<td>PhD</td>
<td>30</td>
</tr>
<tr>
<td>M7</td>
<td>M</td>
<td>Mandarin</td>
<td>Software Engineer</td>
<td>3</td>
<td>3</td>
<td>Bachelor(Hons)</td>
<td>24</td>
</tr>
</tbody>
</table>

We approached a large multinational software development company to provide us with participants working in ICT who would all share the same organisational culture to act as member checkers. We sought to obtain 5 males and 5 females from each of the main NESB language groups we identified: Eastern European (Russian (2), Serbian (3), Croatian (1), Bulgarian (1), Hungarian (1)), Mediterranean (French (3), Greek (3), Spanish (1); Subcontinental (Hindi (3), Tamil (2); Kannada (2); South-East Asia (Indonesian (4), Tagalog (1); Vietnamese (4); East Asia (Cantonese/Chinese (15), Hakka (1), Hokkien (1), Korean (3), Japanese (1). The company were unable to identify any Mediterraneans or male Eastern Europeans in technical roles. However, the 7 member checkers were from language and gender groups covering 79% of the NESB cohort. Each of our 7 member checkers, summarised in Table 1, received the same four scenarios (ones that we identified to have numerical differences and potentially cultural or gender related interpretations) however, each received the scenarios in a different order to avoid order effects. The IDs in Table 1 are used later in Table 2 to identify the source of a quote.

Using thematic analysis (Boyatzis 1998) which uses recurrence (same thread of meaning), repetition (of keywords, phrases or sentences) and forcefulness (volume, inflection, emphatic language) to identify, analyse, describe and report patterns (themes) across qualitative data, we identified some consistent patterns within the interpretations. Initial themes were reviewed against the cultural traits (e.g. individualistic, achievement oriented) identified from the body of literature that we had studied. Our goal was to see if the behaviour patterns corresponded to a cultural dimension so that we could apply the appropriate label. This would position our findings within the culture literature and answer whether the cultural traits identified primarily through study of one or a small number of national cultures, were still evident within a multicultural society. For example, in the box asking for interpretations of why males had responded a certain way males used the following words: active, proactive, power (2), take on chin, not nag, logical, well thought out, take control (2), problem-solving, don’t leave problems to others, take problem head on, assertive (3), authority, decisive, take charge, take responsibility (2), responsible (3), leader has responsibility, take ownership, avoiding problems is bad (5), runaway is bad, cover up is bad (3), frank (4), truthful, straightforward (2), forthright, factual (2), simple, courage, confident, bad to ask for help (2), clarification-seeking advice okay if junior, well-organised, ambitious (2), like challenge, need achievements acknowledged (3), should give praise and feedback (4), team oriented (3), professional, egoistic (2), work hard to get job done (2).

When females were asked why the females had responded a certain way they used the following words: don’t want to waste time/efficiency (4), avoid overtime/overwork/too much work (3), negotiating (2), involving team, teamwork, professional (13), make own choice, methodical, analytical, consider friends, work and responsibilities, avoid hurting feelings/upsetting others (2), communication and discussion good (6), open communication (3), manipulation/lying/sneak is bad (4), bad to runaway from problem, bad to “pass buck”, good to solve issues, good to take action, avoid confrontation, good to express respect and acknowledge achievement, career, success (3), actively involved, proactive, want to work, do good job, [need to] take responsibility/ownership, gentle, attached, avoid upset, should observe hierarchy (3), hierarchical boundaries, seniority.
The words in bold indicate concepts not expressed by the opposite gender. Interestingly the same individual may have used that term when referring to culture but not to their gender. For example, M2 used the term ambitious 5 times in her interpretations from a cultural viewpoint but not once in her gender interpretation. Terms which arose consistently in the NESB cultural interpretations for members of both genders included: relationships, open communication, avoid conflict/trouble/hurt, deserve recognition/give praise where due, should know place/not interfere, backseat role, teamwork and building is important, individual achievement deserves recognition, praise may be given and received differently, praise should come from superior, shouldn’t say no to a request from a superior or coworkers, consensual decision making, need to go with flow, follow order and social standing, observe seniority, customer is god, group decision, democratic processes rather than authoritative, respect other person, consider well-being of others, okay to have personal/individual career goals, take responsibility for own work and problems, avoid conflict.

Our own interpretations of the responses were based on cultural characteristics identified in the literature including: Collectivism/Individualism (Hofstede 1980; Triandis et al 1988; Trompenaars and Hampden-Turner 1997), Ascription/Achievement-Oriented (Trompenaars and Hampden-Turner 1997) (overlapping with Newman, Summer and Warren’s (1977) improvement/maintaining status quo and merit-based/relationship-based)); and High/Low Context (Hall 1990); concern for face (Ho 1976), diffuse/specific (Trompenaars and Hampden-Turner 1997), affective/neutral (Trompenaars and Hampden-Turner 1997) (similar to Newman, Summer and Warren’s (1977) objective/emotional). In table 2 we provide a summary and snapshot of comments and responses. The interpretations of responses by the same language group revealed a great deal of similarity with the gender responses. This is not surprising since gender and culture are social constructs. Sometimes the response was identical and expressed as “as/see above”. Sometimes a cultural explanation was expressed as “fe/males in this culture …”.

In some cases the strong connection between gender and culture was apparent because a gender specific interpretation was given in the cultural interpretation. For example “females rarely speak up their opinions and more likely to take a ‘backseat’ role even if they don’t agree” (F1). Via this process of qualitative analysis of the member interpretations we identified the following themes: concern for group, group decision making, importance of observing hierarchy and seniority, open communication, importance of relationships, taking responsibility and making decisions, with some individuals expressing individual achievement, ambition, professional and career orientations. Given that the majority of our NESB cohort and member checkers were from cultures identified as collectivist, the more individualistic views offered may indicate that these individuals were attracted to Australia because it is individualistic (in Hofstede’s (1980) studies of 50 cultures Australia was identified as the second most individualistic country after the US). Alternatively, Australia’s culture is likely to have influenced their own views.

When we compare these interpretations with characterisations of genders and the national cultural literature we see that some changes have occurred. These changes may be the result of changes over time, original incorrect characterisations or the change in these individuals due to a process of multiculturalisation. We note that the males when asked to interpret from the point of view of the males, males used terms stereotypically (see Broverman 1972) associated with males (power, challenge, ambitious, assertive, decisive, control). Similarly, our females did the same (gentle, avoiding hurting (ie. Nurturing and caring), following hierarchy (knowing their place)). When it came to a cultural interpretation some of these gender boundaries were relaxed and any of these characteristics could have been used to describe why a response was taken from a cultural viewpoint.
<table>
<thead>
<tr>
<th>Characterisation of Option</th>
<th>Anglo</th>
<th>Interpretation</th>
<th>NESB Interpretation</th>
<th>Member View of NESB response (not of interpretation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protest about a bad decision made by a superior</td>
<td>+ve</td>
<td>Assertive</td>
<td>-ve</td>
<td>Accept authority, cautious in speaking mind. Greater respect for structure than own opinion.</td>
</tr>
<tr>
<td>Question a superior’s decision you see as outdated</td>
<td>Not afraid to put forward your own skills/knowledge</td>
<td>Accept authority, cautious in speaking mind. Greater respect for structure than own opinion.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do what you are told without question even if you know a better way hoping to be rewarded later</td>
<td>-ve</td>
<td>Experience (achievement) may outstrip and overtake seniority/rank</td>
<td>+ve to very +ve</td>
<td>Cognizant of power distance, seeking improvement of your position through approval of superiors</td>
</tr>
<tr>
<td>Let the boss pass on the praise and be responsible for the consequences</td>
<td>-ve</td>
<td>Your achievements entitle you to acknowledge the achievements of others</td>
<td>+ve</td>
<td>Concern to follow the hierarchy, observe power distance</td>
</tr>
<tr>
<td>Acknowledge the good work of another even though it might prompt them to look for a better paid job</td>
<td>+ve</td>
<td>Achievements should be acknowledged and rewarded, it’s a competitive world. Speak your mind.</td>
<td>-ve</td>
<td>Not taking on what you see as someone else’s role. Focus on the long term and on the organization/group rather than gratification of the individual.</td>
</tr>
<tr>
<td>Let team member with insufficient skills carry on and someone else do the work</td>
<td>-ve</td>
<td>Not fair, everyone should pull their weight. Project and short term focused</td>
<td>Marginal +ve</td>
<td>Nurturing, avoid hurting feelings</td>
</tr>
<tr>
<td>Work unpaid overtime to fix problem</td>
<td>-ve</td>
<td>Materialistic, egotistic (can do a better job) Concern for self, materialistic</td>
<td>+ve</td>
<td>Wanting to please others, longer term focus. Think of greater good, concern for group</td>
</tr>
<tr>
<td>Put in extra work to do the job yourself rather than ignore or use incompetent people</td>
<td>+ve</td>
<td>Prefer to work alone, trust their own abilities</td>
<td>Neutral</td>
<td>Prefer to work as group, trust the group</td>
</tr>
<tr>
<td>Getting a mentor or advice</td>
<td>Mildly +ve</td>
<td>Willing to take advice on board but would rather do/solve themselves</td>
<td>+ve</td>
<td>More inclined to work with others in their ingroup, make use of networking and personal relationships, loss of face</td>
</tr>
<tr>
<td>Use polite communication to resolve the issue</td>
<td>Neutral</td>
<td>Communication has a role to achieve outcomes and results, offending people to get the work done may be unavoidable</td>
<td>Very +ve</td>
<td>Projects and tasks will come and go, but relationships are to be nurtured and thus communication must consider the impact on people.</td>
</tr>
</tbody>
</table>

Table 2: Summary of selected answer options and results.
5 DISCUSSION

The small and variable sample size for individual scenarios and the subtleties which can exist in culturally-influenced responses meant that quantitative statistical analysis did not reveal statistically significant results. In the study by Chow, Shields and Wu (1999) involving a comparison of the openness of knowledge sharing within US and Chinese organizations, qualitative analysis of the interviews and responses to closed and open-ended questions using two IT workplace scenarios were able to reveal patterns of differences, which quantitative analysis had been unable to detect and, “in one instance, potentially misrepresented the underlying motivations and behaviors” (Chow, Shields and Wu, 1999, p.580). The value of employing and comparing the results of quantitative and qualitative analysis are captured in the sentiment “Statistics tell one story; [qualitative] research tell the story behind that story” Trauth (2002, p.5).

In our methodology we first identified interesting differences by looking at the numbers and graphs for each answer-option and then qualitatively interpreting the type of behaviour embodied in the scenario and answer-option. However, we were interested to determine if answer-options without (interesting) differences also contained culture-related concepts. After qualitatively evaluating and classifying all 125 answer options, according to the characteristics and dimensions identified from the culture literature, we found that 75 combinations out of the 125 potentially contained cultural issues. That fact that we only found differences worth investigating in 31 of these 75 culturally-relevant scenarios reveals the subtleties which exist in examining culture in a multicultural context.

One approach to multiculturalism is to manage culturally-based expectations (Evaristo 2007). This is firstly only possible if the differences in expectations are understood. Expectations could be managed by enhancing cultural sensitivity training to go beyond fact and generalization to include (virtual) workshops at which individuals of different cultures describe their expectations to specific situations (Evaristo 2007). Indeed this is similar to the workplace scenario-based approach we used to develop our model. We conjecture that the 31 scenarios which showed some cultural differences could be good starting points for discussion.

Understanding the influences of culture is understudied. People are a major component of information systems and people’s attitudes to technology, knowledge codification and sharing are shown in other studies to be influenced by culture (e.g. Ardichvili et al. 2006; Chow, Deng and Ho 2000; Michailova and Hutchings 2006). From an IS practitioner’s point of view, culture is particularly relevant to knowledge management. For example, understanding cultural influences across the organization, which may vary not only between offices in different countries and states but also within the same geographic location will have implications for the design and implementation of strategies such as the viability of using: email for knowledge sharing, the role of Communities of Practices and group decision support systems and other team-based activities such as system development and project management. Similarly, while diversity is seen as essential for innovation, a past-oriented (Kluckholn and Strodtbeck’s 1961), collectivist and ascription oriented culture may be less inclined to diverge from what has been done in the past or question what the group wants, particularly senior members of the group or accept new group members. Management will need to understand the basis of this reluctance and ensure that the new team members are appropriately integrated into the ingroup perhaps through formally organized social events or team building activities, known as managed socialization (Moitra and Kumar 2007). One strategy is to work with existing ingroups and knowledge intermediaries while new intra-organizational groups are formed to facilitate knowledge sharing (Michailova and Hutchings 2006). Ely and Thomas (2001) found that diversity of itself does not necessarily bring benefit. An integration and learning perspective was seen to be the key to success by providing rationale, guidance and motivation to deliver sustained and maximized benefits from diversity. Lau and Murnighan (1998) found that the key to handling diversity was to understand the faultlines within a group so that they could be understood and managed. Cultural differences potentially pose a fault-line and thus understanding the differences can assist in managing and
composing groups. Harrison et al (2002) discusses the notion of deep (psychological) versus shallow (demographic) diversity. By bringing deeply embedded differences and similarities to the foreground, including our culture-based belief systems; can promote deep level diversity leading to better social integration and resulting in better performance (Harrison et al 2002). Our study which reveals differences in the ways employees may put their knowledge into action and how ways of thinking and behaving can be aligned to cultural influences can promote appreciation of deep diversity.

In our study we take the common view of ‘multiculturalism’ to mean people from different nationalities/countries/ethnic backgrounds and a multicultural organization to mean its employees come from a wide range of cultural backgrounds. To find interesting patterns of knowledge usage behaviour, similar types of cultural groups may need combining. For example, while we believe the United States, United Kingdom, Australia, Canada and New Zealand have different cultural traits it will be much harder to find differences in behaviour as the differences within these countries is probably at least as great as the differences between the nations. On the other hand if we compare these Anglo countries with Latin countries such Italy, France, Spain, Portugal, Argentina, Brazil, Chile, etc. differences are expected to be noticeable. Indeed when it comes to acting upon the range of behaviours, for example with respect to knowledge management practices and policies, it would not be feasible to try to accommodate each individual ethnic group. The fact that we were able to identify patterns of behaviour by combining ethnic groups supports this approach. While not explored in this study, smaller groupings, say four rather than two cohorts based on ethnic background, could reveal even stronger patterns, but with fewer numbers the validity would be reduced.

In answer to our research questions we were able to find patterns in the responses of ICT workers on the basis of their cultural background identified via languages spoken at home. Understanding culture in the not so distant past, promoted tolerance and assisted trade. Understanding culture in the increasingly multicultural organizations and societies of today is no longer an outward looking concept. Nowadays awareness of our own culturally-influenced attitudes and behaviors and those of the people we work and live with is essential if we are going to survive. Education and training can ensure that we “know” the same things, but our culture will influence how we value, interpret, utilize and share that knowledge. Similar to techniques used to help employees to evaluate their own personality, leadership and/or decision making styles and those of their colleagues in order to work together more effectively, we recommend that organizations train their employees, managers in particular, to understand and recognize cultural differences to improve human resource, technology and knowledge management practices within an organization. Identification of responses to workplace scenarios can contribute to this understanding.

6 REFERENCES


Hofstede, G. and M. Bond (1988). The Confucian Connection: from cultural roots to economic growth Organizational Dynamics. 16(4), 4-21.


## AN INVESTIGATION OF HOW CULTURE IMPACTS GLOBAL WORK: UNPACKING THE LAYERS OF CULTURE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0163.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Cross-cultural issues, Cross-national study, Culture fit / differences / heritage/ intelligence / issues / theory / values, Collaboration</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
AN INVESTIGATION OF HOW CULTURE IMPACTS
GLOBAL WORK: UNPACKING THE LAYERS OF CULTURE

Chand, Donald, Bentley University, IPM Department, 175 Forest Street, Waltham, MA 02452, USA, dchand@bentley.edu

David, Gary, Bentley University, Sociology Department, 175 Forest Street, Waltham, MA 02452, USA, g david@bentley.edu

Galliers, Robert, Bentley University, Provost Office, 175 Forest Street, Waltham, MA 02452, USA, rgalliers@bentley.edu

Kumar, Senthil, Icfai University, Management Department, Hyderabad, India, sengon@yahoo.com

Abstract

One manifestation of today’s globalization is a new type of work environment where employees who are living in different parts of the world regularly engage in collaborative activity. Because these workers are socialized in different cultural environments, an important research issue is how to understand workers behaviors and effectively manage in this global workplace were workers communicate on a daily basis but do not meet each other and may not even know the life styles of their team members. Based on an ethnographic analysis of GLOBALIS, a multinational firm in the financial sector with IT solution centers in the US, Ireland and India, this paper presents a conceptual model of how culture was impacting the behaviors of global teams at GLOBALIS. In this model the behaviors of globally team members is primarily governed by corporate norms, project requirements and workplace culture. National culture, which has been identified in cross-cultural research as a key factor in global work, is just one of many demographic variables like age, education, professional standards, inter-relationships among sites, etc. that impact workplace culture. Thus, this paper reinforces the current multi-faceted view of culture as a set of forces that impact values and behaviors and it also clarifies some of the clouded opinions on whether or how national culture differences impact today’s globally distributed work.

Keywords: Global Work, National Culture, Corporate Culture, Project Culture, Professional Culture, Workplace Culture, Time, Offshore Outsourcing, Ethnographic Case Study, Survey Research.
1 INTRODUCTION

One manifestation of globalization is the movement of multinational firms across their countries’ geographic boundaries in order to access qualified labor pools, tap into hubs of talent, reduce costs through wage differentials, and achieve proximity to customers, thereby becoming truly global in their operations. This movement is changing the way in which work is performed. In this new form of work organization, employees routinely work with team members who reside in other nations. This new way of working is made feasible by advances in information and communication technologies (ICTs). ICTs enable workers to interact, communicate and coordinate their work across great distances, such as is found in virtual global software development teams.

While ICTs are a necessary condition for such work, it is by no means sufficient by itself. While human determination, commitment and innovation have yielded success, distance has shown to contribute complexity to the communication, coordination and control processes in global software development (Herbsleb et al. 2000). Because workers live in different countries and thereby are socialized in different national cultures, investigation of the role of culture in global IT work is an important research topic (Borchers 2003, Chand et al. 2007, Cramton & Webber 2005, Galliers 2003, Galliers 2008, Krishna et al. 2004, Symons & Stenzel 2007; Walsham 2002). Notwithstanding, there are many views on the role of national culture in global IT work that vary from ‘culture does not matter’ to ‘the success of global projects requires a keen understanding and management of cultural differences’ (Carmel & Tjia 2005).

Even though culture is a frequently used term, there is no commonly agreed upon definition regarding specifically what it means (Straub et al. 2002). In daily usage the word culture is associated with observable things such as food, language, literature, music, art, etc. that characterize a community or a group. At the country level we quickly get into trouble with this ‘observable things’ view of culture. For example, is it appropriate to say that India a country with 15 major languages in use, conflicting religions and many races has a uniform national culture? Sir John Strachey who had spent many years as a member of the Governor-General Council in India observed that “Scotland is more like Spain than Bengal is like the Punjab (Guha 2007). This observable diversity does not mean that India does not have a unifying essence. This example of India just points out that the observable things view is insufficient to fully capture the concept of national culture.

Edward T. Hall (1959,1981), a pioneer in the field of cross-culture studies, conceptualizes culture as a means of communication through “silent languages.” He identified five silent languages of cross-cultural difference, namely the use of time, physical space between people, attitude toward material goods, friendship and bounding, and forms of agreement. His thesis is that one can use these silent languages to highlight the cultural differences among groups, communities and countries. For example, on how societies used time Hall observed that Americans think of time as linear and discrete, are monochromic and tend to do one thing at a time, adhere to schedules, believe that meeting deadlines and keeping appointments is important, and they consider keeping someone waiting to be a sign of disrespect. On agreements, Hall observed that Americans do not rely on communication that is deeply embedded in cultural context and therefore they tend to negotiate contracts that are explicit, complete, and literal.

Geert Hofstede, a Dutch academic widely known for his pioneering research on culture, surveyed 116,000 employees of IBM, a large multinational company, spread across fifty countries and three regions and two points in time (Hofstede 1980, 2001). The questionnaire items measured work-related values. Factor analysis of the responses grouped the questions into four categories that he titled as Individualism-Collectivism, Power Distance, Uncertainty Avoidance and Masculinity-Femininity.
Later he added a fifth dimension of culture. When he grouped the responses by countries for each of these dimensions and mapped the responses to an index value between 0 and 100, the country indexes revealed differences between countries along these dimensions.

Hofstede built his theory of culture (Hofstede 1980) based on this study. His key finding was that country-based differences persisted among the responses of the IBM workforce despite the ‘strong’ corporate norms of IBM. To explain this finding Hofstede defined culture as “the collective programming of the mind that distinguishes the members of one category of people from another” and this collective programming of the mind occurs through the socialization processes. This definition frames culture as a set of norms and values that people are socialized into, while growing up, by their family, neighborhood and the schools.

Schein’s (1985) sees culture represented as the layers of an onion. The outer layer is what people generally associate with culture, namely the observable things related to behaviors such as arriving late to events, maintaining leisure time completely separate from work time, the inability to say no, etc., and artifacts such as language, food, literature, etc. The middle layer refers to the norms and values which a community holds. The norms define how to act and what is considered right or wrong, and values tell what is considered good or bad. The core represents the underlying assumptions. According to Schein, to understand the root causes underlying behavior or values, the underlying assumptions have to be surfaced.

Grounding their analysis on social identity theory, Gallivan and Srite (2005) argue that culture should be viewed as a richly layered set of forces that go beyond national culture dimensions and corporate culture norms in shaping individuals’ beliefs and actions.

Divergent conceptions of time, based on national cultural differences, are frequently cited as a fundamental barrier to collaborative work (Hall 1959, Hall 1983, Cusner & Brislin 1996, Olson & Olson 2003). Perhaps one reason why time is so often cited is because divergent adherences to schedules and deadlines are so readily observable, and thus reportable. Also, in the business and IT world, time is a crucial variable in defining whether a project is successful and/or on time. In terms of cultural differences as to how time is perceived, the more distant a culture is seen to be geographically and normatively, the greater the expected difference. This means that in teams that are globally dispersed, time is assumed to be a critical problem in developing collaboration.

This paper explores the role of culture in global work by investigating the use and attitude towards time by the employees of GLOBALIS that were distributed across multiple sites in the US, Ireland and India. Our study is exploratory in the sense that the examination of the IT unit of one multinational firm cannot stand for the national cultures of US, Ireland or India. Further, the results we derived based on an empirical analysis of IT workers cannot automatically be generalized to other occupational groups. Thus, what the inductive analysis of this study provides is a nuanced understanding of the sub-cultures and factors that affect workers’ behaviors in global work.

In the next section we describe the research site, the data sources and our research methodology. After this we present our ethnographic and survey data that is used later to construct and synthesize the conundrum of culture layers into a new model of how these different layers of cultures impact global work. The paper ends with a discussion of the implications of this model and the future research that are needed to place this model on a sounder footing.

---

1 In subsequent work with Chinese scholars, Hofstede added a fifth dimension called short-term v. long-term orientation
2 RESEARCH SITE, RESEARCH METHODOLOGY AND DATA GATHERING

Over the last twenty-five years GLOBALIS, a Fortune-500 firm in the US, has successfully launched and operated multiple IT software and systems services facilities, called solution centers or sites from now on in the paper. Today, GLOBALIS has solution centers in multiple cities in the US, in Ireland and in India. Although GLOBALIS has considerable experience in establishing solution centers and is well aware of industry’s best practices for dealing with issues and problems of geographic dispersion: distance, time zone differences, cultural differences, loss of communication richness, etc., it is facing difficulty in unifying these sites into a cohesive community. The data used in this paper was collected over a span of three years.

We interviewed forty employees of GLOBALIS that included six senior management personnel, six project managers and twenty-eight workers associated with the four projects. These interviews occurred in both individual and group format and most of the interviews happened in person. Each interview lasted approximately one hour in length. These interviews were semi-structured and conversational in nature, and they covered a range of topics related to GLOBALIS’ global delivery systems.

We made repeated visits to the onshore and offshore sites associated with the project. During these visits, we also observed the nature of the work associated with the project, especially meetings and other situations where people from the various sites interact with one another. The Ireland sites were visited in 2003 for one week, and again in 2006 for another week. The India sites were visited over ten days in 2006. During the course of the project, the sites in the New England area were visited intermittently in order to attend meetings, conduct interviews, and deliver reports.

We observed eight engagement managers monthly video-conference meetings over a period of ten months, first six at the US site, seventh at one of the Ireland sites and the eighth at one of the two sites in India. We monitored two team meetings conducted via conference calls between the team leader in the US and testers in India. We also surveyed the workers associated with the four projects we tracked. The team members on these four projects were distributed across the New England sites, Dallas, Texas, Galway, Ireland and Bangalore, India. In addition, we made six management presentations of our findings at the headquarters site in US, in Dublin and in Gurgaon, and received valuable feedback.

Unlike intercultural communication research, which imposes a cultural framework on the interactions, and assumes that the source of problematic interactions is cultural, we were interested in “what more” (Garfinkel 1996) could be accounted for outside of the corpus of literature and research methodologies associated with the study of globally distributed work teams. Thus, we were not looking for the impact of national culture on virtual team collaboration per se. We were not assuming that the work process was problematic. Rather, we were interested in what emerged from the project, rather than trying to confirm previous studies or testing existing theory. In this way, we were following a workplace studies approach abstaining from broad a priori application of theoretical concepts and opting instead for a detailed analysis of in-situ social order (Suchman 1987, Orr 1996, Luff et al 2000).

Our analysis is based on the following four data gathering situations: engagement-managers monthly video conferences, cross-cultural training course, survey of attitude and use of time and our visits of the sites in Ireland, India and the US.

2.1 Monthly Video Conference Meetings

When we started our study, GLOBALIS had just made a decision to establish a new solution center in Bangalore and was engaged in defining and implementing a new global delivery process (GDP). We attended six monthly engagement managers meetings at the headquarters (HQ) of GLOBALIS in the New England region in the US. They were two hours meetings that started at 8 AM eastern standard
time. The management teams at Dallas, Dublin and Gurgaon were connected to these meetings via a video interface. The Dublin and Gurgaon sites were always on the meeting agenda, but their role was relegated to making presentations on the status of their projects, resources and issues. When we observed the same meeting from the Dublin and Gurgaon sites, we noticed that the Dublin and Gurgaon participants would put their audio part on mute, in effect cutting them off from the dialog taking place at the HQ site. Twice these monthly video conferences started late because the video equipment at the HQ site malfunctioned.

When we probed why the Dublin and Gurgaon sites mute their systems, we found that whenever anyone at the remote sites speaks, the camera focuses on that person and the face of the speaker appears on the television screen in the meeting room at the HQ site. This would happen with slightest utterance of a sound such as someone sneezing. To avoid disturbing the conversation, the Dublin and Gurgaon sites would mute the audio portion of their systems.

We observed that every video conference meeting was “run” by a manager at the HQ site, meaning time was kept from HQ, the meeting began from HQ, the agenda was followed or altered from HQ, etc. Questions asked from the distributed sites tended to be clarification questions, while the questions from HQ were accountability questions, (why something was not done, project delays, employee allocations, etc.). We never observed accountability questions raised by people at the distributed sites. Thus, accountability was a uni-directional practice originating from HQ.

In addition, all meetings were scheduled to start at 8 AM eastern standard time in the US. This meant that it was 1:00pm in Ireland, 7:00pm in India, 7:00am in Texas and 6:00am in Utah. Thus, people in New England were beginning their workday with a meeting, while workers in Ireland were having their day interrupted and people in Texas, Utah and India had to work outside their normal work hours. This is an indication of who has to adjust to whom. But more significantly, the people at the headquarters behaved as the decision makers and the people at the remote site behaved as vendors, suggesting an unequal relationship among the sites.

We also observed that the people at the two Indian sites did not understand at least 40% of the discussion taking place among the participants in the meeting room at the HQ site. This was partially because of the quality of the audio signal and partially because they did not understand the accent and were not fully cognizant to the implications of the issues being discussed. Furthermore, the participants in the HQ meeting room often used humor and jokes to convey their positions or even displeasure, and most of this humor did not carry over because most Indian were not familiar with the common American expressions.

In summary, our video-conference observations suggested a hierarchical relationship among the sites and the expression and usage of the English language in the US made it hard for the Indian participants to understand the video conference conversation.

2.2 Cross-Cultural Training

GLOBALIS contracted with a consulting company to train their employees on how to do business with India. This training was a one day program where managers were exposed to generalized traits and characteristics of Indian culture while developing an awareness of their own culture. We attended this training program and found that like most cross-cultural training programs, the focus of this training was on exposing the American workers to those cultural traits of Indians that were different from Americans such as American society is very individualistic and Indian society is group-oriented; Americans are egalitarian and Indians are very hierarchical; Americans view time as linear whereas time in India is cyclic and event driven; American are monochromatic, that is, they prefer to work on one task at a time, and Indians are polychromatic and they often engage in multitasking; etc. No equivalent training was given to US managers on doing business in Ireland. More significantly, the Irish managers were never given any training on doing business in India. On the other hand, when we interviewed the Indian workers, we found out that they were given a short 3 hours program on working with US. Unfortunately, the people we interviewed could not remember much about that
training except that they were told not to use “but” in their email communications with US workers and managers.

When we asked the workers whether national culture differences impact their work, the response was always “they did not see national culture playing any significant role.” However, when we walked into the cafeteria for lunch at a New England site we noticed that all the workers from India who were in the States for a training program were sitting on one table whereas their American counterparts were distributed on other tables away from their Indian visitors. The exception was one Indian woman trainee who was sitting with other American women workers. When we asked the male Indian workers why they did not socially mix with their American colleague at lunch time, their response was that they did not know enough about the American pop culture to engage in informal chit chat. When we asked the Indian female worker who was sitting with other American women what they talked about, she said that since she was getting married soon, the conversation dwelt on weddings in India and in the US and how the bride dresses, etc.

Although workers in the US, Ireland and India told us that their work suffers from many communication problems but none of these workers attributed any of those problems to culture differences. However, the managers felt that cultural differences were really impacting behaviors. We present three stories, one by Scott, who was a project manager at one of the New England sites, the second by Seamus, a project manager at one of the Dublin site in Ireland and the third by the site manager, Alka, in India.

Scott was managing a team of US and Irish workers working on a critical project. Late in the project life cycle new team members from India were added to the project as additional resources to keep the project on schedule. Scott observed that the new team members from India were not as responsive and prompt as the original team members. After attending the cross-cultural training Scott realized that because time in Indian culture is elastic, the Indian team members may not fully grasp the project’s sensitivity to deadlines and schedules. To remedy this situation, he said he micro managed the Indian team but could not improve the behavior.

Seamus was the manager of a team which consisted of Irish analysts, architects and integrators and Indian software engineers and testers. The project involved upgrading of an existing system for an old valued customer, and the Irish team members had extensive knowledge of the problem domain and the current system. Seamus asked the Indian team to start coding from a set of incomplete specifications. The Indian team refused. Seamus was very frustrated and he complained to us that the Indian workers are too process-based and they don’t understand the relationship culture of GLOBALIS. This is an example of a communication conflict that has been framed in terms of professional cultures.

During our conversation with the site manager, Alka, in Gurgaon, India we inquired whether she had more information on the content of the three hours of cross-cultural training provided to the Indian workers. She did not know the details and she asked us to tell her what we have heard from our interviews. We told her that people do not remember the details except that they were being trained to communicate more directly. We suggested that because Hall (1981) had identified US to be a low-context culture, the training course may be focusing on teaching how to engage in direct and detailed communications. Alka mentioned that one source of communication difficulty is that ‘we Indians have difficulty saying No’ and that often leads to ambiguity and misunderstanding. We asked her to provide us an example. She tried to explain this in the following manner. When my counterpart in the US tells me that Alka I have given my word to the customer that the new system will go live on October 1, are you sure you can meet this deadline? Alka feels that even when she knows that it will be hard to meet the deadline, she cannot just say no and instead her response will be “Jack, you know it is hard to predict all that can happen when you are doing software development, especially when the technology we are using is new. But you know us, we are committed and will do our utmost to make the deadline.”
2.3 Attitude and Use of Time Survey

Using a well-tested time-at-work questionnaire (Schriber & Gutek 1987) which is shown in Table 1, we surveyed the GLOBALIS employees who were engaged in the four projects we tracked. Respondents indicated their level of agreement on a five point Likert scale (Likert 1932) with each of the twenty-five statements based on their perceptions of what was most typical within their organization.

<table>
<thead>
<tr>
<th>Temporal Dimensions</th>
<th>What are we assessing?</th>
<th>Questions (® - Scores are reversed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schedule &amp; Deadlines</td>
<td>Importance of meeting deadlines and staying on schedule</td>
<td>1. People here feel that deadlines don’t really matter ®</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Staying on schedule is important here</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. It is important to meet our deadlines</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. We don’t pay much attention to schedules ®</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. No one gets upset when you miss a deadline ®</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6. All our work is tightly scheduled</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7. It is important to be “on time” for everything</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8. People do most of their work under deadlines</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9. People do things when they are ready, not on a schedule ®</td>
</tr>
<tr>
<td>Punctuality</td>
<td>Importance of arriving for work on time</td>
<td>1. People get upset when you are late for work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. People don’t care what time you arrive for work®</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. No one cares if you are late returning from a meal break ®</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. If people arrive an hour late for work, they will feel “rushed” all day</td>
</tr>
<tr>
<td>Time Boundaries</td>
<td>How far works interferes with personal time?</td>
<td>1. People usually expect to take their work home with them</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. People expect to leave at the end of the day without worrying about their work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. People rarely get work-related calls during &quot;off&quot; hours (like nights and weekends) ®</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. When people go on vacation, they are expected to tell their boss how to reach them</td>
</tr>
<tr>
<td>Awareness of time use</td>
<td>To what extent time use is planned?</td>
<td>1. Most people don’t think about how they use their time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. People worry about using their time well ®</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. People here plan their time carefully ®</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. People expect you to know how long it will take you to do something ®</td>
</tr>
</tbody>
</table>

Table 1-Time-at-Work Questionnaire

Figure 1 on the next page shows the mean values for each of the four temporal constructs which are aggregated for the whole company and by each country. On the schedule and deadline dimension the mean values for US, Ireland and India are 4.11, 4.26 and 4.13. This indicates that, irrespective of the differences in the national cultures, GLOBALIS workers in the US, Ireland, and India all proclaim the same orientation to schedules and deadlines. Similarly, the mean values on punctuality for the US, Ireland and India are almost identical, thus indicating that punctuality at GLOBALIS is also valued with the same uniformity across the global sites in these three countries.

The work time/non-work time boundary construct is a measure that attempts to assess how far work interferes with personal life. A mean value of 2.87 on the time boundaries dimension being less than 3.0 suggests that GLOBALIS workers perceive a good work/life balance at their work place. The mean value for both the US and India is 3.08, while the mean value for Ireland is 2.71. This indicates that work/life balance is perceived to be slightly better by Irish workers in comparison to those at the US and Indian sites.

The awareness of time use construct measures the extent to which workers plan and manage their use of time. A mean value of 2.38 suggests that the work environment of GLOBALIS is such that workers cannot always plan and control their use of work time. The mean values of 2.20, 2.39 and 2.45 for workers in the US, Ireland and India respectively would indicate that, irrespective of location, workers believe they have difficulty in planning and controlling their work time.
Figure 1 - Temporal Dimensions by Nationality

In summary, schedule and deadlines are uniformly valued and taken very seriously by GLOBALIS workers across geographic boundaries, punctuality is valued equally across these globally distributed worksites, and Irish workers report a stronger work/life balance than US and India workers. It should be mentioned that we had stratified the responses on project type, age, and job-level, and that data will be presented later when we will synthesize these diverse findings into a unified culture model.

2.4 Visits to GLOBALIS Sites in India, Ireland and US

Because these observations are varied we have identified each with a label.

2.4.1 Observation 1 - Punctuality

When we visited the sites in India, we experienced punctuality and an awareness of time. For instance, while the cross-cultural training on India observed that Indian culture has an “elastic” view of time, meetings in India always started promptly at the assigned time, unless of course we were the ones who were running late. In other words, there was no moment at which we were significantly delayed because of a lack of time sensitivity. We had the same experience at all the other sites of GLOBALIS.

2.4.2 Observation 2 - Age & Tenure

During our ten day visit to the GLOBALIS sites in India, we were immediately struck by how young the GLOBALIS employees in India were. We discovered that the average age of the Indian employees was in the mid-twenties and roughly 90% of the employees had less than 2 years of experience with the firm. These Indian workers had to interface with US employees who are typically 10 to 15 years older than them and who have over 10 years of experience with the systems, application and the firm. We even came across American GLOBALIS employees who had worked on the same project over the span of their career in the organization.

2.4.3 Observation 3 - Chit-chat

At the two Indian sites it was readily apparent that the workers engaged in a lot more chit-chat than what we saw at any of the sites we visited in the US. Even while working people talked over the booths, and the chit-chat was often embedded with work talk where someone would ask a questions and receive an answer from someone else. We estimated that if the work time / social time divide in the US is 80% to 20 %, work time/ social time divide in India is closer to 60% and 40%.
2.4.4 Observation 4-Attrition

GLOBALIS has been successful in recruiting experienced IT staff in India, with most having started their careers with an Indian IT firm (such as Infosys, TCS, Satyam, Wipro, etc.) with many moving to a US IT firm (e.g., IBM, Microsoft, Oracle, Motorola, Hewlett-Packard, etc.) in India as their second IT job. They were attracted to GLOBALIS in order to gain domain knowledge of financial systems and technologies. These people came to GLOBALIS expecting to work directly with the clients. In addition, we were told that the expectation in the IT market place is such that a new employee expects to move to project leadership position in two years, and if s/he is not sent to the US for training or system implementation within two years of joining the firm, then his or her career is perceived to be in a crisis. A lack in any of these expectations often leads to the worker to leave for career aspirations and opportunity with another firm, making attrition a real issue.

2.4.5 Observation 5- SLAs

A manager in US told us that he gets complaints from his staff about the poor quality of work done in India but when he talks to the manager in India he is presented with the service level agreement data that shows the Indian site is meeting the agreed SLAs. We saw an example of this dichotomy during our India visit. We observed how a customer support staff in Bangalore handled a customer call from a US customer. The support staff had real difficulty understanding the name of the customer and took at least 10 minutes to find the record of the customer. We could anticipate the customer was frustrated but the support staff continued repeating the script. Then suddenly the phone went dead and the support staff entered this as a satisfactory closed call. It is possible the customer was frustrated and may have decided to call back again hoping to catch another support staff, but our support person closed the call, indicating that the customer query was resolved.

3 UNPACKING CULTURE

Our time-at-work survey of GLOBALIS workers found that the attitudes towards punctuality and deadlines and schedules were uniform across US, Ireland and India. This finding is not aligned with the common perceptions of the attitude of Indians towards time. For example, observing the Indians who live in the US, Acharya (2008) states “Although many Indian Americans have achieved considerable professional success ... one thing we do not seem to learn is the concept of time. Except for activities related to one’s profession – where one has to be on time or lose a job – there is absolutely no regard for appointed time in other non-professional or social activities. We seem to think that all activities start on Indian Standard Time, and not on announced time.”

Acharya recognizes the casual attitude to punctuality in Indian culture, but he also points out that in the workplace this attitude changes. Within the cultural frame our time-and-work survey finding is consistent with Laurent’s (1986) argument that adherence to corporate norms can modify behaviors that may be embedded in national culture. Our attitude-toward-time survey suggests that any study of how culture impacts worker behaviors in global virtual work must include the analysis of corporate culture.

We surveyed the same 137 employees to assess GLOBALIS’ culture, using the instrument developed by Sarros et al (2002). It profiles corporate culture across seven dimensions: Competitiveness; Emphasis on Reward; Innovation; Performance Orientation; Social Responsibility; Stability and Supportiveness. As shown in Table 2, all seven dimensions received a mean score above 3.0, indicating that GLOBALIS employees report placing importance to all seven dimensions. The mean values for the US, Ireland and India individually are also greater than 3.0, and no significant statistical difference was observed in five of these seven culture dimensions. We can therefore conclude that GLOBALIS’ corporate values are uniform across its work sites.
### Table 2-Culture Profile Summary

The diffusion of GLOBALIS’ corporate culture across sites could be a reason why GLOBALIS employees indicated in our survey uniform valuing of the temporal dimensions of time across countries. This would suggest that GLOBALIS’ corporate norms are modifying the temporal attitudes that are embedded in national cultures. However, as observed by Scott and described earlier in the data gathering section that the actual behavior of the Indian software engineers who were added late to a project he was managing does not match the self-reported finding of the change in attitude toward temporal dimensions.

Furthermore, our chit-chat observation (observation) is indicating that the work time / social time divide behaviors at GLOBALIS sites are aligned to the observed national culture differences (Manrai and Manrai 1995). In addition, our interview data is showing (a) a high-context / low context communications styles exist at GLOBALIS and (b) differences in the professional cultures between the Ireland (more customer-oriented) and India (more process-oriented) sites.

Thus, an important research question is: how to reconcile these conflicting observations and findings associated with the study of workers behaviors in global work. Prior research (Huang et al. 2003, Straub et al. 2002, Myers & Tam 2002; Gallivan & Srite 2005) have recognized this conundrum as layers of cultures that impact behavior in global work and they have warned against using national

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Items</th>
<th>Mean US n=20</th>
<th>Ireland n=78</th>
<th>India n=39</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Performance</strong></td>
<td>1. High expectation for performance</td>
<td>3.9</td>
<td>3.96</td>
<td>3.96</td>
</tr>
<tr>
<td>Orientation</td>
<td>2. Enthusiasm for the job</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Being results oriented</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. Being highly organized</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Competitiveness</strong></td>
<td>1. Achievement orientation</td>
<td>3.81</td>
<td>4.01</td>
<td>3.79</td>
</tr>
<tr>
<td></td>
<td>2. Emphasis on quality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Being distinctive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Social</strong></td>
<td>1. Being reflective</td>
<td>3.76</td>
<td>3.43</td>
<td>3.82</td>
</tr>
<tr>
<td>Responsibility</td>
<td>2. Being socially responsible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Having a clear guiding philosophy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. Having a good reputation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stability</strong></td>
<td>1. Security of employment</td>
<td>3.71</td>
<td>3.27</td>
<td>3.79</td>
</tr>
<tr>
<td></td>
<td>2. Low conflict</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Being calm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Supportiveness</strong></td>
<td>1. Being team oriented</td>
<td>3.57</td>
<td>3.33</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>2. Sharing information freely</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Being people oriented</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. Collaboration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Emphasis on</strong></td>
<td>1. Fairness</td>
<td>3.55</td>
<td>3.06</td>
<td>3.62</td>
</tr>
<tr>
<td>Reward</td>
<td>2. Opportunity for professional growth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. High pay for performance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. Praise for performance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Innovation</strong></td>
<td>1. Quick to take advantage of opportunities</td>
<td>3.47</td>
<td>3.75</td>
<td>3.38</td>
</tr>
<tr>
<td></td>
<td>2. Risk taking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Taking individual responsibility</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
culture as a sole operationalization of culture to explain global workplace behavior. This paper will next unify this culture layers conundrum.

3.1 The Concept of Workplace Culture

The co-location of workers at each work site generates a workplace culture which is simply a shared understanding of norms, practices and values. Because the workers are co-located, behaviors and norms are visible to each other and there is more opportunity to engage in informal chit-chat leading to behavioral knowledge sharing. Thus, proximity and visibility enables a form of cohesion among the workers at a site that we are calling as the workplace culture. We will analyze GLOBALIS’ data further to shape and sharpen the workplace culture construct and demonstrate its role in analyzing behavior in global work.

3.1.1 Analysis of Chit-chat observation

When we visited the Bangalore site we noticed workers in groups of two or three talking in the canteen at all hours during the work day. During our visit to the Gurgaon site, we observed continual small talk among the adjacent workers during work hours as well. This behavior was markedly different from that which we observed at the US sites: there, we rarely saw workers chit-chatting during work hours. While undoubtedly some level of small talk occurred, it did not appear to be at the level we observed in India. Furthermore, in the US, the people whom we interviewed came to meet us at the assigned time and their body language indicated that the interview should end at the allotted time. Seemingly, the workers at the US sites went straight back to their cubicle or office right after the interview. We also noticed workers at the US sites walking to the coffee machine, mostly during the morning hours, and always carrying their cup of coffee to their workplace. In contrast, the workers in India, including American managers who were there on temporary assignment, did not seem to be in a hurry to return to their work after scheduled interviews.

These observations are consistent with what Brislin and Kim (2003) call the task and social time variance among cultures during the workday. Manrai and Manrai (1995) studied this phenomenon empirically and found that in big cities in the US, the task time/social time divide is 80:20, while in India, Nepal, Malaysia, and some Latin countries, it is often 50:50. We initially attributed this behavior to national culture differences. Later, at one of our visits to a New England site, we observed a set of Indian workers from the Gurgaon site (who were stationed there for three months in the US for training), behaving similarly to the US workers: they neither engaged in small talk while working nor chatted at the coffee machine at odd hours. This would indicate that, irrespective of their national culture, people readily adapt to the norms and behaviors of the workplace. This led us to further examine the concept and role of workplace culture.

Cramton and Webber (2005) report that the local contexts across which global virtual teams operate are important and they impact communication and coordination across sites. These studies remind us that greater emphasis needs to be placed on studying how culture and communication manifest themselves in specific contexts. Examples of local context in global work are: travel to and from the worksite, accessibility to the communications infrastructure, local holidays, language proficiency and accents, and cultural differences. Our task time / social time divide observation has identified the existence of a workplace culture that impacts workers’ behaviors, irrespective of their national culture orientations.

3.1.2 Analysis of Age and Experience observation

One US GLOBALIS manager told us that, “When I’m asked a question in a meeting, my modus operandi is to generate a conversation by asking the group what they think about the issue before giving an answer.” When this manager used this approach with his Indian subordinates in Bangalore, “no one spoke up.” Apparently, according to him, “this is common Indian worker behavior that US managers and workers encounter.” At the cross-cultural training course we attended, the US
employees reported that Indian workers “are meek and do not push back”. Other stereotype traits pointed out in this course were: “Indians can’t say ‘no’”, and “Indians have a laid-back attitude toward time”. In summary, both the US managers and workers at GLOBALIS have observed and/or encountered many of these behavior patterns in their interactions with Indian workers.

During our ten day visit to the GLOBALIS sites in India, we were able to investigate these issues directly. We observed that the average age of the employees in India is in the mid-twenties, with roughly 90% having less than two years’ experience with the firm. These Indian workers have to interface with US employees who are 10 to 15 years older and who typically have over ten years’ experience with the systems, applications, and the firm. It is not unusual to find American GLOBALIS employees who have worked on the same project throughout their career in the organization. The Indian workers at the Indian sites are not only young, but were recruited from technical IT firms and lacked the application domain knowledge in the finance sector. They behaved like typical new employees. Thus, many of the behaviors attributed to national culture traits may just as easily be attributed to age and experience. For example, as discussed earlier Scott observed that the new Indian workers added to his project were not as prompt and responsive as the original US and Irish workers. When we talked to these Indian workers we learned that GLOBALIS has many databases with a complex system of access rules and the Indian workers, being new to the project and lacking domain knowledge, were not given direct access to these databases. As a consequence they had to discover who had access and then request them to access and transmit those files. This took time and gave the impression to Scott that these Indian workers were not prompt and responsive.

3.1.3 Observation regarding Process versus Relationship Behavior

In our data section we reported earlier that Seamus, a manager at an Ireland site, told us that when he asked the Indian software engineers to start coding from a very general set of requirements, they refused until the requirements were formally specified. Seamus complained that the Indians were being too process-oriented and they did not understand the customer orientation of GLOBALIS.

Because the software development industry in India is invested heavily in capability maturity certifications and people change jobs often, the IT work culture in India is predominantly process-focused. Since GLOBALIS’ recruiting strategy in India is to recruit only experienced people, its workers at the Indian sites arrive after having socialized in process-based, matured software organizations. Thus, GLOBALIS’ recruiting strategy in India and the professional software culture of the Indian IT industry is responsible for the process-focused software professional culture (Cougar et al. 1990) of GLOBALIS in India.

3.1.4 Analysis of video-conferencing observations

Our video conference observations brought out that GLOBALIS sites are not all equal. The sites in the New England region are in the core and Ireland and India sites are clearly on the periphery in terms of decision making. This was underscored by all the project managers being stationed in the US. While some global teams are currently managed by personnel located in Ireland, India does not currently “own” any projects. Thus, the very organizational distribution of work, while speaking of integrated teams conveys the message of hierarchical stratification. This perspective was reinforced in the training sessions being taken by the Americans (who were being trained to manage relationships), and the Indians (who were being trained to do the work). The video-conference interactions brought out the Indian managers at the Indian sites of GLOBALIS, even though they worked for the same company, viewed their American counterparts as their clients or bosses. This was underlined by their previous work experience. Since the vendor model dominates the Indian IT landscape, it is not surprising to see this perspective being adopted by GLOBALIS’ Indian employees. Thus, the Indian workers were struggling to see themselves as team members, a struggle that was enhanced by the age and experience differential between the Indian and US offices. One could say that the workplace culture of the Indian sites is vendor-oriented, and in a vendor-oriented culture, it is difficult to push
back or say ‘no’ directly, especially when being trained to provide a service. This may be a reason why the workers at the Indian sites use a high-context communications style.

Bringing these observations together, we see that certain cultural attributes are related more to the characteristics of the worksites than due to national or corporate culture. They substantiate the existence of workplace culture. The workplace cultures we observed at GLOBALIS can be summarized as follows:

- The work culture at the Indian sites of GLOBALIS is process-based
- GLOBALIS’ work culture in Ireland is more customer-oriented
- Because GLOBALIS employees at the Indian sites view their US and Irish colleagues as clients, GLOBALIS’ Indian sites culture is vendor-oriented
- Although the task time / social time divide at GLOBALIS match national culture traits, it is part of the workplace culture because of the unconscious adaptation of local practice by visitors from other sites.

As discussed above, age and experience, professional standards and certifications, and client/vendor orientation all combine to impact workplace culture.

3.2 Towards a new model of culture in global organizations

Our investigations of attitude and use of time at GLOBALIS suggest that, to understand the role of culture in today’s global work, the national culture construction of culture is insufficient. Rather, corporate and workplace cultures need to be incorporated into the mix. This is consistent with the findings of Perlow and Weeks (2002) and Huang et al. (2003). Perlow and Weeks noted that behaviors of software engineers doing the same type of work in the US and India reflect the combined influence of national, occupational, and organizational layers of culture in the two settings. Huang and colleagues noted different organizational subcultures at work in their study. We integrate these concepts and offer a new culture model of global work in multinational firms such as GLOBALIS, shown in Figure 2. Workplace culture is a key construct and, as argued above, it depends upon a variety of environmental factors, including age and experience of the workers, industry standards and certifications, and differing client/vendor relationship perceptions at different sites.

The addition of project culture in the model is based on our survey data. The responses were stratified on following variables.

- Nationality - US, Ireland, and India
- Project - Bank of America, FMD, PACE, and STARS
- Position - non-management, lower management, middle management, top management
- Age - 20-29, 30-39, 40-49, 50-59

Figure 3 presents the survey participants’ rating based on the four projects. The variation in the mean values (4.59, 4.09 and 3.75) for the four projects in the schedule and deadline dimension in Figure 3 implies that meeting schedule and deadlines are not only uniformly valued across sites, the level of importance assigned to schedule and deadlines depends of on the nature of the project. This suggests that project characteristics impact workers behaviors.
4 IMPLICATIONS AND DISCUSSION

Although culture investigations in global IT work have been an important area of research, the focus often has been on attributing behaviors to national culture differences (Ramingwong and Sajeev 2007, Symons and Stenzel 2007, Krishna et al. 2004, Borchers 2003, Walsham 2002, Heeks et al. 2001). We will review some of the findings of earlier research using the lens of our proposed culture model in global work.

Using Hofstede’s dimensions of national culture, Ramingwong and Sajeev (2007) conclude that the risk of the ‘mum effect’ (or code of silence) is higher in Asia (higher power-distance, lower individualism-collectivism index and higher long-term orientation) than in the West. Our time-use investigations showed that people tend to comply with organizational norms and values, and therefore,
worker behavior would be driven by visible actions of the organization (such as the firing of a whistle blower) rather than workers’ inherent proclivity to the mum effect. Similarly, since the client/vendor relationship impacts workplace behavior, the ‘mum effect’ will also be moderated by the culture of the workplace. In summary, our culture model of global work suggests that an analysis of the ‘mum effect’ must take into account the impacts of corporate and workplace cultures.

After a chance reading of Hofstede’s (1980) book *Culture’s Consequences*, Borchers (2003) began to see how Hofstede’s culture dimensions were playing out in the behaviors among team members in the US, India and Japan. For instance, Borchers reasoned that the difference in behaviors between the US and Indian workers during a ‘bug fixing’ phase on a project was due to Indian culture having a lower value on Hofstede’s individualism-collectiveness index in comparison to the US, and that this enabled the Indian developers to form close interpersonal bonds within the team.

There are, of course, other potential causes for the behavior of the Indian team, which are not considered by Borchers, and do not necessarily relate to Hofstede’s analysis. During our investigations of global work at GLOBALIS, we noticed that the Indian software developers working for multinational firms belonged to a narrow band of Indian culture. Most of them were educated in English-speaking high schools and grew up listening to Western popular music, popular literature and movies. It is this identity with Western pop culture and their youthfulness that can be seen to separate them from Indian society in general and to seed the kind of bonds of friendship that Borchers observed in the Indian software team members. This high degree of social solidarity, often cemented through shared social activities, may be what is fueling their collaborative work rather than simply belonging to “Indian culture.”

Scholars on global work who have explored research-driven best practices for effective management of global software teams (Krishna et al. 2004, Walsham 2002) frame cultural barriers to global coordination on societies’ distinct ways of working. As a result, they attribute observed behaviors in the workplace to national culture differences alone. For example, British managers in an outsourcing relationship with a particular Indian software supplier noted that Indian programmers would not voice criticism in face-to-face meetings but would sometimes send their opinions in email messages after the meetings had disbanded. This behavior is attributed to Indian culture where ‘deference to authority’ is valued.

Again, there are alternative explanations that can be posited upon closer examination, utilizing different analytical lenses. For example, we observed that a number of the Indians we interviewed could not converse fluently in English. Borchers (2003) observed that because of certain Japanese workers’ inability to communicate fluently in English, they preferred to communicate through emails. Thus, a possible reason for some of the Indian workers not voicing criticism in face-to-face interaction could be their inability to effectively converse in English. Furthermore, these Indian employees often had far less experience on the projects, had less time in the company, had little or no domain knowledge, and were interacting with American team members who had been involved in the project from the outset. All of these factors may well have contributed to a ‘following’ versus ‘leading’ approach to team interactions.

Heeks et al. (2001) offer an interesting example of cultural dissonance. Apparently project leakage in a UK firm Sierra, called UK-Sierra here, was typically 2 to 3 percent and in its wholly-owned offshore development center in India (India-Sierra), it was between 25 to 30 percent. Conflict arose between UK-Sierra and India-Sierra regarding definitions of leakage. At UK-Sierra, leakage was measured in terms of time spent versus time planned. They reported that “from the Indian staff’s perspective, deadlines mattered but how much time was spent did not, and they felt UK values were insensitive to Indian conditions.” Because the authors analyzed this issue based on their view that “players in this global game still retain cultural values rooted in a particular locale,” (ibid., p. 59), Heeks and colleagues framed the conflict in terms of differences in cultural values between UK and India.

In our study of attitude towards time and its use at GLOBALIS sites in the US, Ireland and India, we saw no evidence to suggest that ‘meeting deadlines’ is a particularly Indian cultural value. Rather, it
was a corporate value of GLOBALIS with which the Indians complied. Also, we observed that local working conditions (such as lack of domain knowledge or of tacit knowledge) cause delays and interfere with the promptness of Indian workers. It is hard to say that the definition of leakage in UK-Sierra is based on certain British values, but it is clear that measures used in one corporate environment may not transfer across geographic boundaries, especially when they make workers look bad. Thus it may not be appropriate, certainly not in all cases, to attribute national culture differences to be the cause of dissonance.

Irrespective of their worksite location, GLOBALIS workers felt that national culture is not an overriding factor in distributed workplace collaboration. However, GLOBALIS managers in the US, Ireland and India all reported that national culture differences matter greatly. Besides the contradictory perceptions between workers and managers regarding the role of culture in global work, we also observed uneven management actions with respect to culture training at GLOBALIS. For example, GLOBALIS’ management contracted with a consulting company to train their US employees on how to do business with India. This training consisted of a one-day program where managers were exposed to generalized traits and characteristics of Indian culture while developing an awareness of their own US culture. We attended this training program and found that, like most cross-cultural training programs, the focus was on exposing the American workers to those cultural traits of Indians that were said to be different from Americans. Attitude and use of time was one of the topics covered. It is perhaps instructive to note that the Irish and Indian managers were not given a similar cross-cultural training opportunity. Our interviews with the Indian workers indicated that some had been given a short, three hour introduction to working with US colleagues, but the people we interviewed could not remember much about this training other than they should not use the word “but” in their email communications with US workers and managers.

During our monitoring of global conference calls during our study we observed that any time someone joined a conference call five minutes or so late, they apologized. Since this appeared to be so irrespective of the location of the participant, we concluded that, in the culture of GLOBALIS, five minutes of tardiness for a conference call is considered significant and requires an apology. As an aside, when we took a class of American students to Bangalore, India, the five Indian industry speakers we had invited to speak to the students were all well over half an hour late for their lectures, yet none of them apologized to the class. Does this suggest that in Indian culture being tardy to a formal event by half an hour is not considered significantly late, but that Indians in GLOBALIS have been acculturated to the company culture? In any case, these observations tend to confirm Edward Hall’s (1959) finding that punctuality is cultural at some level, and Laurent’s (1986) argument that adherence to corporate norms may modify behaviors embedded in national culture.

The purpose of these examples has been to illustrate that the current conceptualization of cross-cultural analyses based on the traditional view arising from Hofstede, namely that workers in global work retain cultural values rooted in national culture, can result in inappropriate attributions (Osland & Bird 2000). Myers and Tan (2002) found wide usage of ‘national culture’ models by many IS researchers, resulting in narrow, simplified perspectives on the effects of national culture. Their review of the IS literature on national culture shows how the concept of national culture was taken for granted by many IS researchers. They criticized the conceptualization of national culture models proposed by Hofstede and other researchers. They proposed a research agenda inviting researchers to develop better lenses through which to move beyond simplistic treatments of national culture in order to investigate the dynamic and complex effects of cross-cultural diversity for the management of today’s global organizations. This study has been an attempt to move the field in the direction proposed by Myers and Tan (2002) and Gallivan and Srite (2005). Our model of culture provides a more nuanced and refined framework to study the impact of cross-cultural differences on global work. Next we will discuss the research directions needed to place our model on a firmer foundation and the research approach needed to move the field forward.

First, most empirical studies in the IS and Software Engineering literatures on global work and global virtual teams have tended to focus almost entirely on national culture differences, primarily on two of
Hofstede’s culture dimensions, namely Power Distance and Individualism/Collectivism. Trompenaars and Hampden-Turner (1997) proposed a set of seven culture dimensions and two among these are same as Hofstede’s. More recently, Gupta and House (2004) proposed nine dimensions of culture which, unlike Hofstede’s, are rooted in theory and they are: Performance Orientation; Assertiveness Orientation; Future Orientation; Human Orientation; Institutional Collectivism; Family Collectivism; Gender Egalitarianism; Power Distance, and Uncertainty Avoidance. These provide a more finely grained approach that is worthy of further consideration. Thus, we argue that there is a need to develop a richer understanding of how these culture dimensions affect global interactions and work.

Second, as argued by Huang and colleagues (2003), the role of corporate culture, and corporate subcultures, in global work requires much more serious consideration. Although there is extensive literature on corporate culture and adoption and use IT in organizations (Gallivan and Srite 2005), the learning from these studies have not been incorporated in the study of global virtual work.

Third, the construct of workplace culture is crucial for a more comprehensive understanding of the factors that impact global work, yet there is little research on this construct in the IS literature. Since our conceptualization of workplace culture is based on the study of just one multi-national firm, the contextual variables we have identified need to be validated in other settings.

Finally, one cannot just engage in traditional survey research alone to build theory (Galliers 1992); a pluralistic approach is required (Mingers 2001). We would further argue that, for the theory to be robust, it must be rooted in practice. Because global work is a relatively new phenomenon, an ethnographic study is needed to uncover the factors that may be affecting behaviors in global virtual work with the findings being used to develop hypotheses for theory testing.

References


COPING WITH CULTURAL AND MATURITY INEQUALITY IN OFFSHORE OUTSOURCING: IS MINIMIZING INTERACTION THE SOLUTION?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0157.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Culture fit / differences / heritage/ intelligence / issues / theory / values, Offshoring / Outsourcing, Global IT management, Global software development</td>
</tr>
</tbody>
</table>
COPING WITH CULTURAL AND MATURITY INEQUALITY IN OFFSHORE OUTSOURCING: IS MINIMIZING INTERACTION THE SOLUTION?

Hertzum, Morten, Roskilde University, Universitetsvej 1, Building 43.2, 4000 Roskilde, Denmark, mhz@ruc.dk

Pries-Heje, Jan, Roskilde University, Universitetsvej 1, Building 43.2, 4000 Roskilde, Denmark, janph@ruc.dk

Abstract

Many companies consider and undertake outsourcing of their software-development activities. Often information systems development is outsourced to vendors in different cultures or with a different level of software-process maturity. Recommendations for managing such outsourcing arrangements typically involve that client and vendor should increase interaction, learn more about the culture of the other part, communicate more, form partnerships, and the like. We have studied a client that did the opposite with a successful outcome. Based on a case study we lay out the story of how interaction between client and vendor on purpose was minimised. What mechanisms were used? What worked and what did not? We conclude that minimizing interaction can be a viable strategy to follow when clients face large cultural and maturity inequality in offshoring their software-development activities.

Keywords: Offshore outsourcing, Culture, Maturity, Minimal-interaction strategy, Extra costs.
1 INTRODUCTION

Information systems (IS) outsourcing can be defined as the practice where an organization purchase goods or services “that were previously provided internally” (Lacity et al. 1993). Many companies consider and undertake outsourcing. It has been estimated that global software outsourcing is going to reach $260 billion by 2009 (Vitharana et al. 2007). The documented benefits of IS outsourcing have been reduction of time and cost, increased quality, improved business performance, and increased ability to concentrate on the core of the business (McFarlane et al. 2004). Recently, outsourcing to another country – so-called offshoring – has become popular because large numbers of technically skilled developers are being educated in countries such as India where wages are low compared to Western Europe and North America, in which skilled developers are in short supply. Offshore outsourcing is defined as “the subcontracting of an activity by a client organization to an independent service provider working from an overseas destination” (Vlaar et al. 2008) or simply as “inter-country outsourcing” (King et al. 2008).

While geographic distance is, thus, a defining characteristic of offshore outsourcing, the challenge is not geography as such but overcoming communication bottlenecks, knowledge asymmetries, psychological dissociation, and socio-cultural differences. General recommendations to manage collaboration at a distance include establishing common ground at the outset and distributing tasks such that only loosely coupled tasks are allocated to different sites (Olson et al. 2000). For offshore outsourcing two specific challenges are that the client and the vendor often have very different cultures and are at very different levels of maturity with respect to their software development processes. To handle these challenges it is frequently recommend to increase interaction, learn more about the culture of the other part, communicate more, form partnerships, or the like (Bhat et al. 2006; Hendry 1995; Krishna et al. 2004).

In this study, we analyse a company that did the opposite, namely minimized interaction between client and vendor, and with a successful outcome, in the sense that they just decided to renew their contract with a large Indian vendor; on the same terms as for the last four years. This approach is contrary to the prevalent recommendations in the literature, and we consider it of interest to study this anomaly. In doing so our research question is: Is minimal interaction between client and vendor a way to overcome cultural and maturity inequality in offshore outsourcing?

The next section sets the context for our study by describing related work on offshore outsourcing, culture and maturity. Section 3 accounts for the method we used in our empirical study, and Section 4 introduces the client and vendor organizations. Section 5 analyses the client’s minimal interaction approach to offshoring and identifies the mechanisms established to succeed with this approach. Section 6 discusses how the minimal interaction approach has affected the client’s software-development activities and at what costs.

2 OFFSHORE OUTSOURCING

Carmel and Agarwal (2002) propose four stages of offshore sourcing of information systems development. An organization is at the first stage if it is still an offshore bystander. While there may be a variety of reasons for remaining at this stage, more and more organizations choose to proceed to one of the subsequent stages. Organizations at the second stage have started experimenting with offshoring, for example through pilot projects, and often their motivation for offshoring is unavailability of onshore developers rather than a proactive focus on offshore possibilities. Loaded wages for skilled Indian developers at 30-50% of onshore wages in Western Europe and North America is an important motivator for offshoring, but rarely realized at this stage. One reason for this is that the stage is transitional; when some experience has been gained and cost savings start to occur, organizations move to the next stage. At the third stage organizations are characterized by a proactive cost focus. A typical recommendation at this stage is to restrict offshoring to non-core and structured
tasks, such as construction based on detailed specifications (e.g., Cusick et al. 2006). Often, onshore project managers receive targets specifying that a certain percentage of the developer hours on their projects should be offshore developer hours. At least one study finds that onshore staff tended to perceive offshore developers as cheap worker-bees who could be ordered around (Levina et al. 2008).

At the fourth stage organizations no longer view offshoring as simply a source of low cost work but have adopted a proactive strategic focus. The strategic objectives pursued at this stage include access to new markets and offshoring of entire projects from requirements through to support, also of projects involving innovation and new product development. The best offshore practices described by Bhat et al. (2006) appear to be directed solely at this fourth stage and focus on achieving shared goals, culture, processes, and responsibilities for client and vendor.

While the vast majority of research approaches offshoring from the client’s point of view, some studies do investigate vendors’ views of offshoring (e.g., Bhat et al. 2006; Oza et al. 2005). A theme common to client and vendor research is asymmetries in knowledge and experience. These asymmetries concern, among other things, the business domain, typically with the client in possession of business knowledge and the vendor less so (e.g., Levina et al. 2008), and development processes, typically with vendors that have more structured development processes than clients (e.g., Oza et al. 2005). Technical knowledge exhibits another type of asymmetry in that the vendor typically employs a large pool of technically skilled developers while this resource is scarcer onshore.

The top risks associated with offshore outsourcing include lack of top management commitment, miscommunication of requirements, inadequate user involvement, failure to manage end user expectations, and poor change control (Iacovou et al. 2008). These risks do not appear to be specific to offshoring but rather to apply to information systems development in general (cf. Schmidt et al. 2001). Some of the top risks associated with offshoring are, however, specific to offshoring, including language barriers, lack of offshore project management know-how by the client, lack of technical or business know-how by the offshore team, and failure to consider all costs (Iacovou et al. 2008). Many client organizations have expected cost reductions from their offshoring arrangements due to the lower offshore wages but have not fully understood all the costs involved in outsourcing (Barthélemy 2001). Dibbern et al. (2008) identify specification costs, design costs, knowledge-transfer costs, coordination costs, and control costs as the five main categories in which clients face extra costs when projects are offshore. The five categories of extra costs relate to the less effective possibilities for communication between client and vendor and the resulting degradation in their mutual awareness of each other’s work and day-to-day activities.

2.1 The role of culture

Hofstede (2001) defines culture as “the collective programming of the mind that distinguishes the members of one group or category of people from another”. Hofstede’s work shows that even within a single organization, national groups of employee exhibit different cultural characteristics. These characteristics have been specified in terms of five cultural dimensions: power distance, uncertainty avoidance, individualism/collectivism, masculinity/femininity, and long-/short-term orientation (Hofstede 2001). It appears that managers in organizations chronically underestimate the magnitude and importance of cultural differences (Hofstede et al. 2005).

Prior surveys indicate that national culture is a leading cause of problems in the offshoring of services (Metters 2008). Metters (2008), for example, refers to a survey where 60 executives involved in offshoring information technology (IT) services cited “cultural differences” as the most important problem in relation to offshoring. Also, Terdiman and Berg’s (2001) framework for evaluating a potential offshoring country has “cultural issues” as one of three main areas. The interest in nearshoring is another indication that similar cultural characteristics, such as ways of doing business, are considered important to outsourcing decisions (Carmel et al. 2007). In contrast to nearshoring, offshoring typically implies profound cultural differences between client and vendor. Never the less one can find countries that may be remote but culturally close such as Great Britain and Australia.
In the concrete Denmark and India (the client and vendor countries of this study) differ along several of Hofstede’s cultural dimensions but in particular with respect to power distance – defined as the extent to which the less powerful members of institutions and organizations expect and accept that power is distributed unequally. In India power distance is very high, in Denmark very low (Hofstede 2001). Consequently, in the Danish business culture, rank and title are less important than in India where hierarchical forms of behaviour are expected. In Denmark subordinates are expected to speak up and offer suggestions; in India superiors and seniors enjoy more respect, and decisions tend to be top-down. This affects, for example, communication styles and ownership of results (Schomer 2006).

Recommendations for handling cultural differences in offshoring arrangements include facilitated communication sessions (Dubé et al. 2001), building consensus on norms for meetings and deadlines (Paré et al. 1999), and other efforts to establish a shared culture (Bhat et al. 2006).

2.2 Maturity and software process improvement

Maturity models are used to improve the performance of organisations, processes, technology, and people. The Capability Maturity Model (CMM) is a framework characterizing a five-step path for software process improvement (Paulk 1995). The path describes key processes and goals at each of the five levels. An organization has to meet the goals at one level to reach the next. For example, to go from the basic level 1 where behaviour is characterized by being ad-hoc and intuitive to level 2, you need to achieve the goals incorporated in six key process areas: requirements management, subcontractor management, project planning, project tracking, quality assurance, and configuration management. CMM became so popular that a large number of other models using the same five-step path were invented, including People-CMM, Integrated Product Development CMM, Systems Acquisition CMM, and Testing Maturity Model. Finally a large number of the CMM models were summoned in CMM-integrated – or just CMMI (Ahern et al. 2003; Chrissis et al. 2003).

In relation to offshoring, it is noteworthy that India has embraced CMMI. Four countries in the world have used the CMMI model extensively: Australia, India, Japan, and the US (IndiaExpressGroup 2003). The highest level of maturity is level 5, and a few years ago 80% of all the companies in the world at level 5 were from India (Mohnot 2003).

In Denmark only one or two companies have reached level 5 (cf. Pries-Heje et al. 2008), and the majority of Danish companies are at level 1. Thus, when Indian and Danish organizations enter into offshoring arrangements there may be huge maturity inequalities between them.

3 METHOD

Our empirical study is a case study based on interviews in one Danish organization. The case study is single-case and embedded (type 2) according to the typology by Yin (1994, p. 39). We have not obtained data from the vendor. Thus, the empirical data are restricted to a client-side perspective on offshoring. One of the authors has worked with the organization since 2003, and has carried out several assessments and training sessions in the organization from 2003-2007. We believe that it is fair to claim that this author has extensive knowledge on how software development is carried out in the organization.

Concerning offshore outsourcing, however, the case study reported here took place in 2008 and was carried out by both authors. We conducted an initial interview with three staff involved in the client’s offshoring at the managerial level. During this interview we got an overview of the client’s offshore-outsourcing history and identified seven persons for in-depth interviews. The interviewees comprised persons involved in or responsible for (1) the start-up of the offshoring activities, (2) the entire course of offshoring activities, (3) the offshoring contract, (4) the offshore development centre, (5) concrete offshoring projects and certification of offshore staff, and (6) improvement of the client’s development processes; and (7) an offshore coordinator recently returned from a long-term placement at the vendor.
The seven in-depth interviews were loosely structured by an interview guide addressing:
- The offshoring arrangement between client and vendor, and its evolution
- Client-vendor interactions at the levels of the offshore agreement, projects, and individual staff
- The creation of a project identity in projects involving offshoring
- The coordination of such projects
- Initiatives undertaken to facilitate offshoring and the lessons learned from them
- Issues relating to differences in the cultural background of onshore and offshore staff

In addition, the interviewees were asked to reflect upon the factors critical to the client’s experience with offshoring. This part of the interviews was based on a walkthrough of Iacovou and Nakatsu’s (2008) ten-item list of top offshoring risk factors.

The interviews were conducted in meeting rooms at the client’s premises, except one interview which for practical reasons was conducted at the authors’ university. The initial interview was documented in written notes; the in-depth interviews were audio-recorded, and subsequently an extensive written record of the main points was produced. The written record included selected quotes, but the interviews, which lasted 1-2 hours, were not transcribed verbatim. The interviews were analyzed by reading through the written records several times, noting issues stated in individual interviews and patterns emerging across interviews. These issues and patterns were then grouped into themes, resulting in the analysis presented in this paper.

4 THE EMPIRICAL SETTING

The client is a Danish organization in the financial sector. It has approximately 850 employees, some 450 of which are directly involved in the development of IT systems. The client has for 40 years developed and hosted services for the Danish banks, particularly with regard to payment solutions. The financial sector is characterized by high volumes of safety-critical transactions and, thereby, a need for efficient and secure systems. Moreover, the financial sector is dynamic with changes in national and, increasingly, international legislation forcing revisions of systems, with mergers and acquisitions among banks necessitating integration or redesign of systems, and with considerable competition among providers of financial services creating a continual pressure for the development of new services.

After an early, unsuccessful attempt with outsourcing in the late 1980s, the client refrained from further attempts during the next decade. In 2000 the client started offshoring to India, and in 2002 they started working with their current vendor. The vendor is an Indian software-development organization, which employs over 8000 software developers and has years of experience as an offshore-outsourcing vendor of financial and other services. While the vendor has been certified at CMM level five since 2002 and CMMI since 2006, the interviewees estimate that the client is at CMM level 1 or 2. The collaboration between the client and the vendor has been going on for six years, and it has been decided to renew the current contract when it expires by the end of 2008.

The client’s rationale for entering into an offshoring relationship was to increase its capacity. This is stated by several interviewees, who also state that thanks to this increase in capacity the client has been able to carry through projects it would otherwise have been unable to take on.

5 OFFSHORE OUTSOURCING WITH MINIMAL INTERACTION

When setting up an offshore arrangement some interaction is required to negotiate the terms of cooperation, write a contract, and start working together (Willcocks et al. 2006). In the phase following – the operating phase it has been called (Cullen et al. 2006) – there also needs to be some interaction; the salient question is: how much? At one end, we can talk about minimal interaction. That is, just enough to make things work. One could say that minimal interaction is about paying as little transaction cost (Williamson 1979) as possible. Minimal interaction also entails that as few changes as
possible are made in the client’s and vendor’s internal processes. It should be noted that when one reduces transaction costs the remaining interaction will look as being more intensive at particular contact points. At the other end, we can maximize interaction trying to come as close together as possible. This may involve more communication, more learning about the culture of the other part, trying to level or balance maturity, forming a partnership, and maybe even blur the distinction between a client and a vendor.

5.1 Keeping distance

An illustrative example of minimal interaction is project A in the Danish organisation we are studying. This was the first project the client offshored to the vendor. The project, which lasted three years, consisted of converting an existing system to another platform. That is, the existing system in itself comprised a complete and, by definition, fully accurate specification. Such a task involves little analysis and design compared to programming. This characteristic of project A was the main reason it was chosen for offshoring, and it implied that the client could specify the project very accurately and very easily. This made the project very suited for the client’s minimal interaction strategy because minimal interaction could be attained at low risk.

Also, project A was only economically feasible for the client if it could be offshored. The project showed that the vendor had the technical knowledge required to make the conversion. Very few errors were detected during testing, and some of them turned out to be errors in the “specification”; that is, hitherto unnoticed errors in the old system.

After having completed project A, it was decided to set up a more permanent relationship between the client in Denmark and the vendor in India. It was at this point that the idea of minimized interaction really came into play. A manager says: “The point of departure is that they are vendors. They are not employees. They are a vendor like an external company we cooperate with. The idea was to establish it out there [i.e., at the vendor], so that they can maintain their culture and keep working the way they are used to; and people here [i.e., at the client] work in their way. Actually, reducing the need for intercultural interaction to as little as possible was part of what I was trying to accomplish.”

5.2 Exchanging people

The client has made use of two mechanisms for exchanging people to accomplish offshoring projects while maintaining minimal interaction. Both mechanisms involve intensive interaction but for selected people and selected periods of time. First, offshore developers have been on placements at the client to work with their onshore colleagues. This mimics how new onshore IT developers acquire business knowledge, but in addition to improving the offshore developers’ business knowledge it has also facilitated the general relationship between onshore and offshore staff. However, the placements require that onshore staff has the necessary time for communicating and interacting with the offshore developers; and the placements temporarily cancel the economical effect of offshoring because the offshore developers get onshore wages while they are onshore.

Second, the client has placed an offshore coordinator at the vendor. The few onshore employees who have had this position have been at the vendor on long-term placements. The offshore coordinator has a mediating role involving frequent phone contact with client staff, with whom they are well connected, and participation in project meetings with vendor staff. Collaboration between the offshore coordinator and vendor staff is face-to-face, thus avoiding the limitations of communication and collaboration at a distance and providing more opportunities for becoming aware of cultural and maturity issues in need of attention. Periodic onshore visits have been necessary for the offshore coordinators to maintain their network among the client staff. Moreover, the vendor may occasionally feel that the presence of the offshore coordinators transgresses the client-vendor boundary.

Apart from these two mechanisms, an effort has been made to motivate offshore developers to work for the client for a longer period of time. In the Indian offshore-outsourcing industry it is customary
that IT developers move into the management ranks after only a couple of years as developers. This is very different from the career path of Danish developers, who often work a decade or more as developers in the same business area. This cultural difference threatens the client’s minimal interaction strategy, because the continuous renewal of offshore developers implies that most of them will have insufficient business knowledge. The client has therefore aimed to make their relationship with the vendor sufficiently interesting for offshore developers to make it attractive for them to stay for a long period of time. The onshore placements of offshore developers have been effective in this regard.

5.3 **Exchanging knowledge**

The client has set up business courses at the vendor. The courses have been run by visiting onshore staff and by some of the offshore developers that have been on onshore placements. In some areas of the client’s business, the courses form an entire certification program, which ensures that offshore developers have a basic understanding of the business area for which they develop systems. While the offshore developers are not at a level of business understanding comparable to the onshore staff, their improving business understanding increases their ability to work autonomously and decreases the amount of interaction they need to have with the client.

The courses and certification programs are an attempt to exchange knowledge in concentrated packages and to several offshore developers at a time. This is considered preferable to frequent ad hoc interactions, which are complicated by the geographical separation. Extensive ad hoc interaction is also seen as time consuming, especially to the client, and therefore as contrary to the intention of shifting work from the client to the vendor. A manifestation of this is that a single point of entry has been enforced when offshore staff needs to communicate with onshore quality-assessment staff. This has been decided to protect the majority of the onshore staff from becoming engaged in too many, time-consuming communications. In this case it appear that the client has been more concerned with not making offshoring unpopular among its onshore staff than with providing the offshore staff with access to needed knowledge and the opportunity to gradually build a network.

Restricting access to needed knowledge in order to minimize interaction creates problems because it prolongs the period during which business knowledge is unevenly distributed between client and vendor. As an example, the present assessment of project B – a large, ongoing offshore project – is that it has been hard to strike the proper balance between technical and business development. The project requires both technical and business knowledge, but the uneven distribution of knowledge entails that the vendor, which is involved in the project with a massive 300-400 person years, often has only the technical knowledge. In working on project B, the vendor proceeds on the basis of its technical knowledge and remains unaware of some of the issues that might warrant business considerations. The client is at too great a distance from the vendor’s work to spot such issues and has not been able to specify them up front. As a consequence opportunities for business considerations are unintentionally bypassed, and project B becomes to an excessive extent about technically reprogramming a system.

5.4 **Developing software in two places with minimized interaction**

Today the client uses the vendor in India on a regular basis. Project B provides an interesting example. This project consists of converting a standalone system into a service available to many systems. The client considers such moves toward a more service-oriented architecture crucial to enable reuse across systems and to enable flexible assignment of the development of individual services to onshore or offshore groups with the ability and capacity to take them on. A main reason for completely reprogramming the system is, however, that the existing system has evolved over a long period, and due to extensive changes in the staff working on the system nobody any longer has a comprehensive overview of the programming code. Moreover, the documentation is not trusted to be up-to-date. Thus, it has become exceedingly difficult and costly to make revisions of the system (cf. Naur 1985). The capacity and lower price of the offshore vendor compared to onshore developers make it feasible to solve these difficulties by reprogramming the system from scratch.
However, turning a system into a service is not merely a technical task but also requires considerable business understanding to know the applications to which a service is relevant and the differences in what these applications require from the service. To overcome this challenge the client decided to apply use cases (Cockburn 2000; Jacobson et al. 1992). At first, some of the offshore developers that had been on placements at the client, but had returned to India, were asked to lead the writing of use cases in India. It was agreed to use a writing style with four abstraction levels with the first being mostly business oriented and the fourth very technically oriented. But when the results came in, the more business oriented use-case levels just consisted of pointers to lower levels, and the client discovered that it took way too much effort to review the very detailed technical use cases at the fourth level. Thus, that way of dividing work did not minimize interaction. In a second round, business staff at the client was taught to write use cases. These upper-level use cases were then given to the vendor who wrote the technical levels. This proved to minimize interaction much better.

6 DISCUSSION

Cultural differences between client and vendor are an inherent characteristic of offshore outsourcing. For offshoring to India it is also common that the vendor’s development processes are at a higher maturity level than the client’s development processes (Levina et al. 2008; Vlaar et al. 2008). The case investigated in this study concerns whether such inequalities can be handled by minimizing the interaction between client and vendor.

6.1 Costs and challenges of minimal interaction

In the section above we mentioned project A as an example of minimizing interaction. While the project was successful in the sense that the offshoring arrangement produced a high-quality system, it was restricted in the sense that only a modest part of the activities of a full project were performed by the vendor. The entire project A was offshored to the vendor, but project A was special in the sense that it consisted almost entirely of programming. In this sense project B is a better example of the client’s minimizing interaction strategy, because a larger amount and variety of development activities were offshored to the vendor. Figure 1 shows how work has been divided between the client and the vendor. In the beginning the bottleneck of this model was in the middle. It was difficult to offshore enough coding activities to the vendor and test the quality of the produced code. Today the bottleneck has moved to the front and back ends of the process. About 400 people are ready to work at the vendor site, starting from business oriented use cases and delivering integrated code ready for acceptance. The hard part now, tells a manager at the client, is to get the business people to write enough, high-quality use cases – that is, to decide and specify how they want the business processes to be.

The main limitation of the client’s approach has been that in order to minimize interaction with the vendor it has become necessary to perform considerable extra work. This work is required to enable the vendor to take on tasks in spite of its limited business knowledge. The extra work consists of preparing tasks for offshoring, preparing the vendor for working in the client’s application domain, and assessing the quality of the vendor’s work. In the terminology of Dibbern et al. (2008) this extra work corresponds to specification costs, knowledge-transfer costs, and control costs. While the knowledge-transfer activities and certification programs are intended to gradually make it possible for the client to offshore also the specification of systems and the assessment of work products, the currently offshored activities are somewhat biased toward programming. Thus, the client is succeeding in offshoring programming but, at least currently, at the cost of extra work on other activities. Compared to previous, onshore development the client’s activities have shifted toward the start and end of the development process, see Figure 1.
This shift has important consequences for the client. First, it implies that the client is to a considerable extent doing work in order not to have to do work. The amounts of extra work have not been fully foreseen, and cultural differences entail that the extra work is perceived differently by client and vendor. For example, the vendor organizes activities partly from the implicit perspective that hours are cheap and capacity large, but this perspective is defective when some of the hours (e.g., control activities) are to be performed by the client. It is an ongoing learning process to identify and reduce areas of extra work but also to realize that offshoring is increasing the amount of some of the client’s tasks. Second, the extra work may exceed the capacity of the client staff and thereby prevent the client from offshoring as much work as the vendor would be able to perform. While the bottleneck that initially motivated the client to offshore was perceived as a shortage of programming capacity, it may now emerge as a shortage of capacity to specify systems and control work products. This way, the uneven distribution of business knowledge may be the factor that limits the client’s minimal-interaction approach to offshoring, making a reduction of the knowledge asymmetry central to continued success with this approach. Third, the tasks of the client staff are changing. This implies that client staff increasingly needs a different mix of competences with more focus on business understanding and abilities to facilitate the formulation of requirements, the transformation of requirements into system specifications, and the follow-up on whether developed systems match business requirements. Some client staff may welcome this change of focus; others may be reluctant to give up time for programming in favour of activities at which they feel less proficient and comfortable.

6.2 Conway’s law

Conway’s law (Conway 1968) states that “organizations which design systems … are constrained to produce designs which are copies of the communication structures of these organizations.” Thus, the communication bottleneck between client and vendor in offshoring arrangements will lead to system designs that reproduce this structure. Conway concludes that flexibility of organization is central to effective design. Flexibility is needed to be able to adjust the organizational structure to a system.
architecture that matches the needs of the use situation. Because designers’ understanding of these needs will likely evolve during the development process, flexibility of organization is required throughout the development process, not just when projects are set up.

Conway argues that especially for large systems the required flexibility is rarely present and that the structures of large systems therefore tend to disintegrate during development. This disintegration is the result of a three-step process. First, when designers realize that a system will be large they are tempted to assign too many people to the project. This temptation is exacerbated by access to a large pool of development staff, as is typical in offshoring. Second, in a large project the communication paths must be restricted in order to avoid that communication consumes all people’s time, as exemplified by the single point of entry enforced between offshore staff and the client’s quality-assessment staff. This causes the communication structure to disintegrate. Third, Conway’s law ensures that the disintegration of the communication structure will be reproduced in the system structure, which therefore also disintegrates. This argument appears pertinent to offshoring because the client gets access to the vendor’s large pool of development staff and because the communication between client and vendor is already restricted by their physical separation (e.g., Herbsleb 2007; Herbsleb et al. 1999).

In projects A and B we can clearly explain part of what we see by using Conway’s law. Project A, for example, complied with Conway’s law by reproducing the organizational separation between an onshore group with business knowledge and an offshore group with technical knowledge in the system: The system was completely rebuilt technically but remained completely unchanged functionally (as it was planned). This was not a problem for project A itself, because it was the client’s first offshore project and a lot was learned from it. However, project B gives some indication that the client’s aim of eventually offshoring entire projects from requirements to implementation is still hampered by the uneven distribution of business knowledge. This makes communication about business knowledge a central bottleneck because it leads to missed opportunities in the offshore-developed systems.

7 CONCLUSION

In October 2008 while we were making interviews in the Danish organisation it was decided to renew the contract with the vendor without changing anything but the ending-date of the contract. Thus, the way the interaction had worked between client and vendor was considered a success. We have shown in this paper that the strategy followed was one of minimizing interaction. Concretely, we have shown how keeping distance, exchanging people, and exchanging knowledge can be used to develop software in a way that minimizes interaction. It should be noted, however, that achieving minimized interaction requires a lot of work. It is not a cheap solution; the price (and time) invested by the client was larger than expected.

We conclude that minimizing interaction can be a viable strategy to follow when clients and vendors face large cultural and maturity inequality in offshore outsourcing.

Acknowledgements

This work has been supported by the Danish Agency for Science, Technology, and Innovation through its funding of the SourceIT project. We wish to thank our colleagues in the SourceIT project for inspiring discussions and master-thesis students Sofia Knudsen and Magnus Hansen for their support in documenting the interviews. Special thanks are due to the interviewees for their openness and their willingness to take part in our study in spite of their busy schedules.
References


TRACKING BASED PRODUCT AUTHENTICATION: CATCHING INTRUDERS IN THE SUPPLY CHAIN

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0282.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>RFID, Ubiquitous systems, Supply networks, Business Intelligence</td>
</tr>
</tbody>
</table>
Oertel, Nina, SAP Research CEC Karlsruhe, Vincenz-Priessnitz Strasse 1, 76131 Karlsruhe, Germany, nina.oertel@sap.com

Abstract

Product counterfeiting is a growing problem worldwide, threatening the health of consumers and reducing company profits. By detecting and intercepting counterfeits before they reach the customer, the problem can be mitigated. In this paper, an approach to detect counterfeit items based on their claimed history is presented. The necessary data is provided by tracking infrastructures that enable the recording and retrieval of movements of individual items in the supply chain based on unique identifiers assigned to products. If the movement history of an item deviates from the movements of genuine items that have been learned before, a warning about a potential counterfeit is issued. Counterfeiter activities that are possible in a tracking enabled environment are modelled and the capability of the proposed approach to detect these strategies is assessed.

Keywords: Anti-Counterfeiting, Authentication, Supply Chain, RFID
1 INTRODUCTION

Counterfeiting is growing in terms of industries and countries affected, as well as the number of fake goods seized by customs each year. The OECD (2007) estimates that the international trade in counterfeit products was up to 200 billion $ in 2005. Products subject to counterfeiting range from luxury handbags, sneakers and liquor to plane spare parts and pharmaceuticals. A counterfeit is any product that infringes intellectual property rights such as trademarks, patents and designs. This includes items produced in factory overruns - by the same manufacturers that produce genuine products during normal working hours - although they do not differ in appearance from genuine products.

Each counterfeit that is sold to a customer affects the rights holder negatively. Potential consequences range from reducing the sales of genuine products and affecting brand value to disappointed or even physically harmed customers (OECD 2007, Staake et al. 2008). Rights holders attempt to minimize the negative consequences by implementing anti-counterfeiting measures that aim at cutting off counterfeit supply and increasing the number of detected counterfeits. Product authentication approaches, i.e., methods that aim at deciding whether a product is counterfeited or genuine, enable stakeholders such as customers, customs and supply chain partners to check products and contribute thus directly to an increased counterfeit detection rate. Moreover, the higher the detection rate, the less profitable is the business for the counterfeiter as revenues decrease with each detected counterfeit, while counterfeit production costs persist. For each counterfeiter organization there is a detection rate that reduces the profits to zero and drives the counterfeiter out of business. Additionally, as profits decline, counterfeiters might target products that are less well protected and thus promise larger profits. This contributes to decreasing the supply of counterfeits.

Traditional authentication approaches, e.g., the check of holograms, base authentication on something an item has. Authentication can also be based on something an object did in the past (Stajano 2002). With the advent of item serialization – facilitated by radio-frequency identification (RFID) technology - and tracking infrastructures, history information about individual items is available. Serialization means that each item is assigned a unique identifier (UID) that can be stored on an RFID tag attached to the item. Suitable readers are deployed in the supply chain. When an item passes, an event is created that contains UID, time, location and further context information. Tracking infrastructures then enable the storing and sharing of events between companies. To illustrate how this information can be used for authentication, assume that product #123 was seen in Paris at 12:00 h, followed by a sighting in Tokyo at 12:05 h. As item #123 cannot physically have moved over this distance in such a short period of time, it can be concluded that the item is carrying a duplicated serial number and is therefore probably a counterfeit.

While the potential of tracking for counterfeit detection has been recognized (Koh et al. 2003; Staake et al. 2005; Inaba 2008), few concrete methods have been suggested so far. The general difficulty is to specify rules operating on the item history that accurately distinguish between counterfeit and genuine products, while keeping the specification effort manageable. The research question is thus how to design a counterfeit detection approach based on tracking data that addresses these requirements. The approach proposed in this paper learns rules governing the behaviour of genuine items from available tracking data, thus eliminating the need for manual specification. The warning about potential counterfeits is based on anomaly detection. The research thus contributes to the general aim of developing methods and systems for analysing tracking data to support business decisions (Asif & Mandviwalla 2005; Curtin et al. 2007).

The research method pursued is reflected in the structure of the paper: In the next chapter, related work is reviewed and shortcomings are identified. Chapter 3 gives an overview of item tracking and the proposed authentication approach while chapter 4 presents the methods for learning and counterfeit detection in detail. To evaluate the approach, a model of counterfeiter behaviour in a
tracking enabled world is developed and utilized to assess the detection capabilities of the approach, before concluding and suggesting areas for future work.

2 RELATED WORK

A large share of authentication approaches in use today are based on adding a hard to imitate, artificial feature to the product, e.g. holograms, printing in colour-changing ink or chemical markers (OECD 2007). Authentication with these approaches requires special equipment, expertise or can only be performed one product at a time. As the counterfeit detection rate depends not only on the detection capability of an approach, but also on the check rate, i.e. the fraction of products that are actually checked, approaches that do not allow for bulk authentication in the supply chain will usually result in lower detection rates.

With the advent of RFID technology, cryptographic authentication approaches that make use of the computational capabilities of RFID tags have been proposed (Juels 2006). Cryptographic authentication cannot be performed with low-cost RFID tags as they have only very limited processing power. Once a method has been devised to perfectly imitate a security feature – for instance by hacking cryptographic RFID tags - the approach is no longer secure and counterfeits become undetectable.

Some proposed authentication approaches utilize item serialization, i.e. the fact that each genuine item was assigned a UID during manufacturing. It was proposed to check the validity of the serial number of products against a database of all serial numbers assigned during manufacturing (Lei et al. 2005; Johnston 2005). In the context of tracking infrastructures, this approach was suggested to be used for pharmaceuticals (Koh et al. 2003). Drawbacks of this approach include the low protection against UID copies and the possibility of counterfeiters to skim the database for valid numbers.

When more tracking data is available, the full product history can be made accessible and checked by the prospective buyer for plausibility (Staake et al. 2005). Some types of plausibility rules are suggested by Illic et al. (2009). For specifying individual rules, background knowledge and manual input is required, e.g. the maximum allowed item stay time per location must be configured. The drawback of these approaches is the effort imposed on the user for rule specification and decision.

Few approaches have been proposed that learn rules from available data. It has been suggested to learn which product transitions between any two locations are plausible (Illic et al. 2009, Lehtonen et al. 2007). By considering only two locations, these rules are not precise enough to detect some counterfeit types. The approaches use a probabilistic decision function that classifies items as counterfeit if the observed movements are infrequent. This poses a problem if genuine items perform unusual movements. The Hidden-Markov-Model based approach by Lehtonen et al. (2007) detects counterfeits based on similarity with previously detected counterfeits, but does not adapt well to new counterfeiter strategies.

The counterfeit detection approach proposed in here extends on previous work by learning rules from tracking data, but adds precision by considering the full context information contained in events, event timing, and the full sequence of events instead of two consecutive events. The decision will be based on the discovery of anomalies in event sequences. The concept of anomaly detection has been applied in fraud and computer intrusion detection. In both domains, anomalies caused by fraudster activity are sought among sequences of events, e.g. credit card transactions, phone calls, or system calls (Bolton & Hand 2002). Fraud or intrusions are detected by identifying deviations from previously learned profiles (Forrest 1997; Hilas & Sahalos 2005). A main difference to counterfeit detection is that these systems operate on continuous event streams, while a product has a lifecycle and the corresponding event sequence has a beginning, end, and a regular structure determined by the underlying business processes.
3 OVERVIEW OF TRACKING-BASED COUNTERFEIT DETECTION

3.1 Serialization and Tracking

Serialization and tracking infrastructures are necessary prerequisites for tracking-based counterfeit detection. The UID may be stored on an RFID tag attached to a product, facilitating the simultaneous identification of multiple items without line of sight. However, tracking is not limited to RFID tagged products, and in principle any technology able to provide UIDs can be employed, e.g., printed 2D-barcodes. To enable tracking, suitable reading devices need to be distributed in the supply chain that capture events when items pass. The attributes to be contained in an event are usually determined by the tracking infrastructure. For example, the attributes (UID, time, location, business step) specify a simple event format. An event instance contains values in the domain of each attribute, e.g., (123, 12:00 h, Paris Shop A, Sale). In a business setting, further useful attributes include the organizational unit, associated business transactions, lifecycle information, and aggregation with other products. The captured events are submitted to a local database – the repository - for future reference. A tracking infrastructure provides methods for finding and assembling the events pertaining to a specific item, which are usually distributed over multiple repositories and even multiple companies.

The trace of an item is defined as the sequence of all events containing the item UID ordered by their timestamp. The trace may not only contain information about the physical movements from location to location, but also about the underlying business processes and the actions performed. This is illustrated by the traces of two sample products in Table 1. Both items are produced in Lyon and then delivered to a distribution centre in Paris. While item #123 does not pass the quality control and is therefore disposed of, item #126 is shipped to a retailer and subsequently sold to a customer.

<table>
<thead>
<tr>
<th>UID</th>
<th>Event</th>
<th>Time</th>
<th>Location</th>
<th>Step</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
<td>1</td>
<td>12.08.2008;12:00 h</td>
<td>Lyon Plant A</td>
<td>Production</td>
</tr>
<tr>
<td>123</td>
<td>2</td>
<td>22.08.2008;12:00 h</td>
<td>Paris Distribution Centre A</td>
<td>Receiving</td>
</tr>
<tr>
<td>123</td>
<td>3</td>
<td>22.08.2008;14:00 h</td>
<td>Paris Distribution Centre A</td>
<td>Quality Control</td>
</tr>
<tr>
<td>123</td>
<td>4</td>
<td>22.08.2008;15:00 h</td>
<td>Paris Distribution Centre A</td>
<td>Disposal</td>
</tr>
<tr>
<td>126</td>
<td>1</td>
<td>23.08.2008;11:00 h</td>
<td>Lyon Plant A</td>
<td>Production</td>
</tr>
<tr>
<td>126</td>
<td>2</td>
<td>25.08.2008;12:00 h</td>
<td>Paris Distribution Centre A</td>
<td>Receiving</td>
</tr>
<tr>
<td>126</td>
<td>3</td>
<td>26.08.2008;12:00 h</td>
<td>Paris Distribution Centre A</td>
<td>Shipping</td>
</tr>
<tr>
<td>126</td>
<td>4</td>
<td>26.08.2008;14:00 h</td>
<td>Paris Shop B</td>
<td>Receiving</td>
</tr>
<tr>
<td>126</td>
<td>5</td>
<td>27.08.2008; 11:00 h</td>
<td>Paris Ship B</td>
<td>Sale</td>
</tr>
</tbody>
</table>

Table 1. Sample traces of two items represented in an event format that contains the two context attributes location (represented by city and facility) and process step.

Serialization and tracking are not yet in widespread use, although trials of limited scope are performed. However, some product types, e.g., luxury watches or plane parts, already carry serial numbers and proprietary systems are used to track them over their lifecycle. Standardization of tracking infrastructures is under way, most notably in the context of the EPCglobal framework, a set of standards targeted at enabling an interorganisational, distributed tracking infrastructure based on globally unique identifiers (EPCglobal 2007b). Part of the framework is also a specification of an event format suitable for a business context (EPCglobal 2007a). Government mandates requiring serialization and tracking for high risk products such as pharmaceuticals might further foster the adoption of tracking based systems (e.g. FDA 2004).
The event format presupposed for the proposed counterfeit detection approach is generic in so far as only a UID, timestamp and at least one recurring attribute is required. The approach is thus independent from a specific standard such as the EPCglobal framework and can be adapted to the event format utilized by any tracking system that fulfils these basic requirements.

3.2 Components of the Authentication Approach

An overview of the components of the proposed anti-counterfeiting approach and the data flows between them is given in figure 1. The model learner takes as input a set of sample traces from the tracking repository and infers from them the rules according to which items move. The output is a model of the supply chain containing all rules. The learning step can be repeated periodically to update the model with current traces. For every product type, a separate model should be constructed, so that the trace of a handbag is not compared to traces of car brakes. The Electronic Product Code (EPC) standard contains the product type identifier as prefix in the UID, followed by the actual serial number of the item (EPCglobal 2007b). For all products identified by EPCs and UIDs constructed in a similar manner, the product type is directly available from the events.

The quality of the learned model depends on the quality of the traces contained in the sample. The sample must fulfil the following requirements:

- The traces should cover all potential product movements, including legitimate exceptions in product routing or process execution, as well as seasonal variations.
- All traces must belong to genuine items, otherwise the movements exhibited by counterfeit products will later be accepted as normal.

To fulfil the first requirement, the sample size must be chosen large enough and cover a sufficiently long period of time. Furthermore, commonly used data cleansing techniques aimed at filtering out outliers and exceptions based on low observation frequencies should not be applied. The problem of separating events triggered by legitimate users from events generated by intruders is well known from the intrusion detection domain. But unlike intrusion detection, which operates in the virtual space, counterfeit detection deals with actual physical objects. This allows for authenticating items with non-tracking based approaches, e.g., relying on security features or expert judgement. It is also possible to send designated test items from the manufacturer through the supply chain, authenticating them at each read point, and basing the sample on their traces. Only traces of items confirmed to be genuine should be admitted to the sample. Another possibility to ensure that only traces of genuine items are contained is to visualize the model with a suitable representation (e.g. Illie et al. 2009, Hoffman et al. 2008) and have it inspected by an expert.

The decider takes the trace of the product to be authenticated, the learned model and possibly also the position at which the authentication is performed as input. Position information is required because items might also be authenticated at intermediate points in the supply chain. In fact, every capturing of an event may be coupled with authentication, thus leading to a high check rate. This allows to intercept counterfeits early, before they reach the customer. The position information can be supplied by the user, the reading device or by an authentication service. The decider compares the trace reported by the item to be authenticated with the supply chain model (up to the authentication position) and computes the authentication result. If a mismatch is found, the item is a potential counterfeit and a warning is issued. If no anomaly is detected, the trace is reported to be plausible. These cautious results are justified because even for plausible traces, it cannot be guaranteed that the item is genuine. Likewise, deviating traces might be caused by errors, e.g., wrong item routing and need not necessarily be caused by counterfeiting activity. The detailed algorithms used by the decider and the model learner are presented in the next chapter.
4 AUTHENTICATION APPROACH

4.1 Representing Item Movements

The supply chain model shall in a general yet precise way describe the movements of genuine items. The movements as represented in tracking data are characterized by

- the events contained in the trace,
- the sequence of events, and
- the timing of events in the trace relative to each other.

Before describing how the model is constructed from sample traces, it will be described how each of these three concepts is represented.

The attributes specified by an event format can be divided into recurring and non-recurring attributes. The values of recurring attributes can be expected to be observed for further items, while the values of non-recurring attributes have a limited lifespan. The UID – bound by the lifetime of an item - and the event time – being monotonically increasing – are non-recurring. In the above example (table 1), the attributes location and process step are recurring. As their values are determined by the underlying business processes and the design of the supply chain, the values can be expected to be observed also for future items. The set of recurring attributes can be determined by the user (as this is necessary only once per event format) or automatically by scanning the sample traces for attribute values recurring over the sample period. A unique combination of values of recurring attributes will be called a station. Examples of stations are (Lyon Plant A, Production) or (Paris Distribution Centre A, Receiving). Each event maps to exactly one station, while a station represents many different events.

To accurately represent an event, its position in the trace and prior events must be considered in addition to the attribute values contained. Take for example a process in which items are shipped from a central distribution centre to shops, and, in case they are not sold, sent back to the distribution centre and shipped to outlet centres. Some items will pass the shipping ramp - the station (Paris Distribution Centre A, Shipping) - twice, but depending on the position in the trace, the legitimate next event will vary, being either the reception at a shop or at an outlet centre. Therefore, the concept of a node is introduced to represent an event at a specific position in the trace. Each node references a station that indicates which attribute values are contained in the event represented by the node. Multiple nodes (representing events at different positions in the trace) may reference the same station.

The ordering of events in a trace is represented by transitions between nodes. A transition is a directed arc from a start node to a successor node. The timing of events relative to each other is given by the difference in the timestamps of two consecutive events in a trace. The time is therefore a property of
the transition between event representations. When considering multiple items that share the same transition in their traces, the average, minimum and maximum *transition times* are computed for the transition.

4.2 Constructing the supply chain model

Given stations, nodes, transitions and transition times as building blocks, a model of the legitimate behaviour of items in the supply chain can be constructed from a set of traces of genuine items. Put shortly, the model is built up by transforming each trace into a sequence of nodes connected by transitions and then merging sequences sharing the same prefix together. The resulting model is a tree shaped directed acyclic graph containing nodes and transitions, and an additional look-up table mapping the nodes to stations. The model represents all prior observed sequences of movements of genuine items.

The graph is initialized with a root node and a terminal node. The root node is a common predecessor of all nodes representing potential first events in traces. The root node connects for example the traces of items produced in Lyon and items produced in Tokyo. The terminal node serves as a common end node for all traces. It is needed to mark the potential last events so that the completeness of traces can be checked.

Each trace in the input set is processed in the following way: If the station corresponding to the current event is not yet contained in the model, the station is created and added. Starting with the root node of the graph as the current node, it is tested whether the station corresponding to the current event is referenced by one of the successor nodes. Successor nodes are all nodes to which a transition exists. If such a node is found, the average, minimum and maximum time of the transition between the current and the successor node are updated with the difference between the timestamps of the last and current event from the trace. If no suitable successor node is found, a new successor node referencing the station corresponding to the current event is created. A transition between the current and the successor node is added and the transition times are initialized. As long as there are more events in the trace, the successor node then becomes the new current node and the next event from the trace is processed. If the last event of a trace is reached, a transition to the terminal node is added to indicate that the trace ends at this point.

Figure 2 presents a sample supply chain model learned by this algorithm. Stations are labelled with characters, while nodes are labelled with consecutive integers. The model has the following properties: All direct successors of the root node represent valid start events of traces, the predecessors of the terminal node represent valid last events. The transition from node 8 to the terminal node indicates that \((A, B, F, B)\) is a valid trace. Although the node pairs 4:5 and 8:9 reference the same stations, the preceding nodes 3 respectively 7 influence not only the transition times between stations B and D, but also the set of valid successor events (represented by E respectively G). It is important to note that events may not only have implications for the directly following event, but also on events that happen at any later point in time. These higher level dependencies (spanning multiple events) are commonly found in supply chains. For example, an unpacking event is only valid if items have been packed earlier. Nodes 12-14 showcase a potential repetition of events. By representing events as nodes it is possible to determine how many repetitions are allowed (maximum 3 in the example) and how transition times may change with the number of repetitions.
4.3 Item Authentication

The decider classifies a trace as plausible if its trace conforms to the supply chain model. If this is the case, a similar trace has been observed for a genuine item, rendering the item in question unsuspicious. For a trace to conform to the model, the following conditions must be fulfilled:

- All events in the trace must be valid, i.e., for each event the corresponding station must be contained in the model
- The events in the trace must form a valid sequence, i.e., there exists a path in the graph on which all stations corresponding to events in the trace are connected by transitions
- The trace must be complete, i.e., the path starts at the root node and ends at the terminal node
- The timing of events in the trace is plausible, i.e., the time difference between any two events is within reasonable boundaries of the time span observed for the corresponding transition

If any of these conditions is violated, the item is classified as a potential counterfeit. The four types of mismatches are illustrated in figure 3. The completeness condition needs to be relaxed if an item is authenticated at an intermediate point in the supply chain. For example, if all traces represented by the model end at a shop, but an item is authenticated at a distribution centre, it will unjustifiably be classified as counterfeit. Therefore, if no information about the authentication context is available, the test for a trace ending with the terminal node may be skipped. In case some information about the authentication context is available, transitions from all stations matching with this context to the terminal node should be added temporarily to the supply chain model. The context information may correspond directly to a station, e.g., authentication is performed at the reader located at (Paris Distribution Centre A, Shipping) or only to some attribute values contained in a station, e.g., (?, Sale).

The decider processes the current trace as follows: For each event, the corresponding station is computed and searched in the look-up table. If the station is not found, the item is classified as potential counterfeit. Starting with the root node in the model, a node referencing the current station is searched among the successor nodes. If no such node exists, the item is again a potential counterfeit. If a valid successor node is found, the difference of event times of the last and current event is compared to the corresponding transition time. We propose to check whether the event time difference is smaller than the minimum or greater than the maximum transition time, each extended by a small margin to account for minor variations in event timing. The extension may, e.g., be set to 10% of the difference.
between average time and minimum time for the lower bound and to 10% of the difference between maximum and average time for the upper bound. If all tests are passed, the successor node becomes the new current node and the next event is processed. If the last event in the trace is reached, it is additionally tested whether a transition from the current node to the terminal node exists.

![Figure 3](image_url)

**Figure 3.** Mismatch between supply chain model and trace: a) invalid event, b) incomplete trace, c) invalid transition, d) invalid transition time

### 5 EVALUATION

The information about counterfeiter strategies and the routes on which counterfeited items are put on the market is limited due to the clandestine nature of the operations. To assess the detection capabilities of the proposed approach, it is moreover necessary to systematically explore the possibilities of counterfeiters in tracking enabled supply chains, for which the available data is not suitable. A model of the potential counterfeiter activities is therefore suggested, the effect of the behaviour on item traces is analyzed and used to evaluate the detection capabilities of the proposed approach. In future work, it is planned to computationally verify the results of the analysis by simulating the modelled counterfeiter strategies and normal supply chain operations and assessing the detection capability.

Suppose all items of a certain type are equipped with UIDs and a tracking infrastructure covering the supply chain is in place. The first challenge for the counterfeiter is to obtain UIDs for the counterfeited goods. One option is to completely omit the UID, but this strategy should be easily detectable during authentication. For actually obtaining a UID, the options include guessing identifiers, transferring the UIDs of genuine products to counterfeits or copying the UIDs found on genuine products. If identifiers are guessed randomly, the resulting UID may in rare cases be a duplicate of a valid UID, but most likely it will be invalid, i.e., no genuine item carrying the same UID exists. To transfer UIDs, counterfeiters may remove RFID tags from genuine items and reapply them to counterfeits, or they may seek access to UIDs of disposed products and reuse them. Any UID found on a counterfeit will have at least one of these properties: it is invalid, has been transferred or is duplicated.

The counterfeited items must be distributed, either by injecting them in licit channels or by distributing them outside of the legitimate supply chain, for example by smuggling, selling on flea markets or through online shops. For the resulting item trace, it is important whether there are read points on the path of the item or not. Illicit distribution channels are likely to contain no read points and thus be invisible, while licit channels can be assumed to be visible. The combined effects of UID and distribution strategies on traces will now be analyzed. The various scenarios are illustrated in figure 4.

An item with an *invalid UID* that is traded completely through invisible channels has no trace. This can be detected immediately. Items with invalid UIDs that are injected in the licit supply chains will appear suddenly after the injection point (figure 4a). Some events at the beginning of their trace are missing and the trace will fail the completeness check as the first event in the trace is not a successor
of the root node in the model. However, if a counterfeiter manages to inject counterfeits in the licit channel before the first event is usually captured – which is most likely in the production facilities of genuine items – and it is ensured that this manipulation does not distort the further routing of the item, the resulting trace will fully correspond to the model and the counterfeit will not be detected. If distortions surface later in the trace (not only in terms of location but also in terms of wrong business steps or missing transaction information) the counterfeit will also be detected (figure 4g). An additional check for UID validity offered by the rights holder and strict number management can mitigate this threat.

In case a counterfeit carries a transferred UID (valid and unique), the events in the trace were triggered by the movements of two items: Up to the point in time when the UID was removed, the events were created for a genuine product and the beginning of the trace will conform to the model. After that, the counterfeit carries the stolen UID. Depending on the distribution of the counterfeit, the following possibilities for the continuation of the trace exist: If the counterfeit is distributed outside visible channels, the trace ends early (figure 4b). Unless the last captured event was a valid terminal event, the completeness check will fail. If the counterfeit is injected in the licit supply chain, the sequence of events will only be valid if the counterfeit directly replaces a genuine item, the time needed for the product exchange does not lead to a time threshold violation and the manipulation does not lead to distortions on the future path of the counterfeit. If the injection is not a direct replacement but takes place further upstream (figure 4c), downstream (figure 4d) or in another branch of the supply chain (figure 4e), the model will contain no transition allowing this sequence of events. If the replacement leads to an exceeded transition time, invalid transitions or events later in the trace (figure 4g), it will also be detected.

If counterfeits with duplicate UIDs are distributed through invisible channels, they can be detected as long as the trace of the genuine item carrying the UID is not complete (figure 4a). If the duplicates are injected in the licit supply chain, the trace that is retrieved for the UID is a mix of all sub-traces created by the multiple items carrying the same UID (figure 4f). As soon as the first item with a copied UID is injected in the supply chain, this will result in an invalid trace as a transition between the last event captured for the genuine item (from which the UID was copied) and the first event triggered by the counterfeit is missing. However, if the counterfeiter manages to inject the counterfeit at exactly the station where the next event for the genuine item is expected, the injection will remain undetected, but only as long as the genuine item does not arrive at this station. As soon as the genuine item triggers the next event, the trace becomes invalid as a transition is missing and the counterfeit can be detected at the next sighting. Note that in this specific case, a genuine item will be classified as counterfeit. As counterfeiters will probably put the same copied UID on many items, it may be acceptable to misclassify one genuine item if this allows detecting mass copied items.

![Figure 4](image_url)

**Figure 4. Consequences of counterfeiter strategies on trace data**
From this analysis follows that a wide range of potential counterfeiter activities result in traces that deviate from the traces of genuine items and can thus be detected by the proposed approach. Only few options are open to counterfeiters to “construct” traces for counterfeits that are indistinguishable from the traces of originals. Moreover, each of these remaining options requires considerable effort in terms of knowledge and access to the licit supply chain.

Tracking based authentication enables a high check rate compared to traditional authentication, as automated event capturing can be combined with authentication and, in case of RFID, this can be done for many items simultaneously. The possibility to check items by their UID might also be extended to end customers, further increasing the check rate. The detection capability of the approach for low effort strategies such as identifier guessing without access to the licit supply chain is given, as well as the capability to detect the strategy most prevalent today - the distribution through illicit channels. Additional safeguards such as random inspection of inconspicuous goods, investigation of the sources of detected counterfeits facilitated by tracking data and UID validation may help in limiting the number of counterfeits that cannot be spotted with the proposed approach. In combination with a high check rate, this may result in detection rates that are high enough to drive counterfeiters out of business or make them target more vulnerable products. In any case, tracking based counterfeit detection will raise the bar for counterfeiters and each detected counterfeit decreases the overall negative consequences of counterfeiting.

6 CONCLUSION AND OUTLOOK

In this paper, a counterfeit detection approach based on item serialization and tracking was proposed. Rules governing the behaviour of genuine items are learned from available tracking data. The traces of items to be authenticated are compared to these rules and a counterfeit warning is issued if no match with a rule is found and the trace is thus distorted. By modelling the potential behaviour of counterfeiters in a tracking enabled environment and analyzing the consequences of these actions on item traces, it could be shown that the approach is able to detect a large share of the modelled strategies. It was also analyzed which strategies are undetectable by the suggested approach as they result in traces that are indistinguishable from the traces of originals. These strategies have in common that they require prerequisites from the counterfeiter such as access to the licit supply chain and knowledge about the current location of items. It was argued that the combination of the detection capability and the attainable check rate may lead to favourable counterfeit detection rates, particularly in comparison with traditional authentication approaches.

As some organization are already tracking items or are preparing to do so in the future, they may wish to consider the suggested approach to discourage counterfeiting of their products. Organizations that are affected severely by counterfeiting might assess the suitability of the approach for their specific context. Of particular importance is the trust that companies have in their supply chain partners, their expected willingness to participate in such an approach, as well as the current and expected future strategies of counterfeiters.

It was demonstrated that counterfeit detection is a problem that is in principle amenable to information technology support, but before tracking based counterfeit detection can be applied in practice, many questions remain to be answered. A major concern that could not be addressed in this paper is data sharing in the supply chain. Organizations are generally reluctant to reveal fine-grained and potentially sensitive business data as contained in tracking events. Data access is a prerequisite for model learning and authentication, so a challenge is to enable these functionalities with minimal or no divulgence of critical information. To address this issue, concepts from the area of privacy preserving data mining and secure multi party computation may be assessed. Furthermore, mechanisms for avoiding high false positive rates are required, as a large share of genuine items for which a counterfeit warning is issued might hamper the practical applicability of the approach. Given current rates of read errors in RFID based systems, issuing a warning for any deviation might be a too harsh decision criteria in real world operations and the decision logic may be further refined to account for this requirement.
Operator models need to be developed, e.g., a service hosted by a trusted third party, the tracking infrastructure provider, or a distributed model learning by every supply chain participant. The enforceability, cost and benefit distribution in various supply chain settings might also be explored. How the approach can be extended to end customers, the security implications of this step and user acceptance of a mobile or web based authentication system are further areas of interest.

The proposed tracking based authentication approach might prove to be a useful complement to traditional authentication approaches. Given item serialization and tracking are in place as prerequisites, the proposed approach may help to increase security in the supply chain and to lessen the negative consequences on health and finances caused by counterfeiting.

References


Storing Data on RFID Tags: A Standards-Based Approach

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0382.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>RFID, Standards, Case Study, Ubiquitous systems</td>
</tr>
</tbody>
</table>
Abstract

The potential of Radio Frequency Identification (RFID) for increasing supply chain efficiency has been repeatedly stressed by practitioners and researchers alike. The cross-company usage of RFID applications can only work if the collaborating companies agree on the syntax and semantic used. EPCglobal, an international industry consortium, has specified a stack of specifications that enable a standardized identifier to be stored on the RFID tag and all object related data to be kept on the network. Such a standardized concept does not yet exist to store object related data on RFID tags. To minimize the coordination effort as well as the emerging interoperability or integration problems and, therefore, also guarantee wide-spread adoption of the data-on-tag approach, it is advisable to build on existing standards for the storage of data on RFID tags. In this paper, we recommend applying the ISO 13584 standard for parts libraries (PLIB). We conceptualize how to use ISO 13584 to store data on RFID tags and use a case study on a kitchen furniture manufacturer, which uses RFID tagged components in a cross-company application with its suppliers, to develop a scenario for the storage of data on RFID tags.

Keywords: RFID, Data-on-Tag, Standardization, ISO 13584, PLIB.
1 INTRODUCTION

The potential of Radio Frequency Identification (RFID) for increasing supply chain efficiency has been repeatedly stressed by practitioners and researchers alike (Niederman et al. 2007). The most widely spread practice of using RFID tags on components, products and logistical units as they move through the supply chain follows the GS1 EPCglobal approach (EPCglobal 2007) of storing only an identifier on the tag and all related data in the supply chain participants’ information systems (either using the official EPC or a company specific ID). The main advantages of this approach are the facts that RFID tags are relatively cheap if they only have to store an identifier and do not need memory for user generated content; that it is easy to standardize the identifier, e.g. on the basis of the European Article Number (EAN), which is uniquely assigned by GS1; and that there is no need to encrypt a simple identifier because the access to the data on the network is restricted.

On the other hand, there are several factors that support storing data on the tag (Günther et al. 2008, pp. 143-144). The first factor addresses the need for fast data access – when the IT infrastructure must meet real-time requirements and bottlenecks happen during back-end queries. For such cases, data on tags may help ensure quick access to the required information. The second factor concerns the dependency of the business process, including production, on the back-end system. Storing relevant data on the tag can help production to be kept up without being connected to the back-end system – at least temporarily. The third factor refers to the reliability of the back-end system: storing data on the tag facilitates decentralization and helps avoid single points of failure. This can be relevant if the existing IT infrastructure is not optimized for reliability, e.g. if no redundant system is in place.

The advantages described for both approaches for data storage serve as the disadvantages for the other concept at the same time. But especially because of the differences, Diekmann et al. (2007) claim that these methods are not antithetic but complementary and should be integrated into a consolidated approach; this guarantees that the relevant data is always accessible. In their work, two case studies that employ a combined data-on-network and data-on-tag approach were presented.

When extending RFID applications to inter-organizational processes, the standardization of data formats and data content becomes crucial (Hasselbring 2000). For the data-on-network approach, researchers and companies respectively founded the Auto-ID Center and EPCglobal consortium and have developed the Electronic Product Code (EPC) to uniquely identify physical products (Brock 2001). The data format specification includes a 96-bit code with a fixed, 8-bit header. The standardization of data content is achieved by relying on existing standards (e.g. the Serial Shipping Container Code). The most famous usage of this EPC combines the European Article Number (EAN) with a serial number for each object.

The cross-company usage of data on RFID tags can also only work if the collaborating companies agree upon the syntax and semantic used. Standardization initiatives have taken the first steps in this direction. The German Association of the Automotive Industry (VDA) published a recommendation about the usage of RFID for container management in the automotive supply chain (VDA 2008). The syntax in the user memory is specified by the alternation of data field identifier and value. The semantic of the data fields (IDs, description, data type, etc.) are described in a table in the recommendation, e.g. vehicle identification, maximum quantity of parts, purchase order number, etc. A similar approach for tracking tyres individually has been proposed by the Automotive Industry Action Group (AIAG). In this recommendation, the auto-industry-specific data such as the global location number that identifies the facility where the tyre is made; the tyre cure date and the country of origin should be stored in the user memory (RFID Journal 2005).

Both recommendations to store data on RFID tags will minimize the coordination effort as well as the emerging interoperability or integration problems for all companies in the automotive industry that want to introduce these kinds of applications. However, many different, possibly competing standards for numerous applications within and across industries are not desirable; therefore, to guarantee the
The widespread adoption of the data-on-tag approach, it is advisable to build on existing standards for the storage of data on RFID tags.

In this paper, we recommend applying the part libraries (PLIB) standard ISO 13584. This standard has been widely discussed as a reference model for developing product classification systems and standardized property lists (Leukel et al. 2006b). Major industry consortia have incorporated this standard into their specifications for B2B data exchange (e.g. BMEcat 2005), product classification systems (e.g. eCl@ss, UNSPSC) and property dictionaries (e.g. DINsml). Implementing the PLIB concept for the storage of data on RFID tags promises to avoid heterogeneity and maximize interoperability in cross-company RFID applications.

In Section 2, we conceptualize the usage of ISO 13584 for storing data on RFID tags. A case study of a German kitchen furniture manufacturer is applied in Section 3 to develop a scenario for the storage of data on RFID tags. The company took part in a joint research project on using RFID tagged components in a cross-company application with its suppliers. Finally, Section 4 concludes the paper.

2 PROPERTY-BASED CONCEPT OF ISO 13584

For our purpose, Part 42 of ISO 13584 (ISO 1998), which describes its conceptual model, is of primary interest. ISO 13584 was originally developed to describe technical product data, i.e. functional and physical characteristics, on the basis of unambiguous, semantically well-defined, globally unique properties. Its usage for commercial product data has been proven as well (Leukel et al. 2006a). To describe how the conceptual model of ISO 13584 has to be implemented, the definition and usage of properties have to be distinguished.

2.1 Definition of ISO 13584-compliant properties

The goal of defining ISO 13584-compliant properties is to make them available and accessible in standardized online dictionaries. For instance, the German Institute for Standardization (DIN) Properties Dictionary (http://www.dinsml.net) is based on ISO 13584. That means that each property:

- is identified with a global unique identifier;
- is described with a set of mandatory and optional attributes (e.g. description, unit, data type), which are specified in the ISO 13584's information model;
- is assigned to a set of references to product classes which define context the properties can be used;
- was defined following a standardized process.

The number of properties in the dictionary is continuously growing. Any company can submit new properties to the standardization procedure to be included in the dictionary. The read access to the online dictionary is free of charge, but companies that want to use it without restriction have to purchase a license. This license includes the passing of the properties used to other companies that are involved in their business process (e.g. suppliers and customers). When describing products on the basis of properties, it can happen that the property values are dependent on each other. For instance, if a liquid is described with the property “volume”, it depends on the temperature of the liquid. The property “temperature” is called the condition in this context. To solve this issue, ISO 13584 defines three different types of properties: non-dependent properties, dependent properties and conditions.

2.2 Usage of ISO 13584-compliant properties for the storage of data on RFID tags

RFID technology in supply chain management can be used to track components, finished products or logistical units (e.g. containers, pallets, cases). These objects have certain characteristics which might be stored on their respective RFID tag. Following ISO 13584, all data have to be expressed in form of property-value pairs. This information has to be very precise, e.g. concerning meaning of the property.
or unit of the value, but to minimize the amount of data storage needed on tags, only the ID of the standardized property (which includes this precise information) and the value should be stored.

Figure 1. Logical model for property usage in XML Schema format

When storing these property-value pairs on tags, the syntax and semantic layers have to be distinguished. Semantically, all three different types of properties have to be supported; therefore, it is not enough to provide only the property id and value for a property: for a dependent property, the reference to the value of the condition has also to be given. One option for this usage of property-value pairs is described in Figure 1, which uses a graphical representation of an XML Schema. The reference to the value of the condition is called “PropertyConditionIDRef” in this model.

To illustrate the usage of the three different types of properties, the example in Figure 2 shows the independent property “colour” and the dependent property “optical density”, which is dependent on the condition “optical glass type”. The representation format follows the model in Figure 1.

Although we have used an XML format based on the developed XML Schema to describe the semantics of using property-based product descriptions, it is not necessary to utilize this format as the syntax. The main disadvantage appears when comparing the payload of this short example with the XML element names etc. – which make up 91% of the characters and symbols – which are not essential for the content.

2.3 Benefits of using ISO 13584

Standards in general contribute to the harmonization of interfaces between heterogeneous systems and, for this reason, increase interoperability. This may result in decreasing coordination efforts and wider usage. In the context of storing data on RFID tags, besides the technical interoperability, data interoperability is of great importance. Typical problems include the following:

• data could be misinterpreted because the information is not understandable;
• different data models could describe the same information;
• and different information could have the same description in individual data models.

For the mapping between different data models, a high coordination effort is needed to overcome these problems, if it is at all possible.

From another point of view, competing standards do not solve this issue if companies do not know which standard they should choose. In information systems literature this is known as the standardization problem (Westarp et al. 2000). The different standardization efforts introduced in Section 1 for the usage of RFID in the automotive industry highlight this problem.

The property based concept of ISO 13584 addresses both problems. First, the standardized properties are precisely defined according to the ISO 13584 data model, which includes language independent verbal definitions as well as additional information regarding units, data types, etc. This results in easier data exchange via standardized interfaces, higher data quality, and a reduction of data redundancy (Pohn 2006). Second, using ISO 13584 implies following a bottom-up approach because with properties small pieces of information are standardized and so can be applied to very different
circumstances. Other standardization initiatives (e.g. AIAG and VDA) can create their own standards based on these standardized properties. In this way, ISO 13584 is not competing with other initiatives and, even more importantly, with the EPCglobal approach.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<Properties xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <Property>
    <!-- Colour -->
    <PropertyID>DIN-AAB245-002</PropertyID>
    <PropertyValue>blue</PropertyValue>
  </Property>

  <Property>
    <!-- Optical Glass Type -->
    <PropertyID>DIN-AAA179-002</PropertyID>
    <PropertyConditionID>ID1</PropertyConditionID>
    <PropertyValue>BK7</PropertyValue>
    <PropertyValueCondition>
      <PropertyConditionIDRef>ID1</PropertyConditionIDRef>
    </PropertyValueCondition>
  </Property>

  <Property>
    <!-- Optical Glass Type -->
    <PropertyID>DIN-AAA179-002</PropertyID>
    <PropertyConditionID>ID2</PropertyConditionID>
    <PropertyValue>PYREX</PropertyValue>
    <PropertyValueCondition>
      <PropertyConditionIDRef>ID2</PropertyConditionIDRef>
    </PropertyValueCondition>
  </Property>

  <Property>
    <!-- Optical Density -->
    <PropertyID>DIN-AAB097-002</PropertyID>
    <PropertyValueDependent>
      <PropertyValue>2.51</PropertyValue>
      <PropertyConditionIDRef>ID1</PropertyConditionIDRef>
    </PropertyValueDependent>
  </Property>

  <Property>
    <!-- Optical Density -->
    <PropertyID>DIN-AAB097-002</PropertyID>
    <PropertyValueDependent>
      <PropertyValue>2.23</PropertyValue>
      <PropertyConditionIDRef>ID2</PropertyConditionIDRef>
    </PropertyValueDependent>
  </Property>
</Properties>
```

Figure 2. Example for the usage of different types of properties

2.4 Extraction of ISO 13584-compliant Properties

If a company chooses to use ISO 13584-compliant properties for a certain use case, the properties usually do not have to be developed from scratch. The existing data sources in the companies’ information systems about the object under investigation have to be taken into consideration. Data models provide a useful basis for the identification of properties because the attributes describing an entity in an ER diagram or a class in a UML class diagram can often be transferred into a property, while associations between entities or classes refer to the type of property (dependent, independent, etc.). A methodology for extracting properties from an XML Schema has already been developed by Leukel et al. (2006a).

Once the required properties are determined, the corresponding properties in the property dictionary have to be found. Three cases could appear in general: first, a corresponding property in the dictionary exists and the semantic of the usage is the same. In this case, the property ID from the dictionary and the values from the existing instances can be used without further processing. Second, a corresponding property in the dictionary exists, but its property definition differs in usage, e.g. about data type or measurement unit. Here, the property ID from the dictionary can be used if the instances can be
transformed into the required semantic. Finally, the company could require properties that have not been defined in the dictionary yet. This is highly dependent on the type of object, e.g. some industries are more actively working together with the dictionary operator. In this case, the new property should be created in the dictionary. As already mentioned in Section 3.1, this creation has to follow a certain standardization process that has several phases (initiation, evaluation, etc.), which takes 25 weeks at most.

3 CASE STUDY OF A GERMAN KITCHEN MANUFACTURER

Wellmann GmbH & Co. KG (Wellmann) based in Enger, Germany, is a kitchen manufacturing specialist. Founded in 1953 by Gustav Wellmann, the company was taken over by Alno AG in 2003. In its supply chain, Wellmann’s position is that of the Original Equipment Manufacturer (OEM). The structure of its suppliers is very heterogeneous and varies from small factories to industrial producers and logistic service providers. The kitchens produced are offered by several retailers – mostly under their own brand names. The business goal of Wellmann is to provide high quality, complete kitchens on schedule for competitive prices, despite numerous variants and a high share of individual and special parts.

This case study covers all processes, including procurement and production, that are relevant for introducing RFID at Wellmann. This introduction is currently being analyzed in a joint research project by the company and two universities and is being partly financed by the German Federal Ministry of Economics and Technology. The implementation of RFID in Wellmann’s logistical and production processes is done in several successive steps. In the initial steps, only parts with the most critical logistical importance (Strassner 2005), i.e. parts of high value, are tagged with RFID transponders. These are kitchen cupboard fronts that are custom ordered; produced by an external supplier; and delivered just-in-time for assembly at Wellmann.

3.1 Scenario description

The production of cupboards with fronts follows a structured process that consists of ordering parts, picking parts, drilling fronts, assembling components and assembling whole cupboards. The process starts with a customer who plans his or her kitchen at a retailer which in turn creates a custom order for Wellmann. As soon as the route planning for the custom orders is finished, the orders for the externally sourced parts are automatically placed by Wellmann’s Enterprise Resource Planning (ERP) system once a day to all affected suppliers. Each position in the electronic order is always associated to its corresponding position in the custom order. In this case study, we focus on fronts, which can either consist of wood or glass. Glass fronts are either framed in wood on all four sides or only have wood on the top and bottom. The combination of individual dimensions, different types of handles and other characteristics makes the production of cupboards with fronts vary greatly.

Parts are moved through the factory in special containers, which are identified with a unique twelve digit Transport Group number (TG number). Within this container, the parts are identified with a twelve digit Transport position number (TG position number). Additionally, each part is identified with a twelve digit Transport order number (TA number). All numbers are newly assigned before each production step. To control all production steps, the numbers of all objects are stored in Wellmann’s ERP software. The data model for identification numbers is depicted in Figure 3.

![Figure 3. UML model describing the identification scheme](image-url)
The supplier produces the fronts ordered and supplies them just-in-time when the assembly of the kitchen is scheduled. The delivery of the fronts is scheduled for two working days before the cupboards are sent to the customers. The supplier prints a delivery note, which contains only the TG number and the supplier’s delivery note number as a bar code. Additionally, all parts are listed with their corresponding TG position numbers, TA numbers and characteristics (e.g. dimensions and colour). On an attached negative delivery note, all parts from the order that could not be delivered are listed in the same way. Besides these delivery notes, each front has a label attached that contains the TA number for its first processing step as a bar code and the description of its characteristics as well as its production and delivery date.

When the fronts are received at Wellmann, the receiver first scans the TG number on the delivery note or the TA number off of a label attached to one of the fronts. With either of these possibilities, the list of ordered parts is received from the ERP system and displayed. To check the completeness of the delivered parts, the receiver can scan the barcodes (TA number) of each front label and compare the lists. Additionally, the receiver could enter the TG position numbers or the TA number of the parts that have not been delivered into the system. The receiving process is finalized by storing the delivery note number in the ERP system. This triggers a rescheduling of the downstream activities, where the undelivered fronts are removed from further processing steps.

For further processing, the fronts have to be buffered directly in front of the next work station. In this case, the fronts are drilled at a CNC (Computer Numerical Control) jig boring machine. Each bore program is retrieved from a central server over the network and always relates to one special front. This program determines the bore template as well as the depth of the drill holes and the components that have to be assembled (e.g. hinges and cushion for the doors). During the production of one batch, which consists of several front containers, changes with the bore program might occur. For this reason, the boring machine is loaded with programs once a new container arrives; therefore, the TG number barcode of the container is scanned from the picking list that was created for this work step. The operational sequence of the boring programs depends on the position of the fronts in the container; therefore, the picking process of fronts into the container has to be done very accurately. If there is one front missing, the bore program for this front has to be removed manually. The operator, who equips the machine with fronts, removes the missing front by selecting the position on the display of the CNC machine. Because the machine works automatically after the operator has equipped the machine with the custom fronts from a container, mistakes have a great effect on the overall process if they are destroyed. In such a case, a complete delivery of the kitchen to the customer at the scheduled time is not possible any more.

After this processing step, a new picking list is printed. This contains the next processing step, the container ID (TG number) and lists the fronts in this container (TG position and TA numbers). After each processing step, this data is retrieved from the ERP system. Additionally, the completion of the processing step is recorded. This acknowledgment is necessary because the production order has to follow the planned and scheduled production process. After the acknowledgement in the ERP system, the next processing step is unlocked.

For the transport to the following processing step, the fronts are put into a new container, which depends on the type of cupboard and the new TG position numbers that are assigned for their following processing step. At the final working station, the fronts are assembled with their bodies that have the same TG position number. The bodies are delivered on an assembly line. For a smooth process, the sequence of the fronts plays an important role again, to avoid the downtime of the assembly line. Cupboards whose fronts have not been delivered have to be set aside. The bodies are buffered until the fronts are produced and provided.

3.2 RFID Process Benefits

In contrast to barcode technology, RFID offers the possibility to simultaneously identify several objects without contact, without line of sight and without human interaction. For these reasons, two
general effects are achieved: first, manual effort can be decreased (e.g. faster receiving and shipping processes); and second, costs incurred due to errors (e.g. at the boring machine) can be reduced.

The general potential from using RFID at Wellmann can be achieved with the data-on-network approach; however, in the production environment where components are tagged and not consumer products, the usage of the EPC is not suitable. Instead, the TA number, which is assigned by Wellmann and pushed to the supplier, is stored as a unique identifier. All manual process steps described in the scenario above are positively affected by the introduction of RFID: the registration of fronts at the receiving business step, the manual deletion of missing fronts at the boring machine as well as the control of the fronts’ sequence. An automatic matching of front and bore program ID helps reduce costs incurred by errors (the higher price for emergency orders for fronts, etc.) and reduce the risk that a kitchen can not be completely delivered to the customer.

Additional to the benefits that can be reached with the data-on-network approach, there are advantages to Wellmann by storing data on the RFID tag. First, the data related to the processing step can be accessed without connection to the network. Second, the bore programs for the boring machine including the specific parameters and characteristics of the fronts can be stored on the transponder. These bore programs can be accessed directly before a front is processed. Furthermore, the completion of the processing step, the retrieval of the following steps and the printing of this plan can be omitted when all necessary data are stored on the RFID tag. The processing progress can be stored on the tags as well as a trigger to release it for the following processing step. In case of a failure or breakdown of Wellmann’s local network, the production is not affected.

For Wellmann’s supplier in this case study, the introduction of RFID in cooperation with Wellmann offers benefits, too. Certainly, the supplier also benefits from the reduction of manual effort (e.g. the shipping process and creation of delivery notes). Additionally, the conflict potential in the transferring of title and risk can be reduced.

Wellmann wants to establish a standardized approach for storing data on RFID tags instead of developing a separate solution with every business partner. Although the prototype described in this case study encompasses only the business relationship to the front supplier, it is Wellmann’s intended target to expand the RFID solution to other suppliers, logistic service providers and retailers. Those, in turn, work together with other manufacturers and will benefit from a standardized approach as well. Searching for a prospective solution not exclusively for this use case, Wellmann chose to store the data on RFID tags in the form of ISO 13584-compliant properties.

![Figure 4. UML diagram showing the objects' characteristics](image)

### 3.3 Extraction of ISO 13584-compliant properties

The effects of the RFID introduction described in the scenario above affect the existing processes only. In this case study, RFID does not act as an enabler for new processes with a transformational effect (Straube et al. 2007); that is why all relevant data that should be stored on RFID tags are
implicitly contained in the current business processes and should be extracted from them instead of being newly created. To define the properties compliant to the ISO 13584 specification, at least the name of the property, the type of the property and the data type have to be extracted. Further information, such as IDs and units, can be manually added later. The needed information can be found in the information model of Wellmann’s ERP system.

The processing steps of the fronts, which are the objects of investigation in this case study, are receiving, producing and assembling. The identifiers related to the fronts as well to the containers the fronts are transported in have already been described in Figure 3. Considering this information model, the RFID tag is put on the fronts, which means that the TA number and TG position number can be extracted as properties (compare Table 1). Because of the association of the class fronts to exactly one container, the TG number can be added as a property as well. Attention has to be paid to the “one or more” association to the processing steps. This indicates that all three properties are not unambiguous, but depend on the ID of the disposition number of the processing step. For this reason, the disposition number builds the condition for the tree dependent properties.

Further characteristics of fronts and their related objects are depicted in Figure 4. In this information model, coming from the class front, all other classes are connected with an association multiplicity of exactly one. That is why no dependent properties have to be created because of multiplicities in the class diagram. Other dependencies (e.g. glass type and density) have to be manually detected. The attributes of the class front can be extracted as independent properties along with their corresponding data types. The attributes of the related classes can be extracted as well, but if the names of the attributes are not unambiguous for the front, the class name has to be added to the property (e.g. cupboard length). The extracted characteristics are shown in Table 1.

The properties of the bore program for the fronts can be extracted from Wellmann’s manufacturing execution system. In general, there exist two options: first, the whole bore program could be stored as one property that has the data type BLOB (binary large object). Second, the properties for the bore program can be extracted from the bore program. The bore program is encoded in PrimeFact’s XNC format (Smeerdijk 2006). The subset of this XML Schema, which is used by Wellmann, is depicted in Figure 5. The ISO 13584-compliant properties are extracted with the methodology proposed in Leukel et al. (2006a) and added to Table 1 (where the attribute SubProgram is renamed to bore program ID).

```
<?xml version="1.0" encoding="UTF-8"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified">
  <xsd:element name="Component">
    <xsd:complexType>
      <xsd:sequence>
        [...]
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
  <xsd:element name="Operations">
    <xsd:complexType>
      <xsd:sequence>
        [...]
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
  <xsd:element name="Position">
    <xsd:complexType>
      <xsd:attribute name="X" type="xsd:double"/>
      <xsd:attribute name="Y" type="xsd:double"/>
      <xsd:attribute name="Z" type="xsd:double"/>
    </xsd:complexType>
  </xsd:element>
  <xsd:element name="SubProgram">
    <xsd:complexType>
      <xsd:attribute name="SubProgram" type="xsd:string"/>
    </xsd:complexType>
  </xsd:element>
</xsd:schema>
```

*Figure 5. XML Schema for bore programs based on PrimeFact’s XNC format*
<table>
<thead>
<tr>
<th>Ser. No.</th>
<th>Preferred Property Name</th>
<th>Example of Value</th>
<th>Unit</th>
<th>Property Data Type</th>
<th>Type of Property</th>
<th>Dependent on Ser. No.</th>
<th>Possible Identification in DINsml</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Name</td>
<td>Front</td>
<td>String</td>
<td>Non-dependent</td>
<td></td>
<td></td>
<td>DIN-AAA054-002</td>
</tr>
<tr>
<td>2</td>
<td>Length</td>
<td>0.625 m</td>
<td>Double</td>
<td>Non-dependent</td>
<td></td>
<td></td>
<td>DIN-AAA357-002</td>
</tr>
<tr>
<td>3</td>
<td>Width</td>
<td>0.02 m</td>
<td>Double</td>
<td>Non-dependent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Height</td>
<td>0.2 m</td>
<td>Double</td>
<td>Non-dependent</td>
<td></td>
<td></td>
<td>DIN-AAB517-003</td>
</tr>
<tr>
<td>5</td>
<td>Colour</td>
<td>Blue</td>
<td>String</td>
<td>Non-dependent</td>
<td></td>
<td></td>
<td>DIN-AAB245-002</td>
</tr>
<tr>
<td>6</td>
<td>Glass Type</td>
<td>BK7</td>
<td>String</td>
<td>Condition</td>
<td></td>
<td></td>
<td>DIN-AAB197-002</td>
</tr>
<tr>
<td>7</td>
<td>Optical Density</td>
<td>2.51 D</td>
<td>Double</td>
<td>Dependent</td>
<td>6</td>
<td></td>
<td>DIN-AAB097-002</td>
</tr>
<tr>
<td>8</td>
<td>Delivery Date</td>
<td>2008-11-24 08:24:19</td>
<td>Datetime</td>
<td>Non-dependent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Custom order number</td>
<td>VD234</td>
<td>String</td>
<td>Non-dependent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Disposition number</td>
<td>34K13Z3</td>
<td>String</td>
<td>Condition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>TA number</td>
<td>1234003</td>
<td>Integer</td>
<td>Dependent</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>TG number</td>
<td>1234</td>
<td>Integer</td>
<td>Dependent</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>TG position number</td>
<td>003</td>
<td>Integer</td>
<td>Dependent</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Bore Program ID</td>
<td>98735123</td>
<td>Integer</td>
<td>Condition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>X-axis Value</td>
<td>482.5 mm</td>
<td>Double</td>
<td>Dependent</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Y-axis Value</td>
<td>563.0 mm</td>
<td>Double</td>
<td>Dependent</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Bore Depth</td>
<td>5.9 mm</td>
<td>Double</td>
<td>Dependent</td>
<td>14</td>
<td></td>
<td>DIN-AAA080-002</td>
</tr>
<tr>
<td>18</td>
<td>Bore Diameter</td>
<td>8.0 mm</td>
<td>Double</td>
<td>Dependent</td>
<td>14</td>
<td></td>
<td>DIN-AAA788-002</td>
</tr>
<tr>
<td>19</td>
<td>Bore Program</td>
<td>AD8976B E7620AB</td>
<td>BLOB</td>
<td>Non-dependent</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 1. Extracted ISO 13584-compliant properties*
3.4 Prototypical Implementation

To realize the RFID project at Wellmann, special requirements and challenges have to be considered. First, only one type of RFID tag has to be used for applications in logistics and production. A combined approach of the data-on-tag and data-on-network concepts has to be applied. Second, passive RFID tags are preferred over active ones because of the price differences. For the data-on-network approach, RFID tags according to EPC Gen 2 (ISO/IEC 18000-6) in the UHF frequency range have succeeded in the market and should also be used at Wellmann for this reason.

Nevertheless, most RFID tags that comply with these specifications are only produced with the 96 bit memory for the EPC. For the reasonable usage of the data-on-tag approach at least 1024 bits of additional memory is necessary. The additional cost of this type of RFID tag is currently about 20 Eurocents. Individual companies have to decide for their own RFID business cases whether such a solution is economically reasonable or not. In the case of Wellmann, between 20,000 and 30,000 fronts, which are the subject of investigation in this case study, have to be tagged each year (253 working days). The extra costs of 5000 Euro per year at Wellmann can be justified by the reduction of costs incurred due to errors at the boring machine. Assuming that only 25% of the ca. 5 cupboards per day (1265 per year) that are mis-bored can be saved from errors, which saves about 20 Euro per cupboard for recycling, material and extra logistics, the annual savings add up to 6325 Euro. The qualitative benefits, such as the reduced dependency on the backend system and the number of additional satisfied customers, can be added, but it is not so easy to evaluate them monetarily. In this costs consideration, only the extra costs and savings of using a combined data-on-tag and data-on-network approach were considered, not general RFID benefits and the costs for implementation, maintenance and training.

To prepare for the introduction of RFID at Wellmann, technical tests in the laboratory were first conducted with RFID tags with 512 bits of extra memory. Later on, RFID tags with 1024 bits were chosen because of the needed memory as already stated above. For this prototypical implementation, a proprietary syntactical format – following the VDA recommendation (VDA 2008) with a separator – was chosen to store the data on tags. Since the ISO 13584 standard, which is recommended in this paper, only covers the semantics layer, further research on the syntactical layer is necessary.

4 CONCLUSIONS AND OUTLOOK

A company that considers the introduction of RFID technology has to develop a business case and calculate the related costs and benefits. The business case determines which of the following three scenarios to pursue: (i) the RFID tag stores the identifier only and all object-related data is stored externally (typically a networked databases); (ii) all object-related data is stored on the RFID tag; or (iii) one uses a hybrid approach using both paradigms for different applications. Due to various reasons, for the company described in the case study the hybrid approach is appropriate.

If companies want to use RFID technology in a cross-company application, standards have to be considered. For the data-on-network approach, the EPCglobal standards provide an appropriate solution. A standard for the data-on-tag concept had been missing.

In this paper, we recommended using ISO 13584 for the standardized storage of data on RFID tags. A standard in general contributes to the harmonization of interfaces between heterogeneous systems and, for this reason, increases interoperability, data quality and reduces data redundancy. The properties in our approach are precisely defined according to the ISO 13584 data model, which includes language independent verbal definitions as well as additional information regarding units, data types, etc.

We have shown that using ISO 13584 is a suitable approach for the data-on-tag concept by presenting a case study from Wellmann, a German kitchen manufacturer. Within this case study, we explained how the properties were extracted from the existing information models. For the processes under
consideration, all necessary data could be transformed into independent, dependent and conditional properties. Thus, at least for this application ISO 13584 is appropriate.

The next steps for a successful implementation of ISO 13584 consist of standardizing those properties that have not yet been standardized in the DINsml property dictionary. Subsequently, and after gaining experience with RFID technology, the scope of the application will be expanded.

Although we used a single case study approach in a very specific industry, the goal of using ISO 13584 is that the approach can be used for all other industries and RFID applications. This is because using ISO 13584 implies following a bottom-up approach. With properties, small pieces of information are standardized and can be applied to different circumstances. In this case study, we applied them to the production of custom ordered kitchen cupboards. Other standardization initiatives (e.g. AIAG and VDA in the automotive industry) can create their own standards for certain applications (e.g. container management and theft prevention) based on these or other standardized properties.

References


## P2P Architecture for Ubiquitous Supply Chain Systems

<table>
<thead>
<tr>
<th><strong>Journal:</strong></th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manuscript ID:</strong></td>
<td>ECIS2009-0580.R1</td>
</tr>
<tr>
<td><strong>Submission Type:</strong></td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td><strong>Keyword:</strong></td>
<td>RFID, Peer-to-peer networks / file sharing, Supply chain management, Ubiquitous systems</td>
</tr>
</tbody>
</table>
P2P ARCHITECTURE FOR UBIQUITOUS SUPPLY CHAIN SYSTEMS

Schoenemann, Nina, University of Cologne, Pohligstrasse 1, 50969 Koeln, Germany, schoenemann@wim.uni-koeln.de
Fischbach, Kai, University of Cologne, Pohligstrasse 1, 50969 Koeln, Germany, fischbach@wim.uni-koeln.de
Schoder, Detlef, University of Cologne, Pohligstrasse 1, 50969 Koeln, Germany, schoder@wim.uni-koeln.de

Abstract

Ubiquitous information technologies like RFID allow for immediate, extensive and fine-grained capture of real world information. Scalable and efficient networks for exchange of this vast amount of information amongst companies are crucial for the economic exploitation of benefits of ubiquitous information technologies. Existing networks bear several limitations like risks of single-point-of-failures or bottlenecks, unequally distributed power and burdens as well as inflexibility through stringent structures and formats. In particular there is a need for improving the scalability of solutions and ensuring autonomy of network participants. In this paper we introduce a Peer-to-Peer-based architecture for exchanging distributed information, which are shared among participants of a supply chain facilitated with ubiquitous information technologies. This architecture builds on the well-established EPCglobal standards, but can be implemented as an autonomous network. Unlike other architectures it does not need central coordination mechanisms, because it is based on self-organizing Peer-to-Peer protocols. We argue that our architecture supports business processes especially of small and medium-sized enterprises better than other architectures. We provide a discussion about requirements for solutions and a simulation-based analysis of the proposed architecture.

Keywords: RFID, Peer-to-peer networks, Supply chain management, Ubiquitous systems
INTRODUCTION

Ubiquitous information technologies merge the virtual world with the real world in a seamless way, making embedded information and computing capabilities available anytime and anywhere (Weiser 1991). Gathering real world information with ubiquitous IT enables precise monitoring and controlling of physical objects, their current state, whereabouts, movements and their environmental conditions. Radio Frequency Identification (RFID) is a basic and mature ubiquitous IT, which is applied to automatically identify individual objects and to link related individual information. It has already been thoroughly tested and is in actual use in several companies. Case studies (Fosso Wamba et al. 2008, Loebbecke 2007) as well as analytical studies (Bendavid et al. 2008, Lee & Özer 2007) document its potential both to improve efficiency of business processes, their management and controlling, and to facilitate new services and concepts within the scope of an integrated Supply Chain Management (SCM) (McFarlane & Sheffi 2003, Angeles 2005). Examples are track and tracing of products, automatic inventory and item-level product lifecycle management. Still, RFID systems for SCM, which reach beyond a single organization’s boundaries, are not yet widespread (Gaukler & Seifert & Hausman 2007).

While the decrease of production costs for RFID tags and readers is mainly dependent on technological advancements, the cost-benefit-ratio of the whole system is influenced by design and configuration of the system architecture (Asif & Mandviwalla 2005). A prominent solution is provided by EPCglobal (EPCglobal 2007), a comprehensive framework for standardized, integrated deployment of RFID in supply chains. EPCglobal and other existing solutions for interorganizational RFID systems still bear limitations. Among the main limitations are inefficient, centralized system structures and unbalanced powers between participants (users) and solution providers. Upcoming streams of information may not be manageable by these solutions or become very costly. Central coordination entities could turn out to be bottlenecks or even be attacked by offenders. Furthermore, those entities may also misuse their power. Joint adoption of RFID across the supply chain is still considered unfavorable by many companies. These conditions call for pursuing research in the field of interorganizational ubiquitous IT systems (Ngai et al. 2008, Curtin & Kauffman & Riggins 2007). Therefore we discuss network system architectures and propose a new architecture aiming at overcoming existing limitations. Our architecture enables companies to efficiently exchange locally stored information, captured by RFID systems, over an interorganizational, decentralized network. Our research focuses in particular on adoption of ubiquitous IT by small and medium-sized enterprises (SME).

The remainder of this article is structured as follows: In section two we shortly describe the use of ubiquitous IT in SCM, existing approaches like EPCglobal and related work. The next section comprises our proposal for an architecture based on the Peer-to-Peer (P2P) paradigm. This architecture is then discussed concerning conceptional requirements especially of SME. Further on, a preliminary prototype is presented and tested by means of computational simulations. Finally we conclude with key findings, recommendations, limitations of our work and an outlook to further research.

BACKGROUND

SCM aims at the integration of business processes alongside the whole supply chain in order to increase global competitiveness through flexibility and responsiveness (Bowersox & Closs & Cooper 2007). Information systems play a major role for the integration across the supply chain (Gunasekaran & Ngai 2004). The deployment of RFID supports real-time, ubiquitous monitoring of the integrated supply chain, by making item-level information, e.g. product history and whereabouts, visible for all participants (McFarlane & Sheffi 2003). Other ubiquitous IT like environmental sensors or embedded
systems allow for even richer information services alongside the supply chain, e.g. cold chain monitoring or self-routing packages.

2.1 Supply chain RFID systems

An RFID system commonly consists of three components: Transponder (tag, label), reader and a backend system. A globally unique identifier for item-level identification of physical supply chain objects is needed to allocate information related to individual objects (Främling & Harrison & Brusey 2006). The Electronic Product Code (EPC) builds the identification foundation of the EPCglobal framework. It is an identifying code for objects tagged with RFID, similar to the traditional EAN for barcodes, but on item-level instead of object-class-level. The EPC consists of a header, filter, partition, manager master number, object class and serial number. The first three parts contain format and processing information (e.g. format type and length of object class number). The manager master number indicates the tag owning organization. Each organization that wants to use EPCs on its products has to acquire a master number. In the US and in Europe the EPC is the commonly used identifier. Other identifiers for item-level RFID concurrently were developed and are also in use (e.g. uCode in Japan).

Because memory capacity on tags is limited, extensive information is stored in backend systems connected to the RFID systems by networks (Diekmann & Melski & Schumann 2007). Through RFID-based information services business partners can immediately inform each other about any supply chain event information (e.g. “Packet 88 was shipped in Honolulu at 12:31 UTC”) at time of occurrence and on item-level. Business process information flows can thus be automated, enabling supply chain participants to react to events in near real-time.

To allocate new or unknown services or look up changed service interfaces, the network system has to provide service discovery functions. The EPCglobal infrastructure provides core services to paying members of the EPCglobal network. One of the core services is the EPC Object Name Service (ONS) root. This service is administrated by the company VeriSign and is needed for search of EPC Information Services (EPCIS) offered by other participants of the network. The ONS root links address entries, by leveraging Internet Domain Name Service (DNS), to the manager master number of the organization, which provides the EPCIS. The holder of the master data number may operate local ONS and hierarchically link to further EPCIS or ONS. Non-certified EPCs or object identifiers of different formats cannot be processed by the EPCglobal network. Another core service is the Discovery Service. It is not completely specified yet, but will allow for search on item-level in the future. All core services are implemented and controlled by central authority of the EPCglobal organization. Disadvantages of the hierarchical structure and the centralized implementation of EPC core services are limited scalability and expandability, bottleneck and single-point-of-failure problems, unbalanced loads in the network, disabled self-organization and participants’ loss of control and power. The characteristics of the EPCglobal architecture may be one reason for low adoption of RFID supply chain systems. Companies refuse to join the exclusive club and to pay fees while the system still is in development (Kürschner et al. 2008).

Efforts of EPCglobal and others are under way to improve and amend the framework. An open RFID service discovery with an alternative EPCglobal-like ONS root was brought up by Afilias for discovery of participants’ local information services. Fabian and Günther (2007) elaborate security issues of EPCglobal and recommend geographically distributed name services. Other approaches bypass EPCglobal’s network by developing completely new frameworks. One of those is DIALOG, in which identification of objects follows the scheme ID@URI (Kärkkäinen & Ala-Risku & Främling 2003). Here the address (URI) of information resources is deposited directly on the object’s tag. This approach is quite flexible, but requires additional procedures for tagging, memory capacity on tags and administrating information resources stored at the URI.
2.2 RFID systems based on Peer-to-Peer networks

Requirements on scalability and efficiency of interorganizational RFID systems inspired efforts to create distributed solutions which support real decentralized coordination. Decentralized coordination is regarded as the more appropriate way to face challenges of managing complex supply chain processes in dynamic environments than centralized coordination (Amaral & Uzzi 2007). The concept of P2P networking is aligned to decentralized coordination in large networks of participants, who interconnect to exchange resources (Schoder & Fischbach & Schmitt 2005). Advantages of P2P systems for interorganizational applications are scalability, reduced risks of central bottlenecks, direct transactions, self-organization and failure resistance.

A special type of P2P networks, called structured or Distributed Hash Table (DHT) -based P2P networks, relies on hash table based communication protocols, which can guarantee efficient scalable and reliable information exchange (Risson & Moors 2006). In DHT-based P2P networks each information resource is stored as a key-value-pair and can be looked up with the key, which is an identifying name of the information resource (e.g. a file) mapped to a keyspace. Responsibility for the key-value-pairs is distributed fairly among peers. Each peer is responsible for a specified range of key-value-pairs and stores an address table with a confined number of known peers (neighbors) for routing purposes. Requests for storage or looking up of information are handled by routing those to a neighbor peer, whose hashed identifier is closest to the key in the keyspace. By this means it is ensured, that requests reach the responsible peer after a deterministic number of steps. Joining and leaving of peers is handled dynamically through intelligent repair mechanisms of address tables.

Summing up the current state, interorganizational RFID systems are centralized or proprietary and bear risk of performance weaknesses. Trying to face this, Wakayama et al. (2007) presented a product tracing system, which combines information service discovery using DNS according to EPCglobal with DHT-based service discovery. The major weakness of their approach is the chain-linking of address entries. If participants are not available, chains are broken and information can thus not be retrieved properly. Therefore, pursuing their idea, our architecture aims at efficient routing of information within a self-organizing, scalable, fully decentralized network of autonomous peers, without chain-links.

3 PROPOSED DESIGN OF THE P2P ARCHITECTURE

In this section we specify the design of our P2P architecture for service discovery of ubiquitous IT-based information services provided by participants within supply chains. We start with the description of object identification based on EPC or other identification schemata. Then we will illustrate the processes of providing and of querying information within the network. Lastly we explain the interoperability with EPCglobal.

3.1 Object identification

To identify an object tagged with RFID, a unique name has to be stored at the RFID tag memory. For RFID-based identification of objects the EPC provides a standardized name schema. For an open approach we regard it necessary to include any variants of identifiers in use or desired by applying companies, for instance hardware serials on RFID tags. A hashing function allows the publishing of information assigned to an object by using the hash value of any kind of object identifier. The hash value is mapped to a distinct location (one responsible peer and a predefined number of replication peers) in the network, from where participants can retrieve the entry directly by lookups using the object’s identifier as search key. Object identifiers and participants’ service addresses can be treated separately and multiple address links can be linked to one identifier.
3.2 Providing information to network participants

All RFID-based information generally remains in the participants’ local systems. Only meta-information is made publicly available through stored entries in the network. Two main entry types are distinguished: PublisherProfiles and PublicationRecords. A PublisherProfile defines the interface for an information service of a participant, who provides object information for other participants. It consists of the type of service (e.g. SOAP or EPCIS) and the service address (e.g. URI of the SOAP internet interface). Once an object identifier was captured by a company, it may publish information on that object within the network. For each object there is one PublicationRecord stored in the network. A record comprises a list of RecordEntries, and each RecordEntry consists of a reference to a PublisherProfile and optional information about the entry (e.g. type of event). Optional information is time of the capture event, object identifiers (children) that were aggregated to this object identifier (e.g. a container), a parent identifier, and additional, unspecified data. Each PublisherProfile and each RecordEntry has an expiration time, after which it will be deleted to prevent the network from storing too much or outdated data. Figure 1 shows classes and attributes of publication data, which is stored in the network. The segregation of profiles and records allows for separate update of each. In SCM systems records are potentially updated much more often than profiles. Furthermore, additional certificates or other authentication and authorization data may be stored in the profile.

A participant who wants to provide information sends a RecordEntry together with the object identifier to the P2P network. It will be replicated (depending on the configuration of the peers) and stored in the network. The participant’s profile will be published if it had not yet been published in the network or if he wants to update it.

3.3 Discovery and query of object information

The search for information is either accomplished directly with object identifiers or indirectly with queries, if a relevant identifier is unknown respectively is sought for or if other attributes are relevant for the search, e.g. date of entries. Direct search is simple and can efficiently be carried out in a DHT-based P2P network by taking the object identifiers as keys and looking up the keys in the DHT-keyspace. In contrast, complex query processing in P2P networks is a challenging task. Because in a fully decentralized DHT-based P2P network there is no participant who knows the entire network, an indirect query without an identifying key has to be resolved using distributed indices (Harren et al. 2002). If complex queries should be supported, indices have to be updated subsequently to each publication or periodically.

In addition to query processing the network also offers a publish/subscribe mechanism to realize event driven supply chain management. Information can then be pushed to interested participants immediately after capturing it. The mechanism allows participants to create a topic, subscribe to a topic or publish information linked to a topic, which automatically is sent to all subscribers of that topic. The topic subscriber group is also used for access and publishing restrictions by password.
Technically a topic corresponds with a sub-network of subscriber peers, which is in the same way self-organizing as the basic P2P network.

3.4 Interoperability

The proposed architecture offers interoperability with other systems like the EPCglobal network or ERP systems. Information to be published just has to be linked to any kind of object identifier and the publishing participant has to have or create a PublisherProfile. This means, the participant has to operate an information service at an address that is reachable by his supply chain partners. Figure 2 shows the components of the proposed architecture and how it could interoperate with EPCglobal. RFID reader events are processed by reader systems and middleware (which are provided e.g. by ERP software vendors) and then passed to our system as either binary data stream, XML or through data bases like mySQL. The system then extracts meta-information and publishes it in the DHT-based P2P network. On the other side, a query application can directly lookup meta-information of an object identifier or query meta-information through complex queries by sending a RecordEntry, where all determining attributes (e.g. event time) are filled and sought attributes are left blank. It then receives PublisherProfiles of participants which provide the appropriate information. By calling the service address given in the PublisherProfile, the query application can finally retrieve the information, e.g. from an EPCIS.

Figure 2. Components of the architecture and its interoperability with EPCglobal

4 DISCUSSION OF CONCEPTIONAL REQUIREMENTS

Conceptional requirements on supply chain RFID systems are manifold, depending on application context, supply chain and individual company needs. A detailed discussion regarding EPCglobal can be found in Kürschner et al. (2008). SME play a major role in implementing integrated supply chain RFID systems. On the one hand, as manufacturer, assembler or supplier of (pre-) products SME are in charge of tagging objects with RFID and of providing item-level information first (Gaukler & Seifert & Hausman 2007). On the other hand, SME can often act and operate very flexible, because of lean processes and organizational structures. Therefore, good alignment with innovative ubiquitous IT systems is attainable for SME. Despite that, SME still hesitate, e.g. in Germany 2007 only 3 % used RFID, but already 18 % of large companies (Martzahn & Renner & Pols 2007, p. 32). The following discussion will shed light on requirements of users, with focus on SME.

4.1 Scalability

Internet DNS root server process up to 48 billion queries per day (VeriSign 2008, p. 4). Upcoming item-level information systems will impose significantly higher volumes of information and queries. Furthermore, analogue to the Internet, benefits from RFID increase with amount of users due to
network externalities (Curtin & Kauffman & Riggins 2007, p. 95). Scalability is therefore most important for supply chain RFID systems. It ensures that companies are not restricted but can extensively interact with many business partners. Strong partnership networking is especially crucial for SME.

Based on DNS EPCglobal is organized hierarchically, and it is arguable whether centralized core services (like ONS root and discovery service) are capable of managing future loads. In a DHT P2P network there is no central bottleneck and with hashing algorithms information load is distributed equally. The search for entries with a DHT lookup receives a guaranteed answer even at high number of participants and communication loads. Also a limited response time can be guaranteed. Responses to searches and complex queries either contain the requested information (i.e. the appropriate PublicationRecords) or a Null-answer, if no one yet has published the sought information. DHT protocols like CAN, Chord, Pastry or Tapestry process lookups within $O(\log N)$ steps in networks with $N$ peers (Risson & Moors 2006). The required address tables for routing the lookups only occupy $O(\log N)$ storage space.

### 4.2 Security

Access to and authenticity of supply chain data, which participants insert into the network, should be secured. In centralized systems security mechanism can easily be implemented through authorization of participants and control of data by the authoritative organization. On the other hand, the central authority may abuse its power, because it can take over control about data streams and network access. Furthermore, it is an obvious target for attacks.

In distributed systems other security means are feasible. Authorization can be realized by a public key infrastructure (PKI) and a trusted third party, which offers and validates certificates of participants. An alternative solution is the web of trust (Datta & Hauswirth & Aberer 2003), where authentication of participants is confirmed not by a certifying authority but by other participants. Furthermore, data should be encrypted and transported over public infrastructures via secure protocols like SSL.

DHT-based P2P networks are self-organizing and therefore capable of self-healing after dynamic joinings and leavings of participants or other disturbances - up to some degree (depending on the configuration e.g. of replication mechanisms). Within the networks there is no single point-of-failure, which would be a potential target for internal or external attacks. As long as there is no large group of malicious actors, the network is robust. Besides that, all peers are autonomous and have equal power and control over their own data.

### 4.3 Flexibility

Existing IT infrastructures of SME may not be well equipped, which hampers exploitations of RFID benefits, e.g. because information is not utilized to advance internal information processes. Therefore solutions should be flexible and allow for customized implementation. Dynamic business relationships and supply chain structures should also not be absorbed by static network structures. The ability to quickly respond to changing situations is notably essential for SMEs.

Most interorganizational RFID systems are bind to fixed structures, like the hierarchical infrastructure of EPCglobal, and to special naming schemata for object identifiers, like EPC or ID@URI. But dynamically changing business environments and especially the internal flexibility of SME, e.g. in adapting to customer needs, require flexible systems. Therefore, our network is not bind to any fixed structure. It is capable to handle dynamic groups of participants, and any character string may serve as a key for information provided within the network. For instance, serialized data from XML documents can easily be exchanged in the network. In this way, existing e-business standards (like AS2, ebXML, eCl@ass) can be included with small efforts. This allows for individual configurations by participating companies when they implement such a system and integrate it into their existing SCM or ERP.
systems. The absence of a hierarchical structure supports flexible connections and enhancements. A distributed P2P system as proposed may be operated as a sub-network or a separate private network. In contrast, the EPCglobal architecture requires the establishment of dedicated instances, which operate core services and ensure unique naming.

### 4.4 Autonomy

Often SME are exposed to the power of major customers, who put pressure on them to implement systems suiting only customers’ needs. In the past years large companies, for instance Wal-Mart, mandated suppliers to adopt RFID (Soon & Gutierrez 2008). But sustainable adoption of ubiquitous IT will only be successful, if a company has internal motivation. Provided solutions should support SME in using RFID independently from others. This, of course, holds true for all other companies as well.

The participants of a P2P network are independent from each other and from any organizing instance. There is no central server, whose breakdown would block network transactions or become a bottleneck reducing efficiency and scalability. It is nearly impossible for individual instances to gain control over the network. Costs of implementation and operation of the systems are fairly distributed among all participants. There is no fee which has to be paid to an infrastructure operator. To lower the barriers, an interorganizational RFID system should allow for lean and incremental implementation. It should not cause or intensify dependencies on other organizations. Participants of EPCglobal are bind to the core service infrastructure, which is operated by a commercial US company (VeriSign), as well as to the hierarchical network and fixed data structure. Our architecture is compliant with EPC standards and can be used in combination with EPCglobal or other systems, but it is not bind to the EPC.

### 4.5 Costs and benefits

Major barriers for adoption of RFID are high investment expenditures and unclear benefits (Angeles 2005). Many cost-benefit-analyses are concentrating on costs of RFID tags and readers - but costs for tagging procedures, process reengineering and reorganization, system integration and user training have to be taken into account (Smith & Konsynski 2003). As manufacturer or supplier SME suffer from cost-benefit-asymmetry, because they bear the costs of tagging, while large retailers take credit e.g. for better customer service and lower inventory. In addition, SME are deterred by the complexity of system rollout and integration. After employment of RFID systems the flood of information has to be managed and that often requires well-planned process reengineering. SME traditionally lack capital, time and competencies to accomplish these tasks successfully. Assigning consulting firms or service providers adds even more to the cost-side.

EPCglobal grounds on a complex framework, initially driven by interests of retailers and large corporate groups. Each participant has to pay a periodic subscription fee. In contrast, an open P2P network will be free of charge, because there is no authority to charge it. Despite that, adoption costs and benefits are hard to be determined, because they strongly depend on individual design choices and organizational factors (Dutta & Lee & Whang 2007, p. 646). But within complex supply chains a P2P-based solution encourages a fair approach.

### 5 SIMULATION-BASED ANALYSIS

Supply chains are complex systems, whose investigation is known to be difficult in many respects (Harrison 2007). Under these circumstances computational simulations provide a good method for the study of supply chain systems (Terzi & Cavaliere 2004). Our simulation-based analysis may set an initial point for subsequent empirical studies. By this means we can test the large-scale performance of our P2P system with affordable efforts. After shortly explaining the prototype and setup we will discuss selected results.
5.1 Prototype

The prototype is based on the generic DHT-protocol Pastry (Rowstron & Druschel 2001a), which allows creation of a P2P network as a self-organizing, fault-tolerant overlay in the internet. Pastry routes messages between peers based on a 128-bit hash-value ring provides load balancing and leverages peers’ routing distances in the underlying internet for improved efficiency. Furthermore, we deploy the Pastry application PAST (Rowstron & Druschel 2001b) for hash-based, distributed storage and retrieval of data in Pastry networks. The prototype is implemented in Java and uses the open source Pastry implementation FreePastry1. In addition to the direct search with hashed object identifiers we added two exemplary types of queries: (1) Entries published by a specific participant, and (2) entries of a specific date.

5.2 Simulation setup

The simulation is conducted with a network of \( N \) participants. After bootstrapping the participants stochastically publish RecordEntries. Each participant publishes his PublisherProfile at first entry publishing. The probability of publications per preconfigured time interval \( \text{interval} \) is defined by a parameter \( p_{\text{pub}} \). Participants search (based on hashed EPCs) and query information in the network stochastically with probability \( p_{\text{look}} \) per interval. When receiving answers, they also lookup PublisherProfiles of received RecordEntries.

To examine the ability of flexible self-organizing, we introduce dynamic factors. The probability for leaving and joining of randomly picked participants is given by \( p_{\text{join}} \) and by \( p_{\text{leave}} \). The parameter \( rep \) determines the number of replications which are stored for each entry in the network.

In each simulation experiment we measured the actual dynamics \( \text{dyn} \) (numbers of participants leaving and joining the network per interval), the actual number of activities \( \text{act} \) (publications and lookups) per second, and failures \( \text{fail} \), and the average time till responses \( t_{\text{resp}} \) (excluding failed ones) in seconds. A publication or lookup activity is counted as failure, if there is no respond from the network after a preconfigured timeout (default 30 seconds). A Null-answer is not counted as failure, but as correct response.

5.3 Selected results

During the simulation runs we varied the input parameters for network size, replication, publishing and querying, joinings and leavings. The default time interval was set to 100 milliseconds. We run each experiment with at least 200,000 activities. Table 3 shows some selected results.

There were no significant differences in response times in different network sizes. The responses took 0.27 to 0.63 seconds on average. Moreover, failure rates were not significantly higher in larger networks. In small networks, e.g. with 500 participants, no failures appeared. Failure rates in low dynamic networks did not exceed 0.007 \%. In networks with dynamics of 5 \% of peers joining or leaving each minute failure rates increased slightly, but only up to 0.013 \% (experiment 7 in table 3). The replication factor naturally had an impact on failures. On the one hand, more replications led to lower failure rates, down to 0 \%. On the other hand, more storage capacity is used. But no observable influence on response times could be observed. Storage size for records ranged from less then 1 to a few 100 KB per peer. This obviously depends on amount and expiration time of published data. The used bandwidth for communication was not measured. Simulations of Haebeleren et al. (2005, p. 6) showed that the general overhead for Pastry consumed a bandwidth of 1-10 KB/s at maximum.

---

1 The software version is 2.0_04, downloaded on 26 May 2008. Online: http://freepastry.rice.edu/FreePastry/.
<table>
<thead>
<tr>
<th>N</th>
<th>rep</th>
<th>dyn</th>
<th>act</th>
<th>t.resp</th>
<th>fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>500</td>
<td>0</td>
<td>4.1%</td>
<td>3,303</td>
<td>0.38</td>
</tr>
<tr>
<td>2</td>
<td>500</td>
<td>2</td>
<td>1.7%</td>
<td>18,507</td>
<td>0.32</td>
</tr>
<tr>
<td>3</td>
<td>500</td>
<td>2</td>
<td>2.5%</td>
<td>6,110</td>
<td>0.36</td>
</tr>
<tr>
<td>4</td>
<td>1,000</td>
<td>1</td>
<td>1.1%</td>
<td>2,077</td>
<td>0.42</td>
</tr>
<tr>
<td>5</td>
<td>1,000</td>
<td>2</td>
<td>0.4%</td>
<td>925</td>
<td>0.52</td>
</tr>
<tr>
<td>6</td>
<td>1,000</td>
<td>2</td>
<td>3.8%</td>
<td>8,447</td>
<td>0.47</td>
</tr>
<tr>
<td>7</td>
<td>1,000</td>
<td>2</td>
<td>5.0%</td>
<td>7,072</td>
<td>0.58</td>
</tr>
<tr>
<td>8</td>
<td>1,000</td>
<td>3</td>
<td>4.2%</td>
<td>8,224</td>
<td>0.46</td>
</tr>
<tr>
<td>9</td>
<td>1,000</td>
<td>10</td>
<td>2.7%</td>
<td>4,085</td>
<td>0.46</td>
</tr>
<tr>
<td>10</td>
<td>10,000</td>
<td>2</td>
<td>0.5%</td>
<td>4,804</td>
<td>0.58</td>
</tr>
<tr>
<td>11</td>
<td>10,000</td>
<td>4</td>
<td>0.5%</td>
<td>7,634</td>
<td>0.54</td>
</tr>
<tr>
<td>12</td>
<td>10,000</td>
<td>10</td>
<td>0.6%</td>
<td>9,644</td>
<td>0.27</td>
</tr>
<tr>
<td>13</td>
<td>20,000</td>
<td>2</td>
<td>0.3%</td>
<td>6,829</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Table 3. Selected results of experiments

6 CONCLUSIONS

6.1 Lessons learned

The challenge of implementing RFID systems alongside the supply chain is not trivial and the design of appropriate network architectures is crucial. This article aimed at proposing a P2P architecture to meet general demands for ubiquitous IT systems and some special requirements of SME. To overcome limitations of existing approaches we introduced decentralization of supply chain RFID systems. Good scalability and flexibility are major advantages of decentralized architectures. Sophisticated DHT protocols provide a highly efficient base for distributed supply chain RFID systems.

We illustrated our DHT-based P2P architecture and explicated conceptional requirements with special regards to SME. Further on we presented results of simulation experiments with our prototype for large-scale networks. The results indicated short response times and self-organizing capabilities to handle dynamic participant structures as well as prevention of failures.

6.2 Limitations

This work is still in progress. The prototype is no fully-fledged software and real-life tests still have to be conducted. Advanced security mechanisms are not implemented, but Fabian and Günther (2007) describe P2P-based security mechanisms, that can be applied for supply chain RFID systems. The proposed architecture is not meant to replace existing solutions, but to enrich and amend their further development. With our work we want to give design hints for more fair and scalable implementations.

6.3 Recommendations

Companies should thoroughly analyze available supply chain RFID systems before they start implementation. If they do not want to be dependent on global solution providers and to constrict their flexibility, decentralized architectures should be considered in particular. Nevertheless, technical standards like the EPC have to be taken into account for any interorganizational RFID implementation. SME without in-depth RFID knowledge can use support of consultancies and research institutes, but may keep in mind, that more scalable solutions provide more future potential.
6.4 Future research

Next parameters of the simulation model will be refined on the basis of empirical data taken from analysis of information and material flows in real supply chains. Empirical evaluation and scientific discourse are further steps of our research. With a refined simulation setup and an advanced prototype, we may predict future information flows within RFID-based supply chains and assess the suitability of different architectures more precisely. It is worthwhile to find out, how structures and dynamics of supply chain systems are affected by large-scale adoption of ubiquitous IT. Several alternative solutions will be available concurrently. It is necessary to pursue research on the design and quality of various solutions.

References


Towards A Reference Model for Grassroots Enterprise Mashup Environments

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0585.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>Design Science, Commons-Based Peer Production / User centred content production, Reference modelling, Service oriented architecture (SOA)</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
TOWARDS A REFERENCE MODEL FOR GRASSROOTS ENTERPRISE MASHUP ENVIRONMENTS

Hoyer, Volker, University of St. Gallen (=mcminstitute) and SAP Research St. Gallen, Blumenbergplatz 9, 9000 St. Gallen, Switzerland, volker.hoyer@unisg.ch

Stanoevska-Slabeva, Katarina, University of St. Gallen (=mcminstitute), Blumenbergplatz 9, 9000 St. Gallen, Switzerland, katarina.stanoevska@unisg.ch

Abstract

A new kind of Web-based application, known as Enterprise Mashups, has gained momentum in the last years: Business users with no or limited programming skills are empowered to leverage in a collaborative manner user friendly building blocks and to combine and reuse existing Web-based resources within minutes to new value added applications in order to solve an individual and ad-hoc business problem. Current discussions of the Mashup paradigm in the scientific community are limited on technical aspects. The collaboration and the peer production management aspects of the Mashup development have received less attention yet. In this paper, we propose a reference model for Enterprise Mashups which provides a foundation to develop and to analyse grassroots Enterprise Mashup environments from a managerial and collaborative perspective. By following the design science research approach, we investigate existing reference models and leverage the St. Gallen Media Reference Model (MRM). The development of Enterprise Mashups is structured by market transaction phases similar to electronic markets. The user roles, the necessary processes and the resulting services are modelled according to the views of the MRM. By means of the SAP Research RoofTop Marketplace prototype we demonstrate the application of the designed reference model for grassroots Enterprise Mashups environments.

Keywords: Enterprise Mashups, St. Gallen Media Reference Model, Design Science, Electronic Markets.
1 INTRODUCTION

1.1 Motivation and Problem Scope

The process of development of Web-based business applications follows usually the typical process of software development involving first assessment of user requirements followed by a long process of development and testing. The functionality of the resulting application is actually a compromise of user requirements, as not all user requirements can be considered. As a result, there is a long tail – a term first coined and popularized by Chris Anderson (2004) – of many specific and heterogeneous user requirements or dynamically changing user requirements that are not covered by the IT department (Carrier et al. 2008, Hoyer and Stanojevska-Slabeva 2008).

A possible solution for this problem could be a new development paradigm, known as Enterprise Mashups, which integrates the users from the business units characterized by no programming skills in the software development process (Cherbakov et al. 2007). At the core of the paradigm are two aspects: First, empowerment of the end user to cover ad hoc and long tail needs by reuse and combination of existing software artefacts; and second, broad involvement of users based on the peer production concept. According to Yochai Benkler, who coined the term peer production, “it refers to production systems that depend on individual action that is self-selected and decentralized rather than hierarchically assigned” (Benkler 2006). Thereby, the creative energy of large number of people is used to react flexible on continuous and dynamic changes of the business environment. Instead of long-winded software development processes, existing and new applications are enhanced with interfaces and provided as user friendly building blocks.

Existing research efforts focus mostly on technical aspects like the development of Mashup tools – i.e., IBM Mashup Center (formerly IBM QED Wiki, IBM Mashup Hub, IBM Damia), Intel Mash Maker, Microsoft Popfy, and Kapow Mashup Server, which enable easy integration of available components - or underlying technical concepts and principles - i.e., Maximilien et al. (2008), Yu et al. (2008), Hoyer et al. (2008). The discussion from a collaborative and peer production perspective is still missing in the scientific community discussing the implications, challenges, but also the potential benefits and limitations of the Mashup paradigm in the enterprise context. Important questions from these perspectives are: Who is involved in the Mashup development and what are the roles of the different players in particular of the software development department and the business units? What are the necessary processes to enable and support community building and collaboration?

The goal of this research paper is to fill this gap by designing a reference model which integrates the collaborative and community aspects and can serve as a framework to structure the development and analysis of Enterprise Mashup environments. The general research questions guiding this study are to model the roles and relationship of the interacting users as well as to interlink community and technical aspects within grassroots Enterprise Mashup environments.

1.2 Research Design: Design Science applied

For answering the research questions motivated in the previous section and characterized by a practical nature, engaged research is needed in order to provide rigorous solutions. Design science research aims at solving practical and theoretical problems by creating and evaluating IT artefacts indented to solve identified organizational problems (March and Smith, 1995, Hevner et al. 2004, Peffers et al. 2008). Artefacts represent the final result of a design process. They can be characterised as constructs, model, methods, or instantiations (March and Smith, 1995).

To come to rigorous and relevant research results, we draw upon on Peffers et al. (2008) to specify the subsequent phases of the design science research process applied:
• **Problem Identification and Motivation.** In the motivation, we specify the specific research problem, show the practical relevance and justify the value of a solution. Based on the problem scope, we derive the research questions guiding this paper.

• **Define the Objectives for a Solution.** In the second chapter, we infer the objectives of a solution from the problem definition and knowledge of the state of art. In particular, chapter two defines Enterprise Mashups, analyses existing reference models by means of a literature review and presents the St. Gallen Media Reference Model.

• **Design and Development.** In chapter three, we leverage the St. Gallen Media Reference Model (MRM). As observed by Legner (2008), Hoyer and Stanoevkska (2008), and Carrier (2008) the peer production of Enterprise Mashups has many similarities to electronic markets: Available components are classified and offered by providers and potential consumers search for the most suitable ones and if required pay for the usage. Thus, the required support should enable the matching of supply and demand in a way similar to conventional market phases (knowledge, intention, contract/design, and settlement). We apply these market phases to model the roles and the relationships between the interacting users according to the layers of the MRM.

• **Demonstration.** For demonstration of the designed artefact, we apply the reference model in order to develop the SAP Research Rooftop Marketplace. In particular, we structured the requirement and software design phases according to the reference model for Enterprise Mashups.

The results of each of the above activities are presented in the remaining parts of the paper. Finally, the last chapter closes with a brief summary, limitations of the conducted research and an outlook to further research.

2 RELATED WORK AND OBJECTIVES OF THE SOLUTION

2.1 Enterprise Mashups – Definition and Characteristics

In literature, the exact definition of Enterprise Mashups is open to debate. In this work, we refer to the following definition: “An Enterprise Mashup is a Web-based resource that combines existing resources, be it content, data or application functionality, from more than one resource in enterprise environments by empowering the actual end users to create and adapt individual information centric and situational applications” (Hoyer et al. 2008). By simplifying concepts of Service-Oriented Architecture (SOA) and by enhancing them with the Web 2.0 Philosophy of peer production, Enterprise Mashups focus generally on software integration on the user interface level instead of traditional application or data integration approaches (Daniel et al. 2008). In contrast to SOA that is characterized by high technical complexity of the relevant standards and requiring specialists’ technical knowledge, the simplified Enterprise Mashups enable the integration of the end users with no or limited programming skills in the development process.

The relevant architectural components of the Enterprise Mashup paradigm are resources, widgets, and Mashups (Hoyer et al. 2008) and can be structured in an Enterprise Mashup Stack comprising three layers (see figure 1): **Resources** represent actual contents, data or application functionality that are the core building blocks of Mashups. They are encapsulated via well-defined public interfaces (Application Programming Interfaces; i.e., WSDL, RSS, Atom, etc.) allowing the loosely coupling of existing resources – a major quality stemming from the SOA paradigm. These resources are provided by enterprise systems or by external Web providers (i.e., Amazon, Google, etc.) and are created by traditional developers who are familiar with the technical development concepts. The layer above contains **widgets** which provide simple user interaction mechanism abstracting from the complexity of the underlying resources. For example a widget “Customer Data” might provide results for a predefined query requesting the data for all customers of a sales manager. The creation of these widgets can be done by consultants or key users in the business units who understand the business requirements and know basic development concepts. Finally, end users with no programming skills...
are able to combine and configure such visual widgets according to their individual needs, which results in a **Mashup**. For example, the sales manager wires the "Customer Data" with a map to show the location of the customers.

![Enterprise Mashup Stack (Hoyer et al. 2008) - Meta Model and User Roles](image)

**Figure 1.** Enterprise Mashup Stack (Hoyer et al. 2008) - Meta Model and User Roles

The first key driver of the Enterprise Mashup paradigm is the **lightweight composition style** by reusing existing building blocks in new ways. The Enterprise Mashups paradigm separates between the wiring and piping composition as depicted in Figure 1. The piping composition integrates heterogeneous resources defining composed processing data chains concatenating successive resources. Aggregation, transformation, filter, or sort operations adapt, mix, and manipulate the content of the underlying resources. The visual composition of input and output parameters of widgets on the Mashup layer is called wiring (i.e. the output parameter address of the customer widget is connected to the input parameter of the map widget). In addition to this lightweight composition style, the **mass collaboration principle** is the second key driver. The willingness of users to offer feedback to the Mashup creator, who may be unaware of problems or alternative uses, directly contributes to the adoption of the Mashup and can foster its ongoing improvement (Hoyer et al. 2008). Another important contribution of users is the inclusion of their Mashups in the available pool of components. The willingness of users to provide their Mashups for further reuse increases the number of available components.

### 2.2 Reference Modelling

Despite the popularity of the term reference modelling since the 1990s, there is still no single meaning connected to this term and it is used to designate different approaches. By analysing various definitions, Fettke and Loos (2007) identified the following three main characteristics of reference models: First, **best practices**. A reference model provides best practices for conducting business. Second, **universal applicability**. A reference model does not represent a particular enterprise or system, but a class of domains. Hence a reference model is valid for a class of domains. Third, **reusability**: They can be understood as blueprints for information systems development. Thus a reference model is a conceptual framework that could be reused in multitude of information system projects. Modeling guidelines (Becker et al. 1995) and evaluation criteria of reference models (Frank 2007) are discussed in the scientific community.

In general there are two approaches for creating reference models: either by observing many instances available in practice and extracting common elements into a reference model or by leveraging and adjustment of existing reference models. The first approach is suitable when a sufficient number of instances are available. The second approach might be a suitable option when the underlying phenomenon is not well researched yet, but similarities to other phenomena can be revealed. As
Enterprise Mashups are a recent development there are not proven good practices yet that can be applied for reference model extraction. Thus, the second approach was applied: existing reference models were checked if they fulfil the requirements of Enterprise Mashups environments.

By means of a literature review and by applying the classification framework of Braun and Esswein (2007), we analysed existing reference models that are relevant for the Enterprise Mashups paradigm. Gartner proposes a practitioner reference model that specifies the technical architecture components in Enterprise Mashup environments (Bradely 2007). A practitioner reference model of Forrester uses a similar layer structure like Gartner and the presented Enterprise Mashup Stack. In addition, a phase model is integrated specifying the inputs and information flow (Young 2008). First, the actual content provided by the IT department is provisioned for the Enterprise Mashup environment from both internal and external resources. Second, users from the business units use a so-called Mashup composer to arrange and combine content, as well as determine a visualization paradigm. Third, the mashable components are managed by a Mashup life-cycle manager and shared with others to use in new Mashups if desired. Even though both reference models provide first technical structures of Enterprise Mashup environments, a multi-view concept integrating the managerial perspective is missing. The existing reference models miss in particular support for the collaborative aspect of Enterprise Mashups development and do not provide sufficient support for the peer production process. In order to integrate the different aspects (community, processes, or technical), a multi-view reference model is necessary.

Another specific characteristic of Enterprise Mashup environments is their similarity with electronic markets. Enterprise Mashup environments need to provide besides support for easy integration of software artefacts also support for efficient management and matching of supply and demand for Mashup components. Legner (2007) describes the trading of Web Services according to market transaction phases. Carrier et al. (2008) put the discovery and sharing of mashable elements in the center of the development process to reuse existing assets in new combinations.

In summary, a comprehensive reference model for Enterprise Mashups is required that on the one hand considers technical requirements regarding the easy integration of mashable components and on the other hand support for matching of supply and demand for required Mashups based on the market paradigm. We will incorporate these findings by leveraging a scientific multi-view reference model (The St. Gallen Media Reference Model) that has its roots in electronic markets and that has already been applied successfully for modeling electronic markets (Schmidt and Lindemann 1998), m-commerce communities (Stanoevska-Slabeva 2003), and cross-company electronic collaborations (Schroth and Schmid 2009).

### 2.3 St. Gallen Media Reference Model

The St. Gallen Media Reference Model (Schmidt and Lindemann 1998) provides a framework for specifying IT infrastructures. Under the term medium, we understand platforms based on information and communication technologies, i.e., communication spaces of "social interaction which allow the participant to meet and which embed them in a common physical, logical, and socio-organizational structure" (Schmidt 1997). The media reference consists of two dimensions: The horizontal dimension contains the four phases of a market transaction whereas the vertical dimension is built of four views.

The four view layers structures the different successive interaction goals of the participating agents. The **community view** describes the participating agents, their roles and the organizational structure defining the relationships among roles together with their obligations and rights. The **interaction view** refers to the relevant processes and is based upon the underlying services. The **service view** comprises all services in the four market phases that need to be available on the platform. The four services are: First, the **knowledge phase** is which information about offered services and knowledge and the media platform itself is acquired. Second, the **intention phase** in which agents signal their intentions in terms of offers and demand. Third, the **contract phase** where agents negotiate legal binding contracts and finally the **settlement phase**, in which agents act according to the negotiated contract using the
platform’s settlement services offered for this purpose. Examples of services in the service view are electronic product catalogs in the knowledge phase or payment services in the settlement market phase. Finally, the infrastructure view contains communication protocols and standards which comprise the groundwork for the implementation of services.

3 DESIGN: REFERENCE MODEL FOR ENTERPRISE MASHUPS

As elaborated in the previous chapter, we leverage and adjust the existing St. Gallen Media Reference Model due to its similarities to electronic markets and due to their successful application for managing communities. The driving force is the transfer of typical market transaction phases to the development of software artefacts to address the specific requirements of Enterprise Mashups. In addition to the two dimensions (views and phases), we introduce the architectural Enterprise Mashup Stack as a third dimension of the reference model for Enterprise Mashups as depicted in Figure 2.

Starting with the knowledge phase, available mashable components (Mashup, widget, resource) are classified, rated and explained in different ways to the agents of the Enterprise Mashup environment. Concepts from the Web 2.0 philosophy, like rating, tagging, or recommending are integrated for browsing through the Enterprise Mashup medium. During the intention phase, the concrete offers are provided in a structured manner including the payment mode, the price as well as the delivery conditions. In the contract (design) phase, users select the right mashable component based on the provided information, configure it according to their preferences and combine it with other components. Finally, in the settlement phase the Enterprise Mashup is executed.

In the following, each view of the model is described and modelled by using the well-known conceptual modelling languages Unified Modeling Language (UML), Business Process Modelling Notation (BPMN) and Fundamental Modelling Concepts (FMC).

3.1 Community View

To describe the interacting and connected agents as well as their tasks and roles, we refer to the following interaction model well known in Service-Oriented Architectures (Papazoglou 2003) but also in electronic markets (Legner 2007): A provider develops and publishes a mashable component via an intermediary, where a consumer can find it and subsequently may compose and consume it. As
depicted in Figure 4, the interaction between consumers and providers is always managed by the intermediary. The tasks of the three agent roles are described in the following:

- **Provider.** A provider implements and hosts a Mashup component (resource, widget, or Mashup) which encapsulates the actual content or knowledge. To promote their provided functionalities, the provider annotates the component with relevant information and publishes it to the intermediary through which the component description is published and made discoverable.

- **Intermediary.** An intermediary mediates between providers and consumers similar to electronic marketplaces (Legner 2007). In contrast to traditional SOA-based implementations like UDDI or ebXML, novel forms of intermediaries are currently about to emerge which improve navigation, transparency, and governance. They monitor continuously the parameters (such as availability or response latency) and provide performance metrics and other evaluation results which may be used by the consumers to select the right Mashup component.

- **Consumer.** Based on the available information of a mashable component, a consumer is able to retrieve and compose Mashup components according to his individual preferences. Consumers take also over the role of annotating Mashup components by tagging, recommending, or rating them. The consumers also contribute to the community base of widgets by providing their created and adapted Mashups back in the community pool.

![Figure 3. Agents in the Enterprise Mashup Community.](image)

A critical success factor for Enterprise Mashups is a broad potential user group, familiar with the technology and willing to use it in their daily operational environment. Users from the business units are able to create their own component and provide it to the community (Hoyer and Stanoevska-Slabeva 2008, Carrier et al. 2008). In this sense, users can take over both agent roles (consumer and provider). For example, a sales manager can publish his created Mashup combining different widgets (i.e., the customer and map widget) to the community. One of his colleagues can reuse and consume the Mashup immediately without the involvement of the IT department. The creation of more complex widgets and the adaptation of existing backend services are in responsibility of the developer from the IT department (provider role). Additionally, the IT department takes over the intermediary role. The Mashup components itself are located within the enterprise boundaries (i.e., my sales orders) or are sourced via external Web providers. The foundation of the richness of Enterprise Mashups applications is based on the seamless combination of corporate internal and external information sources. News feeds, a map or a credit cart checking Web Services can be combined ad-hoc with the internal customer data. The continuous growing numbers of available Web-based resources can be observed by the two intermediaries Programmableweb.com or Seedka.com.

### 3.2 Interaction View

Figure 4 depicts a simplified process in BPMN describing the interaction between the three agent roles as presented in the community view before. The detailed interaction process can be found in Hoyer and Stanoevska-Slabeva (2009). According to the findings of section two, the interaction process itself is characterized by permanent loops between the converging design and runtime phases. The need to
adapt the operational environment ad-hoc leads to adding, removing, or replacing existing mashable components. In the following discussion of this paper, we focus on the Mashup layer.

<table>
<thead>
<tr>
<th>Consumer</th>
<th>Intermediary</th>
<th>Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery community, mentor, features, etc.</td>
<td>Aggregate information about users and Widgets</td>
<td>Identify new trends, evaluate new widgets</td>
</tr>
<tr>
<td>Pre Production (Tagging, rating, sharing, etc.)</td>
<td>Providing widgets on the Mashup environment</td>
<td>Developing widgets, defining business model</td>
</tr>
<tr>
<td>Searching for Widgets</td>
<td>Discovery Catalogue Repository</td>
<td>Publishing Widgets</td>
</tr>
<tr>
<td>Accepting business model (payment, licence)</td>
<td>Identifying customer needs</td>
<td>Promoting Widgets</td>
</tr>
<tr>
<td>Composition</td>
<td>Composition matching (interoperability aspects)</td>
<td>Commitment for using widget</td>
</tr>
<tr>
<td>Consumption Widget</td>
<td>Monitoring widget consumption</td>
<td>Providing Widget</td>
</tr>
<tr>
<td>Payment Widget</td>
<td>Billing Widget</td>
<td>Billing Mashup environment usage</td>
</tr>
</tbody>
</table>

Figure 4. Interactions between the three Agent Roles (Mashup Layer).

After registering to the Enterprise Mashup environment, consumers (i.e. the sales manager) are able to discover the community and member profiles. By means of examples in form of short videos, the benefits of the Enterprise Mashup environment are demonstrated to potential users. Only if a huge amount of users are convinced of using the environment, it will exploit its actual potential. By discovering the catalogue of mashable components (in this case widgets), consumers are able to select a widget based on extensive information provided by the provider, intermediary as well as the consumers. Reviews, recommendations, and ratings of colleagues help for selecting a component. In case the consumer accepts the underlying business model (costs, payment model, consumption licence, etc.) of a widget that is defined by the provider, he can compose the component with others by connecting the input and output parameters of the widgets. In contrast to the classical software development, the design of ad-hoc applications uses real resources and no demo systems. In this sense the consumption in the settlement phase differs only from the hidden configuration capability in contrast to the design phase. In case a new business situation comes up, the consumer shifts quickly to the design or intention perspective to adapt the individual operational environment.

3.3 Service View

Based on the described interaction process, we derived the required services for the process steps. According to the role of the intermediary mediating between consumers and providers, the IT department is responsible to act as a service intermediary (Hoyer and Stanoevska-Slabeva 2008). Because business users focus on solving daily business problems in the sales or accounting department and not on creating or adapting their operational environment, Enterprise Mashup platforms have to hide the complexity from the users. The figure below depicts the related services to implement the interaction process by using the Fundamental Modelling Concepts (FMC) notation. In contrast to the technical oriented UML notation, FMC focuses on human comprehension of complex systems.

The growing number of available mashupable components requires adequate discovery services for retrieval purposes. According to the user context (profile, preferences, social network it belongs to) relevant widgets are presented to the consumers who are able to select the relevant Mashup component. Sharing of information, experiences and knowledge within the community is a key driver for Enterprise Mashups. Besides the default semantic annotations (functional and non-functional

---

1 http://www.fmc-modeling.org
qualities) defined by the provider, consumers are able to tag, recommend, or rate the mashable components. A catalogue manages all this information for widgets, Mashups, and users. The design of the widget components (wiring) is handled by the composition service that accesses the catalogue for the required information. As mentioned already before, good enough solutions lead to a converging design and run time. That implicates a direct integration between thes two services. From consumer perspective, no traditional deployment exists. They design their operational environment and execute it immediately. During the execution phase, the Enterprise Mashup environment monitors and protocols the usage of the mashable components. Based on the aggregated statistics resulting in popularity, availability or error rate information, consumers and providers get additional information to select or adapt a component or to publish a new one. Additionally, the accounting of the environment usage is calculated by this data. To administrate and monitor the performance of the Enterprise Mashup environment, the IT department needs adequate services to manage the running system.

Figure 5. Services of Enterprise Mashup Environments (FMC Notation).

3.4 Infrastructure View

In contrast to existing applications (i.e., MS Excel or MS Access) created and managed by business units to address ad-hoc requirements, the technical infrastructure of Enterprise Mashups environments are managed and provided by the corporate IT department. Independent if the users from the business units act as consultant (widget layer) or as end user (Mashup layer), they are able to integrate easily their local resources encapsulated by Mashup components. Consumer-oriented Mashup environments like iGoogle, Netvibes, or Facebook consume mostly light weight Web-based resources sourced via RSS, ATOM, or JSON. Instead Enterprise Mashup environments integrate resources from legacy systems as well. Currently, the major vendors of enterprise systems like SAP, Oracle, or Microsoft enable their applications to service-oriented platforms which are based on established standards. For example, SAP provides more than 1500 so-called enterprise services described by WSDL. However, besides standardized Web Services, the implementation of first Mashup prototypes in Enterprise Mashup environments shows the relevance to integrate other resource types as well, i.e. XML files, data bases, or RPC.

Due to the open and Web-based character of Enterprise Mashups, wide accepted standards or protocols are required. This includes technical standards for the visual composition (piping and wiring) but also standardized accounting methods.
4 DEMONSTRATION – SAP RESEARCH ROOFTOP MARKETPLACE

This section is devoted to apply and demonstrate the design artefact by means of the SAP Research Rooftop Marketplace platform (Hoyer et al, 2009). It represents a prototype which allows the creation and adaptation of Enterprise Mashups according to the individual and heterogeneous needs. We used the designed reference model to develop the prototype by following the marketplace and collaborative character as elaborated in this paper. Thereby, the reference model structured the requirement analysis and the transfer to the technical specification of the platform.

![SAP Research Rooftop Marketplace](image)

**Figure 6. SAP Research Rooftop Marketplace.**

The SAP Research Rooftop Marketplace itself is a Web-based application based on AJAX (Asynchronous JavaScript and XML) and is internally available to all SAP employees without installation. In this sense, the platform is provided by SAP Research acting as an intermediary. By using a Single Sign On (SSO) login process, users (provider and consumer) are able to register to SAP Research Rooftop automatically. After defining the individual user profile, users are able to discover the Mashup community which provides information about Enterprise Mashups in general as well as features and demonstrations (videos) of the SAP Research Rooftop tool in specific. A catalogue allows the discovery of the provided Mashups and widgets. Besides the browsing of predefined categories, the user is able to select top rated, most popular or latest widgets. Each widget is annotated by extensive information provided by all three agent roles (see figure 6). The actual consumer is able to rate, to tag, or to recommend a widget. In case a user has created an Enterprise Mashup, he is able to send a link to his colleague who can consume immediately the component. On the other side, providers can specify a default description and the business model (i.e. the costs) of a widget. The SAP Research Rooftop Marketplace platform itself (intermediary) monitors the consumed widgets and Mashups continuously (i.e., popularity and availability) and provides the information to the consumer and provider.
The composition of widgets is depicted at the bottom of the figure above. The customer data widget is wired with a map by connecting the input and output parameter (orange line) represented by the red bullets on the left (input) and the blue bullets right (output) side of the widget. The configuration and information of a widget can be easily accessed by clicking on the info button of the widget. It allows that the consumer can directly contribute to the community by rating or tagging a widget without changing to a different view which covers the knowledge/intention phase. According to the described marketplace and collaborative characteristic of Enterprise Mashups, the SAP Research Rooftop Marketplace platform integrates these two market phases. The figure above at the bottom right indicates the available information of the “Customer Data” widget and how a consumer can add easily a rating and comment to a widget.

Coming back to the composition environment, the real data of the “Customer Data” widgets are displayed already at the design time as depicted in the figure. If the user selects a customer (Siemens in Munich, Germany) in the widget, the address is updated in the map (in this case Microsoft Virtual Earth). So, there exist no separation between the design and runtime within the SAP Research Rooftop Marketplace prototype. By shifting to the runtime view, only the configuration capabilities like adding new widgets are logged.

5 CONCLUSION

The aim of this paper is the design of a reference model for grassroots Enterprise Mashups environments serving as foundation to develop and structure Enterprise Mashup environments. In order to achieve this, we follow the design science methodology. After defining the main terms related to Enterprise Mashups and reference models, we presented a designed reference model for Enterprise Mashups by leveraging the St. Gallen Media Reference Model. Thereby, we took advantage of the observed and identified similarities to electronic markets and collaborative characteristics. Instead of following the traditional software development phases (requirements, specification, development, testing and deployment), we propose the structure of market transactions. The roles, required services, and relationships between the interacting agents (provider, consumer and intermediary) were modelled. By means of the SAP Research Rooftop Marketplace, we demonstrate the application of the designed reference model for structuring the development of an Enterprise Mashup environment.

What is still missing is a broader application of the reference model for Enterprise Mashups in other areas and its further verification in an iterative design cycle (“Design as a search process”) according to the design science methodology (Peffers et al. 2008, Hevner et al. 2004). Further research will deal with the application of the reference model to structure further Enterprise Mashup environments. In particular, we have applied the reference model in the frame of the EU funded FAST/EzWeb project (http://fast.morfeo-project.eu) that covers the Mashup and widget layer.

The designed reference model for Enterprise Mashups provides furthermore only a first generic framework that helps to understand the organizational and managerial challenges of the Mashup paradigm in enterprise environments. Besides the structuring of requirements for the development of Enterprise Mashup platforms, the generic reference model has to be extended and operationalized with unhandled managerial aspects. The loosely coupled user-friendly building blocks both from internal and external IT systems require also a governance, quality, and security concept defining who is able to access a widget. The inclusion of these aspects would provide a more detailed and different models of the community and interaction view.

References


## HOW FACTORS AFFECTING SELECTION OF IMPLEMENTATION APPROACH INFLUENCE ERP SYSTEM IMPLEMENTATION COSTS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0594.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Enterprise resource planning (ERP) (packaged systems), IS Investments, Implementation, Empirical study</td>
</tr>
</tbody>
</table>
HOW FACTORS AFFECTING SELECTION OF IMPLEMENTATION APPROACH INFLUENCE ERP SYSTEM IMPLEMENTATION COSTS

Johansson, Björn, Copenhagen Business School, Center for Applied ICT, Howitzvej 60, DK-2000 Frederiksberg, Denmark, bj.caict@cbs.dk
Sudzina, Frantisek, Copenhagen Business School, Center for Applied ICT, Howitzvej 60, DK-2000 Frederiksberg, Denmark, fs.caict@cbs.dk

Abstract

Different approaches on how to implement or deploy enterprise resource planning (ERPs) systems exist. Although virtually nobody really doubts importance of ERPs for running a business today, there is a sentiment regarding their implementation – both in terms of time and money. In this paper we investigate relationship between factors influencing selection of a specific implementation approach and companies’ ability to stay on budget when implementing ERPs. The question is: whether factors influencing implementation approach then affects to what extent ERP system implementation costs exceed planned costs for implementation. The questionnaire research, focused on this issue, was conducted in Denmark, Slovakia and Slovenia. Dependent variables were percentage of actual ERP system implementation costs and staying on budget vis-à-vis the planned costs and budgets. The independent variables were implementation approach, country, company size, information strategy, representation of the IT department on board level, and number of implemented modules. Main conclusions are that number of modules influences selection of implementation approach and companies with information strategy are more likely to stay on budget. However, implementation approach does not significantly influences implementation costs and clear relationships between factors influencing selection of implementation approach and costs for ERP implementation could not be found.

Keywords: Enterprise Resource Planning (ERP), IS Investments, Implementation, Empirical Study.
1 INTRODUCTION

Enterprise resource planning (ERPs) systems consist of an integrated set of programs that provide support for core business processes, such as production, input and output logistics, finance and accounting, sales and marketing, and human resources. An ERP system helps different parts of an organization to share data, information to reduce costs, and to improve management of business processes (Aladwani, 2001). Wier, Hunton, and HassabElnaby (2007) argue that ERP systems aim to integrate business processes and ICT into a synchronized suite of procedures, applications and metrics which goes over firms’ boundaries. It can be stated that the success of ERPs to a high extent depends on its implementation. It can also be stated that implementation or as often also labelled deployment is a complex and costly endeavour.

So, although virtually nobody really doubts their importance for running business, there is a sentiment regarding ERP implementation – both in terms of time and money. Cunningham (1999) reports from an investigation of 7500 IT projects conducted by Standish Group in 1998 which discovered that 45% of them were late or over budget. According to the only publicly accessible Standish Group report on ERP implementation projects (Standish Group, 1995) actual cost of projects was, on average, 214% of what small companies planned, 182% of what medium companies planned, and 178% of what large companies planned and took 2.39 times longer than small companies intended, and 2.02 times longer than medium companies intended, and 2.30 times longer than medium companies intended. There are, definitely, several contributing factors to staying on budget and on time. However, in this paper, the question is if different approaches for implementation result in different outcomes when it comes to costs for ERP implementation.

The research question is whether factors influencing implementation approach then affect to what extent ERP system implementation costs exceed planned costs for implementation in European context, which is characterized by, for instance, fixed price policy. It can be stated that cost of implementation is an important part of total cost of ownership (TCO), and therefore it is important to know how large is the disparity between actual and planned total cost of implementation of ERP systems, and how different factors influences costs for ERP implementation. The rest of the paper is as follows: the next section discusses ERP system implementation approaches. The section after that describes the research method and how data were collected and analyzed. The following presents the results of the analysis regarding relationships between the variables and whether the organization stays on budget or not respectively what percentage they spent on implementation related to the budget. The penultimate section then discusses limitations and suggests future research. Finally some conclusions are presented.

2 ERP SYSTEM IMPLEMENTATION APPROACHES

In the paper, implementation is defined as the way how organizations systematically integrate ERPs into the specific organization. This can be done in different ways and that is what we mean with implementation approach. Implementation approach is defined as a systematically structured approach that aims at integrate selected ERP system into the workflow of an organizational structure (http://en.wikipedia.org/wiki/Product_software_implementation_method) One way to distinguish between different implementation approaches is to look into changes in the organization and when these changes take place. This can be described as piecemeal versus concerted implementation (Robey, Ross and Boudreau, 2002). The difference between these two is that in the piecemeal implementation the ERP is implemented first and then changes in the organizations business processes are implemented. The concerted implementation approach means that the ERP and changes in the business processes are implemented at the same time. These different approaches could be related to IT/IS strategy and it could be suggested that if the organization has a formalized information strategy
that probably influences what ERP implementation approach that the organization select. It can also be related to business process reengineering (BPR) (Davenport, 1993; Hammer and Champy, 1993) which has a clear focus on restructuring both the organizational structure as well as the used information system (IS) structure, and it can be stated that this makes the change from the earlier structure of legacy system complex. It is, therefore, interesting to see what influences organizations when they select a specific approach for implementation.

McGillcuddy (2007) states that there is a difference between size of the organization when it comes to the time it takes between the organization starts to implement an ERP to its implemented ERPs go live. The claim is that small businesses have a shorter time than midsized and large organizations. This statement builds on data presented in a report from the Aberdeen Group. In that report it is said that 86% of small enterprises achieved their first go live milestone within the first year, in midsized enterprises the same happened in 64% of the implementing enterprises and when it comes to large enterprises just 47% of them reported that they experienced the first go live milestone within a year. It could be asked if this means that small organizations more often implement ERPs as a big bang approach and that the bigger the organization is, more likely they select a slow phased implementation approach.

In this study, we distinguish between three types of implementation approaches: slow phased, pilot project and big bang implementation. In the literature, there exist two general approaches for how ERPs are implemented, which were popularized in the mid-1990s (Mabert, Soni and Venkataraman, 2003; Markus, Tanis and Van Fenema, 2000b): (1) the “big bang” approach and (2) the phased implementation approach. The “big bang” is an implementation approach that means that the entire organization starts to use the new ERP at the same time. The big bang probably has been planned for a long time and the specific ERP have been adjusted and to some extent tested before the actual big bang, but, what happens is that the organization decides on a specific date for when the ERP should be taken into usage. When that specific day then comes, data are transferred from the old legacy system and all users start to use the new system. This can then be compared to the phased implementation approach. The differences between these are that the phased means that some parts of the organization start to use the new ERP and after a while the next part starts to use it and so on. The phased implementation can be phased in different ways, it could be that, if the organization is situated at different locations, a specific location starts, or it could be that a specific user group starts and so on. The major difference between these two approaches is probably the time it takes. The big bang approach means definitely a shorter time for the roll-out in the entire organization. The phased implementation approach takes longer time, but it is not sure that it takes so much longer time from the first decision on adoption of a new ERP to the time it is in full use. It could be that the big bang implementation approach demands a longer time period for preparing before the big bang. However, it can be stated that although phased implementation is time consuming, it involves less risk compared to the “big bang” approach (Scott and Vessey, 2000). Recent research has also revealed that the phased implementation tends to involve less reengineering efforts.

Parr and Shanks (2000) state that there is a need to further describe implementation approaches into a taxonomy if being able to investigate ERP implementation. They suggest a taxonomy describing three different implementation approaches which they label: Comprehensive, Middle Road and Vanilla implementation. However, in our view this categorization is more related to earlier decisions such as deciding on what ERP package to adopt and/or deciding on if going for “best practices”. But the taxonomy suggested by Parr and Shanks have an interesting further categorization when they talk about characteristics related to each approach in the framework. The characteristics are: 1) physical scope, which means if implementation is made at several places, 2) BPR scope, which consider to what extent reengineering is considered, 3) technical scope, which is about to what extent the adopted ERP is modified, 4) module implementation strategy, considering two different strategies for implementation of ERPs modules, 5) resource scope, which is about the time and budget scope for the implementation. In this paper, the most interesting characteristic from Parr and Shanks to investigate further is the module implementation strategy. What they state about this is that there exist broadly
two different decision points in the module implementation strategy. The first decision is about whether the ERP should be implemented as a skeleton or with full functionality and the second decision is then if the implementation should be done module by module integration to legacy systems or all ERP modules implemented and then integrated to legacy systems (Parr et al., 2000). The latest described approach – all ERP modules implemented – can be compared to big bang implementation while the other one could be compared with phased implementation. Parr and Shanks state that phased implementation, is less risky, but more resource intensive, while the big bang implementation is precarious but a less time consuming option. According to Basoglu et al. (2007), big bang implementation creates adoption problems in the long run, and the reason they state for this is that organizations, when implementing big bang, spend less effort in adjusting the software and the organization to each other. Because of the advantages of a phased implementation, it was of our interest also to figure out how exactly companies approach this issue and what it is that makes an organization select a specific implementation approach. This and the inspiration from (Bernroider and Leseure, 2005) was the reason for splitting the phased implementation into slow phased-in implementation approach (one module at a time) and a pilot project implementing (one module followed by all other modules in one step). Although one could try to divide big bang implementation into two, as e.g. (Madapusi and D'Souza, 2005) did, it could also confuse respondents.

3 DATA AND METHODOLOGY

This exploratory paper is based on a questionnaire survey, conducted in Denmark, Slovakia and Slovenia in May and June 2007. Questionnaire forms accompanied by cover letters were mailed to randomly selected companies. Lists of addresses and information about the number of employees were retrieved from CD-Direct in Denmark, and from respective Statistical Bureaus in Slovakia and Slovenia. In each country, 600 questionnaires were sent to small, 300 to medium enterprises, and 300 to large companies. The number of questionnaires mailed to small companies was double the number of medium and large companies because small companies constitute the highest proportion of companies and based on our personal experience, they are less likely to respond. In total, there were 223 responses (21 from Denmark, 112 from Slovakia, and 90 from Slovenia) out of 3600 mailings, i.e. the response rate was 6.2%.

Respondents were to answer what the actual total cost of ERP system implementation was – whether it was less than planned, as planned, or more than planned. In case that the total implementation cost did not match the planned one, they were asked how many percent less or more they actually spent on implementation. There were 120 responses, which compared actual and planned implementation costs (and provided all required information on independent variables) and 114 provided enough input to calculate the actual percentage. Independent variables were implementation approach, country, company size, representation of the IT department on the board level, information strategy, and number of implemented modules. The questionnaires were sent to companies in Denmark, Slovakia and Slovenia, so therefore one of the independent variables is country. The implementation approaches were big bang, phased-in, and a pilot project implementation. In the analysis, we have analyzed small, medium and large companies. The definition, which we used, stated that companies from 10 to 49 employees are considered to be small, companies from 50 to 249 employees are considered to be medium-sized enterprises, and companies having 250+ employees are considered to be large companies. This definition is consistent with how the European Commission (European Commission, 2003) defines SMEs. Regarding the independent variable information strategy, this should be understood as that the organization has a formal information strategy. Representation of the IT department on the board level means that there is a CIO or alike director for IT on the board level. Therefore, it will be described as CIO in Figure 1-5. Regarding modules, we asked if they implemented modules for (1) finance and controlling, (2) human resources, (3) manufacturing and logistics, (4) sales and distribution. So the figure used for the analysis is not the overall number of modules but the number of the abovementioned groups covered by implemented modules.
In this paper, three relationships were investigated with the aim of finding if and how the factors: CIO, country, company size, information strategy, and number of modules influence costs of ERP implementation in the form of actual costs compared to budgeted costs. This is done by first exploring whether the factors influence selection of a specific ERP implementation approach. This is then followed by exploring whether a specific implementation approach influences companies’ costs when they implement ERPs. The final relationship investigated is then if the factors have a direct influence on companies’ costs when they implement ERPs.

In the first relationship (if factors influence selection of implementation approach) the independent variables: country, company size, CIO, information strategy, and number of modules are investigated with the aim of finding if there exists any relationships between them and the dependent variable implementation approach. In this context, implementation approach is described as being a strategy for implementation that can be done in three different ways: big bang, slow phased, and pilot project. The result of this is shown in figure 2.

The second relationship investigated is then whether a specific ERP implementation approach influences costs of the ERP implementation. This is done from the independent variable implementation approach described as either slow phased, big bang or pilot project implementation. The result of this exploration is shown in Figure 3.

The final relationship is then an investigation whether the independent variables CIO, country, company size, information strategy, and number of modules influence costs of ERP implementation in the form of actual costs compared to budgeted costs. The result of this is presented in Figure 4, which shows the results related to staying on budget, and in Figure 5, which shows the results related to percentage spent. What they both present is how the suggested factors influence actual ERP system implementation cost vis-à-vis planned costs. The first relationship looks into how many companies did not exceed their planned budget. There were only three Slovenian companies, which spent less than planned; they were merged with companies, which spend exactly the amount they planned, since both can be classified as staying on budget. The second relationship investigated focuses on the percentage spent compared to the planned amount.

Regarding the methodology, logistic regression was used for the analysis of the first relationship, analysis of variance (ANOVA) and chi-square test for the second one, and ANOVA and logistic regression for the third one. Multivariate approach was used in both ANOVA and logistic regression.
Additionally, binomial test was used to test if there is a significant difference between the percentage of companies that stayed on budget and 50%; Tukey-Kramer multiple-comparison test was used to identify differences between individual instances of independent variables; t-test and Wilcoxon signed-rank test were used to test if there is a significant difference between the average ratio of actual ERP system implementation cost vis-à-vis planned costs and 100% (i.e., companies spending exactly according to plan). Results of the statistical tests are commented on confidence level $\alpha = 0.05$.

4 RESULTS

In the study, the overall results of the question about selection of implementation approach are the following: 28.5% said they used a slow phased implementation, 20.8% used a pilot project for implementation, while 50.7% used the big bang implementation. The result in our study shows a significant (p-value < 0.001) difference between the percentages and results presented by Palanisamy (2007). The result of our study is as such interesting to compare with statements about the IT productivity paradox (Brynjolfsson, 1993; Brynjolfsson, 2003; Carr, 2004; Hitt and Brynjolfsson, 1996) and statements about the risk of big bang implementation (Parr et al., 2000; Scott et al., 2000). It is also interesting to think about software vendors and distributing partners when analyzing this. It can, definitely, be stated that the way ERPs are implemented, depends on the vendor’s suggestion and since they suggest and provide tools for big bang implementation, it is not that strange that 50% of implementation is done in that way. However, there could also be other factors involved and the idea was to investigate if and how country, company size, CIO, information strategy, and no. of modules influences selection of implementation approach. The results of this are summarized in Figure 2.

![Figure 2 Exploring the relationship of factors influencing selection of ERP implementation approach](image)

The first factor investigated is whether country as such has any influence on selection on ERP implementation approach. The three country investigated are Denmark, Slovakia, and Slovenia. These three countries show definitely some cultural differences and therefore we were interested in whether they differ in what ERP implementation approach that are selected. The result of how country influences selection of implementation approach is shown in Table 1.

Table 1 does not show any significant difference between countries regarding selected implementation approach. There is a small difference when it comes to the percentage of big bang implementation between Denmark and Slovenia, where Denmark shows the highest number of big bang implementation. Otherwise the results are in line with the overall results on ERP implementation approach. It would be interesting to further investigate whether the difference in percentage of big bang implementation is a cultural difference between the three countries. An initial hypothesis on this could be that Slovak and Slovenian companies are more used to work with long time planning and therefore in higher degree go for the slow phased and pilot project implementation in relation to
Danish organizations. It may be assumed that subsidiaries of multinational companies, regardless of location, will have to use the same ERP system and the same implementation approach as selected by headquarters, thus independent from the county, in which a subsidiary is located. On the other hand, the number of multinational companies should be small enough to influence the investigation of relationship between the country and the selected implementation approaches.

<table>
<thead>
<tr>
<th>Influencing factor</th>
<th>Implementation approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Slow phased</td>
</tr>
<tr>
<td>Denmark</td>
<td>19%</td>
</tr>
<tr>
<td>Slovakia</td>
<td>30%</td>
</tr>
<tr>
<td>Slovenia</td>
<td>30%</td>
</tr>
</tbody>
</table>

*Table 1 Relation between country and selected implementation approach*

In the questionnaire, there was no specific question about size of organizations, this information was instead collected from other sources and linked to each response. The organizations were divided into the following size groups: large, midsized, and small using the measure of number of employees. This means that large is when the organization has 250+ employees, midsized 50-249 employees, and small when the organization has less than 50 employees. From the perspective of selection of ERP implementation approach, the size of the organization is of interest. Our basic thoughts about this are that if it is a small organization then it would go for the big bang implementation, and the reverse would then account for large organization. The rationality for this statement would be that a small organization does not have so many users so it would therefore be easier to do a big bang implementation. In the large organization it would be more risky to do a big bang implementation and therefore would it be possible to suggest that slow phased ERP implementation is more often used.

The result from the questionnaire related to organizational size is shown in Table 2.

<table>
<thead>
<tr>
<th>Influencing factor</th>
<th>Implementation approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizational size</td>
<td>Slow phased</td>
</tr>
<tr>
<td>Large</td>
<td>27%</td>
</tr>
<tr>
<td>Midsized</td>
<td>25%</td>
</tr>
<tr>
<td>Small</td>
<td>35%</td>
</tr>
</tbody>
</table>

*Table 2 Relation between organizational size and selected implementation approach*

There is no difference between large, mid-sized, and small organizations when it comes to which approach is most commonly used, and as shown in Table 2 the most commonly used approach is big bang. However, percentage of big bang as used implementation approach decreases with the increasing size of organization. This means that percentage of slow phased implementation is higher in small organizations than it is in large organizations. In one way it could be said that this is strange since it should be easier to do a big bang implementation in a small organization and less risky than in a large organization. One potential explanation to the result could be that implementation approach depends on implemented ERP system, and this could mean that in for instance large organization SAP is more commonly implemented and it could be that the implementation approach is influenced by what ERP that is implemented. Another possible explanation could be that large organization have resources available to do some kind of test implementation which they after having done decide on roll-out in a big bang implementation. This would then mean that they do a big bang implementation after doing a sort of parallel test implementation. For small organizations it can be suggested that they do not have the resources to do that and since the potential impact of a failure of a big bang implementation is smaller they maybe more often directly goes for big bang ERP implementation.
According to Bernroider et al., (2005), who used the same three types of implementation strategy, in small and medium enterprises (which they merged together), the most often used implementation approach was big bang, the second implementation approach was slow phased, the least used was pilot study. In large companies, the most often used implementation approach was big bang, the second implementation approach was pilot study, and the least used was slow phased. Our results are in line with Bernroider’s results but it differs to some extent when it comes to percentage of pilot project and slow phased implementation in large organizations. Unfortunately, although there are several studies conducted in the U.S., such as Madapusi et al. (2005), although investigating only two general implementation strategies – big bang and phased-in, but they cannot be compared to these outcomes, since the definition of a company size differs significantly between the U.S. and European Union.

We aimed to investigate if presence of IT professionals in the board of companies influences the selection of ERP implementation approach. In order to do that we asked if the IS/IT division were represented at board level in the organization. In the paper, we describe representation in the board as having a CIO. The reason for asking about representation at the board level was that this may for instance influence selection of ERP implementation approach. It could be suggested that if there is representation of IT/IS at board level then the risk of implementing ERP as big bang would be considered in more depth and from that it would be possible to state that a higher level of pilot project and slow phased implementation should be the case. The result of this question can be seen in Table 3.

<table>
<thead>
<tr>
<th>Influencing factor</th>
<th>Implementation approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIO</td>
<td>Slow phased</td>
</tr>
<tr>
<td>With CIO</td>
<td>23%</td>
</tr>
<tr>
<td>Without CIO</td>
<td>33%</td>
</tr>
</tbody>
</table>

Table 3 Relation between presence of CIO and selected implementation approach

The results on representation of the IT department at the board level do not show any significant results regarding relation to selection of implementation approach. There is a small difference between the slow phased and pilot project if the organization have a CIO or not, but if summing up slow phased and pilot project and comparing the sum with the big bang implementation approach no difference is seen. The result is to some extent a surprise since it does not show any difference whether the IT department is represented on board level or not on selected implementation approach. Further research may include a deeper investigation of CIOs – their educational background (field of study), risk adversity, leadership style, possibly also related factors like organizational culture and structure and size of IT department. The latter and cooperation between business and IT staff may be related also to information strategy.

In the questionnaire, it was asked whether the organization has an information strategy or not. The answer on this question is rather complex to interpret since having strategy or not could be seen from the perspective of whether the strategy is formalized or not. However, our intention with this question was to distinguish between if the organization has a written formal strategy or not. From the extent of having a formalized strategy or not the idea is then to investigate if it influences selection of a specific ERP implementation approach. The result of this is shown in Table 4.

<table>
<thead>
<tr>
<th>Influencing factor</th>
<th>Implementation approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information strategy</td>
<td>Slow phased</td>
</tr>
<tr>
<td>Yes</td>
<td>23%</td>
</tr>
<tr>
<td>No</td>
<td>35%</td>
</tr>
</tbody>
</table>

Table 4 Relation between information strategy and selected implementation approach
The results from the analysis show that there are no big surprises in whether the organizations have a formal information strategy or not in relation to selected implementation approach. There is a higher extent of slow phased implementation in organizations without formal information strategy, which could indicate that these organizations does not have that clear perspective on whether they should fully implement the specific ERP. However, it could also be that they have a clearer perspective of what they aim at and therefore takes longer time for the actual ERP implementation and focus more on adjustment of the specific ERP as well as adjustment of specific business processes.

The numbers of modules used for the analysis are actually numbers of the groups of processes (finance and controlling, human resources, manufacturing and logistics, sales and distribution) covered by implemented modules, not the overall number of modules.

<table>
<thead>
<tr>
<th>Influencing factor</th>
<th>No. of modules</th>
<th>Slow phased</th>
<th>Pilot project</th>
<th>Big bang</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29%</td>
<td>18%</td>
<td>53%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>33%</td>
<td>40%</td>
<td>27%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>27%</td>
<td>8%</td>
<td>65%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>25%</td>
<td>23%</td>
<td>52%</td>
<td></td>
</tr>
</tbody>
</table>

*Table 5 Relation between no. of modules and selected implementation approach*

There is a significant difference between companies with 2 and 3 modules. The difference is namely in pilot project and big bang implementations. Both groups use slow phased approach in about the same percentage of cases but companies with two modules much more often go for pilot project implementation and companies with three modules choose much more often big bang implementation.

The model for investigation the relationship between implementation approach and ERP implementation costs is shown in Figure 3.

*Figure 3 Relationship between implementation approach and ERP implementation costs*

In order to test relationship between the actual percentages spent vis-à-vis the planned implementation costs and selected implementation method, ANOVA was used. The averages on percentage spent for each implementation approach are shown in Table 6.

<table>
<thead>
<tr>
<th>Influencing factor</th>
<th>Slow phased</th>
<th>Pilot project</th>
<th>Big bang</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage spent</td>
<td>104.7 %</td>
<td>105.9 %</td>
<td>109.3 %</td>
</tr>
</tbody>
</table>

*Table 6 Average percentage spent in relation to implementation approach*
There is no significant relationship between the chosen implementation approach and the actual percentage spent. In order to test relationship between staying on budget and selected implementation method, chi-square test was used. The distribution of the observation is shown in Table 7.

<table>
<thead>
<tr>
<th>Influencing factor</th>
<th>Implementation approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staying on budget</td>
<td>Slow phased</td>
</tr>
<tr>
<td>Yes</td>
<td>76%</td>
</tr>
<tr>
<td>No</td>
<td>24%</td>
</tr>
</tbody>
</table>

Table 7 Relation between staying on budget or not in relation to implementation approach

There is no significant relationship between the chosen implementation approach and staying on budget. Since there is no clear relationship between the chosen implementation approach and implementation costs, we decided to add additional variables into the model.

There is a significant relationship between staying on budget and having a formal information strategy. Companies with formal information strategy seem to be more likely to stay on budget (74.2 %) than companies without information strategy (59.3 %). Overall, 67.5 % of companies stayed on budget; this percentage is significantly different from 50 % (p-value < 0.001), i.e. more than one half on companies actually manages to stay on budget.

Figure 4 Findings from analyzing the first relationship

When analyzing the second relationship, data were transformed into percentages and these percentages were then analyzed. Findings about the second relationship are summarized in Figure 5.

Figure 5 Findings from analyzing the second relationship
ANOVA identified a significant relationship between the percentages of actual spending compared to planned one and information strategy. Companies with a formal information strategy were less over budget (104.2 %) than companies without one (111.3 %).

Overall average was 107.3 %. There is a significant difference between the overall average of 107.3 % and 100% (no disparity between planned and actual costs). P-value is smaller than 0.001 regardless whether t-test for difference between mean and value, or Wilcoxon signed-rank test for difference in medians is used.

Based on the results, it can be summarized that companies with formal information strategy are likely to spend about 7 percentage points less than companies without information strategy. It suggests that ERP system vendors need to be sensitive to companies without information strategy, since these have either wrong expectations of costs or lack technical skills beneficial for ERP system implementation. However, there are also other explanations that are worth mentioning and these are the following. Firstly, it could be that companies with formal strategy are better on making a budget and take more serious in the task of doing that. Secondly, it could also be that they are better in constructing a clearer contract with the implementing partner. Thirdly, it could also be that they have a better control over overall costs and thereby are better in calculating the implementation budget. Fourthly and finally, most likely have organizations with a formal information strategy a clearer view over what they want and thereby do not so many “surprises” show up during the implementation.

5 KNOWN LIMITATIONS AND FUTURE RESEARCH

There are two known limitations of this paper, which are actually inherent for most of questionnaire surveys – response rate and reliability of data. Usually, there is an average response rate of 10 % expected in questionnaire surveys. But a response rate of 80 % and less (that is a case of almost all questionnaire surveys) can already lead to biased results. We tried to overcome the problem by sending out 3600 questionnaires and hoped that the autoselection would not depend on the questions asked. In our opinion, we achieved it, since the percentage of companies being over budget (i.e. ones, which would be more likely to complain about their bad experience) is only 32.5 %, i.e. less than 45 % (which included also projects going over time) mentioned in Cunningham (Cunningham, 1999), and surveyed companies were only 7.3 % over budget, i.e. much less than 114 % for small, 82 % for medium, and 78 % for large companies mentioned in the Standish Group (Standish Group, 1995) report. Regarding the reliability, it is not possible to check it without being allowed to look into accounts and to talk to people involved in the implementation, which would provide insight necessary to understand the accounting data.

Regarding the implementation approach, one could also consider additional factors, such as size of the system, its complexity, organizational hierarchy, and extent of the coverage.

The future research should look into what caused additional costs. For example, customization of ERP is a crucial, lengthy, costly aspect of the implementation of ERP systems (Gefen, 2002). Studies have shown that many organizations exceed their budgets due to the need for more customization than they originally planned (Markus, 2000; Markus, Cornelis and Paul, 2000a; Swan, 1999). Besides customization, companies often run into higher than expected costs for temporary and overtime labor, re-skilling, and training during the implementation process (Markus, 2000; Markus et al., 2000a; Sumner, 2000).

Last but not least, it might be useful to investigate whether additional costs arose because of the misalignment (the gap between the standard version of the ERP system and the organization) or was it spent in order to increase benefits. Investigation of both total costs of ownership and total benefits of ownership might provide a different angle for looking at expenditures.
6 CONCLUSIONS

To sum up, although not all companies manage to stay on budget when it comes to ERP system implementation, the situation in investigated European companies is not too critical. It can be evaluated from two points of view. First, about two thirds of companies still manage to stay on budget. Secondly, companies exceeded their budgets only by 7.3% on average. A contributing factor for Danish, Slovak and Slovenian, i.e. European, companies staying more-or-less on budget is the prevalent fixed price policy for ERP implementation projects in Europe. So, the findings might be generalized in European context but definitely not for the U.S., where effort-based pricing policy is prevalent. A formal information strategy implies more comprehensive planning, so there should be also smaller discrepancies between the plan and the reality. It was found out that the chosen implementation approach does not influence the ability of the company to stay on budget with implementation costs. The research also pointed out that selection of the implementation approach depends on number of modules, which are implemented.

References

COMPETITIVE ADVANTAGES OF ELECTRONIC MARKETPLACES IN THE RETAIL AUTOMOTIVE AND MAINTENANCE, REPAIR AND ORDER (MRO) INDUSTRIES.

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0608.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business value of IT, E-commerce (B2B / B2C / B2G / G2C), IT Strategy, Marketplaces</td>
</tr>
</tbody>
</table>
COMPETITIVE ADVANTAGES OF ELECTRONIC MARKETPLACES IN THE RETAIL AUTOMOTIVE AND MAINTENANCE, REPAIR AND ORDER (MRO) INDUSTRIES.

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0608</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business value of IT, E-commerce (B2B / B2C / B2G / G2C), IT Strategy, Marketplaces</td>
</tr>
</tbody>
</table>
Abstract

We question the strategic potential of EMPs to bring a competitive advantage to their users: their nature and the main differentiating factors that progressively appear over time.

Based on three case studies in the retail, automotive and MRO (Maintenance, Repair and Order) industries, conducted from 2002-2007, we more thoroughly describe where buyers and suppliers perceive strategic advantages in the eValues brought by EMPs. We classify these strategic opportunities according to Porter (1980) and Wiseman’s (1985) typologies: differentiation, cost, innovation, alliance/power gains and growth. Finally, we explore how these different factors have evolved over time in users’ perceptions.

The paper brings an in-depth and longitudinal empirical study of EMPs specific competitive advantages in each industry.

Keywords: Electronic marketplace (EMP), competitive advantage, strategic information system, inter-organizational information system (IOIS), information systems (IS), purchasing, supply, eValues, electronic catalogues, electronic auctions, sourcing, integration, standardization.
1 INTRODUCTION

During the last ten years, electronic marketplaces (EMPs) have been progressively adopted in industrial and service activities. Emarketservices study (Zällh, 2005) points out 52 significant EMPs in 17 industries based on two criteria: reputation and significant global traffic. If EMPs survival since 1999 proves that these intermediaries bring different types of eValues to their users, it is still not clear whether EMPs will only be transactional inter-organizational systems (IOIS) or whether they will hold a strategic potential both for their users: buyer and supplier organizations.

The literature demonstrates the different types of eValues brought by EMPs but does not say if these eValues can bring a competitive advantage to their users. More precisely, we will focus on the following questions: do EMPs drive differentiator factors to suppliers and buyers? What types of eValues could be considered as a competitive advantage? How do these strategic factors evolve over time from 2000 to 2007? Are there differences in these advantages according to the type of industry? What are the long-term competitive advantages that EMPs will provide to buyers and suppliers in their industry?

In the first section, we bring back EMP characteristics and their challenges in purchasing and supply chains. Then, we show the different types of eValues in EMPs. Thirdly, we summarise what the literature tells about the emergence of a competitive advantage in EMPs. In the second section, we present our research methodology. In the third section, we present our results and conclude with a discussion.

2 LITERATURE REVIEW:

2.1 The EMPs: definition, characteristics and challenges for purchasing and supply chains

In an organizational perspective, EMPs can be seen as intermediaries between buyers and suppliers devoted to answering purchasing and supply chain needs. They can also be defined as inter-organizational information systems (IOIS) that interact to create, store, transform, and communicate data between buyers and suppliers.

Behind the term EMP, there are actually a number of different business models, that is to say strategic positioning and models of revenue. Different criteria to classify EMP business models appear in the literature. Kaplan and Sawhney (2000) distinguish EMPs that focus on exchanging MRO (Maintenance, Repair and Operations) products and services like Hubwoo with those that focus on more strategic goods directly linked with manufacturing processes (like WWRE and Supply On). In the same connection, Barratt and Rosdahl (2002) highlight “vertical EMPs” that offer their services to a single industry like WWRE and Supply On, with “horizontal EMPs” that offer their services to all types of industries. Soh and Markus (2002b) go further by proposing three criteria to represent EMPs landscape according to the value proposition of the EMP, the product/market positioning and the value of activities. Finally, the e-business literature aimed at EMPs distinguishes two main types of eValue in EMPs: transactional EMPs aimed at executing exchanges and collaborative EMPs aimed at

---

1 In this article, we will mention buyers and suppliers to talk about organisations being in a position of buyer or seller/supplier in the EMP. We are not directly talking about individuals working as purchaser or seller.
encouraging collaboration between participants (Mahadevan, 2003). In this paper, we will focus on the differences between EMPs positioned in different industries and types of activity.

Whatever its business model, the EMP’s essential purpose is to provide an answer to the purchasing and supply challenges summarized by Carter et al. (2000). They highlight the challenge of electronic commerce to improve communication processes, performance metrics to assess suppliers, sourcing tools to help supplier identification and selection globally, the choice to internalize or externalize MRO purchasing, and the general evolution of exchanges around large-scale buyer and supplier companies. They conclude by describing the increased use of IT in purchasing and supply to reach these challenges.

2.2 The nature of electronic eValues in EMPs

Different types of eValues are brought by EMPs with different ways of classifying these eValues. The literature generally presents two main origins of eValues. The main benefits gained by technology are mentioned by Malone, Benjamin, and Yates (1987). They assert that electronic interconnections will bring three main benefits to markets: electronic communication will accelerate data transportation generating cost reductions; electronic matching will improve sourcing; and electronic integration will facilitate process coupling. In the same vein, Bakos (1997) concludes that IT diffusion in electronic markets will bring price transparency to buyer organizations and increase their knowledge of supplier organizations cost structure. IT will also make it easy to compare standard offers between different suppliers. Finally, IT will bring cost reductions linked with negotiations and transportation.

In an empirical perspective, Kambil and Van Heck (1998) attribute different IT benefits to electronic auctions; they facilitate supplier identification and price negotiation and they improve the coordination of logistic flows. Moreover, IT increases the richness of data exchanged and brings processes that reinforce regulation control. Hence, IT reduces opportunism risks and provides the capacity to track flows as arguments to avoid conflicts.

On the other hand, eValue is described though the role of electronic intermediaries in an industry characterized by its structure and specific needs. Kaplan and Sawhney (2000) highlight aggregation (the capacity to bring many buyers and sellers under the same roof) and matching (an optimized encounter between offer and demand to minimise negotiated prices) as the two main EMPs’ eValue. Aggregation is particularly interesting in the case of a fragmented industry whereas matching serves to reintroduce competition in oligopolistic industries to reduce negotiated prices. The effect of EMPs on industrial structures are not known yet: whereas Malone et al. (1987) predict a reduction of vertical integration with a “move to the market” others such as Clemons and Row (1992) state that electronic exchanges will bring a “move to the middle” with an increase of middle-sized companies connected through electronic networks. Amit and Zott (2001) focus on asset complementarities (Han et al. 2005), innovation synergies, the nature of the participants and the exchanged mechanisms (structure), and finally the governance defined as the interaction rules. Actually, EMPs can be considered as strategic networks (Gulati, Nohria, Zaheer, 2000), that is to say « stable interorganizational ties which are strategically important to participating firms. They may take the form of strategic alliances, joint ventures, long-term buyer-supplier partnerships, and other ties (Gulati et al. 2000: 203). In addition, these networks are characterized by the opportunity to share risks and generate economies of scale (Katz, Shapiro 1995; Shapiro, Varian, 1999), to share knowledge (Dyer, Singh, 1998), to facilitate access to the market (Kogut, 1998), to reduce information asymmetries and to improve coordination. Finally, these networks highlight the key role played by suppliers and customers to create value (Afuah, Tucci, 2001). Rayport et Sviokla (1996) also develop the concept of a virtual value chain where eValue is created by the way of combining informational with physical value chains. Kambil and Van Heck (1998) illustrate this capacity in the Dutch flower industry.
It appears that there are many different nature of eValue in the e-business context but no consensus on its source. eValue is partly made by technology, and partly made by the services provided by EMPs through their intermediation. If the process through which eValues is generated in EMPs is difficult to catch, it is yet possible to catch sight of the types of eValues that will be diffused in all organizations and the ones that will bring a competitive advantage to their users.

2.3 Dynamic evolution of eValues in EMPs: from a transactional value to the emergence of a competitive advantage

One of the key questions in evaluating EMPs impact on markets is to know whether these IOIS will bring a competitive advantage to their users. Will EMPs be a trend or will they shape long-term exchanges in the digital economy?

Information systems (IS) for competitive advantage are defined as IS that drive or formulate the organization’s competitive strategy in order to provide it with (or maintain) a competitive advantage (Wiseman, 1985). In this perspective, IT may be used to help the organization produce at lower cost, to differentiate itself from its competitors, or to identify and concentrate on a particular market segment (Porter, 1980; Porter, Millar, 1985; Clemons, Row, 1991). Porter and Millar (1985) also incorporated the concept of value with the previous meanings of strategic impact. An IS would have strategic impact if it had the potential to add value to a product or a service in at least one stage of the value chain. Moreover, considering all the supply chain of IOIS, IS will provide a competitive advantage if it modifies the structure of the industry, improves the position of the firm, or creates new business opportunities. In the same path, Rackoff and al. (1985) develop the theory of strategic thrusts to identify strategic IS opportunities. Strategic thrusts are major competitive moves (offensive or defensive) made by firms to use IT to create a competitive advantage. The authors identify five strategic thrusts: differentiation, cost, innovation, growth and alliances. In this paper, we will identify the latest with the capacity for suppliers or buyers to gain power in the chain (Cox, 2003).

In this paper, we will cross this framework to the resource-based-view (RBV) theory. This theoretical framework seems to be particularly interesting in analyzing the strategic potential of EMPs (Ordanini, 2005; Wernerfelt, 1984) considers that the growth of a company depends on its capacity to identify and exploit resources that may give high profits. The identification of these resources partly relies on the capacity to exploit markets’ inefficiencies. Barney (1991) shows that a firm’s competitive position is based on resources that it is able to control. In order to sustain a competitive advantage, the resource should be valuable, rare and difficult to imitate or substitute.

Jelassi and Enders (2005) question the capacity of EMPs to generate a competitive advantage for buyers. Some EMPs competitive advantage factors have yet been highlighted in the literature. Soh and Markus (2002a) question the capacity for EMPs to drive collaboration between buyers and suppliers as a potential competitive advantage. Ordanini (2005) identifies nine effects of participation in a digital exchange: process cost reduction, time saving, quality of process, purchasing cost reduction, increased number of suppliers, increased number of customers, sales growth, information and knowledge, partnership and cooperation. Further on, Ordanini (2006) summarizes three main factors that bring an advantage to EMP buyers: standardization, business process integration and IT negotiation tools to aggregate suppliers. Soh et al. (2006) show that price transparency should be the main advantage provided to buyers by EMPs using eAuctions. However, some EMPs do not provide this transparency to buyers in order to attract sufficient suppliers to be able to obtain a critical mass in exchanges. Then, they offer buyers “compensatory benefits” such as information content and IT change management. Before that, in the next section, we will present our methodology.
3 METHODOLOGY:

This investigation is based on three in-depth multiple case studies on EMPs made from 1999-2007: WWRE (World Wide Retail Exchange) in the retail industry, Supply On in the automotive industry and CC-Hubwoo in all types of industries to exchange indirect goods and services. These case studies included interviews conducted with the managers of each EMP, but also with their users, buyers and suppliers. Case studies (Yin, 1994) are applicable when control over events is not needed and when there is a focus on contemporary events and multiple-level analysis. Case studies permit the analysis of many variables.

Previous studies in EMPs provide us with an understanding of each EMP business model. We use these interviews to sum up the initial value proposition of each EMP in order to better analyze the strategic potential of EMPs perceived (Davis, 1989) by their users. We suppose that users were the best positioned stakeholders to have an objective view of the strategic potential of EMPs - hence avoiding the biases of the marketing discourse of EMP managers who have to justify their value to their buyer and supplier customers.

In this research, we focus on users and analyze 28 semi-structured interviews of buyer and supplier organizations: 9 in WWRE, 10 in Hubwoo and 9 in Supply On. In these organizations, we interviewed different profiles of managers: IS managers, Purchasing managers in buyer organizations, Sales managers in supply organizations, Supply Chain managers, and CEOs. We also use secondary data documents such as cases studies and interviews shown on EMPs web sites, roll out documents, users return on investments analysis.

We process data in two steps. Firstly, we use N’Vivo in order to codify the competitive advantage factors enhanced by buyers and suppliers according to Porter’s (1980) typology: differentiation, cost, innovation, alliance/power gains and growth in bold in Table 1). To do so, we first transcribe interviews and codify them in tabulars. We link “differentiation” factors when they were presented by interviewees in comparison with their competitors with key words or meanings such as competitive advantage, success key factor, and unique resource. Besides, we complete the “Innovation” factors codified according to a previous typology (Authors, 2006).

Secondly, we explore the evolution of the frequency of these different items over time. To do so, we note the number of items that appear during the different key periods of the evolution of EMPs: their start and survival period from 1999-2002, the consolidation of EMP business models from 2003-2005 that correspond to the consolidation of the advantages perceived by users, and finally, the stabilization of their activity since 2005. We also enrich this frequency approach to all paragraphs that describe the dynamic evolution of the perceived value brought by EMPs. This longitudinal analysis helps us to identify the orientation of the strategic potential of each EMP, as well as differences between industries. Thirdly, we use these case studies to highlight the strategic potential of EMPs that progressively appear in user discourses.

4 CASE STUDIES:

4.1 WWRE in the retail industry:
WWRE was founded in 2000 with the main middle-sized retailers worldwide (Tesco, Ahold, Kmart, Casino, Auchan etc…) with the exception of Wal-Mart and Carrefour. The former decided to develop its own IT tools in a proprietary mode, whereas the latter decided to join GNX – Global Net Exchange), which became a competitor for WWRE. Since 2000, with the overloading of customer mass consumption in Europe and US and the increased competition between suppliers in new markets such as China, retailers encounter difficulties in maintaining their margins. The industry is then looking for other growth opportunities such as mergers and management and organizational innovations. WWRE is then created, with I2 and Ariba as IT partners, to boost B2B exchanges.

Historically, the retail industry is one of the most advanced in electronic linkage. As an example, EDI covers 90% of exchanges between retailers and their main suppliers and is recognized to have provided major improvements in delays, reduction of data mistakes (Holland, 2003). However, the lack of IT integration between suppliers and retailers leads to numerous mistakes in the supply process (notably orders). The industry tends towards synchronizing the informational and physical supply chains for the logistics delivery (Rayport, Sviokla, 1996). The purchasing process has been managed with traditional face-to-face negotiations in central purchasing departments. WWRE initially offered a large set of IT tools and services covering the whole exchange process. Finally, users only adopted e-RFX, e-Auctions and e-Catalogues.

As summarized in Table 1, buyers see three main differentiation factors in WWRE. First, EMPs accelerate the use of electronic auctions with price gains. Second, they recognize the promising potential of global data synchronization in future transaction gains. Third, they already include knowledge management gained by shared returns on experience on the implementation of IT tools between members. The latest is shared with suppliers’ perspective. Suppliers also describe WWRE as a customer-oriented IS that incites them to orientate their sell-side information systems according to customer needs. The EMP initiates privileged links with retailers that may turn out to be lock in benefits in the future.

Cost gains are shared between buyers and suppliers and seen as directly linked with their participation in EMPs. The intermediary improves data standardization and quality as well as the standardization of the supplier communication process. Finally, the EMP favors resource mutualisation between members and hence reduces investments to manage electronic exchanges.

Another strategic potential is the innovative dimension of the EMP. For retailers, innovation lies in the capacity to easily compare supplier offers thanks to process standardization or even to access international supplier data bases owing to the EMP partnership. Besides, by facilitating retailers with tools such as eAuctions or more collaborative ones such as CPFR (Collaborative Planning, Forecasting and Replenishment), EMPs bring fuzzy relational modes with suppliers from collaboration to hard competition. Finally, the reporting tools of EMPs provide retailers with decision-making tools helping them to know the products that bring higher sales and margins in stores. In addition, suppliers highlight the improvement of data quality through synchronization. This leads suppliers to better control the information sent to retailers. Suppliers also mention the fuzzy relational modes as a risk factor in exchanges.

Alliance and power gains to retailers characterize the main strategic potential of WWRE. First, retailers build alliances with other retailers to have a lobbying activity, for instance on the choice of standards at an international scale. The EMP also provides middle-sized retailers with an international dimension, as it is the case for Casino or Auchan. Finally, the fuzzy supplier relationship management is seen as a weapon (Wiseman, 1985) to empower retailers’ position in the chain. In this context, some suppliers also make alliances with other suppliers to serve their own interests.

EMP participation is recognized to bring growth both for retailers and suppliers. Retailers’ main gains lie in electronic auctions in a short-term perspective and data synchronization in a long-term
perspective. For suppliers, the EMP is not mentioned as a factor of growth in itself, but since 2005 all interviewees are convinced of its strategic potential.

How do these strategic factors evolve over time?

The initial interviews in 2002 showed that the buyers and suppliers who were willing to participate in the EMP initially did not have a clear vision of the strategic impact of EMPs. The initial approach was to consider the different tools proposed by the intermediary without being able to foresee the strategic potential such as the implications for supplier relationships or alliances with other retailers. Retailers see EMPs as a way to adopt IT tools for purchasing and supply chains: “Initially, the objective of WWRE was to adopt common tools between retailers, without any more vision” (Retailer). Hence, the main gains expected were transactional: cost reduction, process reengineering (Davenport, 1993), and technological communication with suppliers.

Progressively, between 2003 and 2004 the main items are linked with the innovative potential of EMPs and the first alliance agreements emerge in discourses in 2004.

Finally, at the beginning of 2005 and after the stabilization period of their revenue model, the EMP is seen as way to obtain a competitive advantage. Growth arguments are kinked with the EMP IT tools. Some of them, such as eAuctions, are seen to have brought short-term competitive advantages between 1999 and 2003 when all retailers adopted eAuctions tools in their daily activity, initially in indirect goods (household appliances, bins) and then in food products (vitamins, vegetable bins, corns). In addition, long-term growth perspectives appear with the progressive implementation of an international global data synchronization network, connecting retailers and their suppliers, with the same standardized product and supplier data, necessary for the implementation of electronic catalogues. “We consider global data synchronization as the technological foundation of the business relationship with retailers with potential applications to electronic tenders, promotion management, and product assortment in stores” (Retailer). Other collaborative tools such as CPFR and software for shared visibility of supplier promotions in stores may be the third IT tool generation. Here again, whereas WWRE initially in 2000 highlighted collaboration as a key strategic advantage of the EMP, this advantage is still to be realized.

4.2 Supply On in the automotive industry:

Supply On was founded in 2000 in Germany by Bosch, ZF, Ina and Continental to manage B2B exchanges of direct goods and services in the automotive industry. In the 1980s, manufacturers were really manufacturing vehicles and assembling parts, only dealing with a few Tier 1 suppliers. This situation evolved in the 1990s with an increasing number of parts and technical issues in car manufacturing that led manufacturers to externalize to Tier 1, and then Tier 2 part of the engineering and assembly lines. In 2000, Tier 1 suppliers were in charge of dealing with a higher engineering and supply chain complexity as well as an increased number of communications with their own suppliers. Today, the industry is characterized by close business links between vehicle manufacturers and Tiers 1 suppliers, but also, to a lesser extent, by links between Tier 1 and Tier 2 suppliers. This is also a call to drastically improve B2B coordination and communication processes in the whole chain owing to adapted IOIS.

In 2008, Supply On links more than 65000 Tier 1 suppliers with Tier 2 (and possibly other smaller suppliers), that represent 75% of the top automotive suppliers in different areas of activity such as electronic, pneumatic, metals, turned and mill parts (Supply On internal report, “We make our global supplier management easy”, 2008).
With the help of support services such as training and a hot line, the main value proposition of Supply On is presented by the company in three areas: engineering and sourcing (Business Directory, eAuctions tools), supply Chain (EDI, Web EDI and inventory collaborative tools (VMI)) and Supplier Relationship Management Quality: (supplier performance assessment. These tools are electronically integrated since 2006.

Let us compare the evolution of the strategic potential of Supply On as perceived by buyers and suppliers over time.

From the buyer perspective, the main **differentiation factor** is **sourcing**. Supply On helps buyers to find new suppliers, specifically when they are looking for competences they do not have in their portfolio, in order to manufacture new commodities. Since 2006, Supply On has reinforced its supplier portfolio by opening an international desk in China (Shanghai) and near the historical American manufacturers in Detroit. In addition, they highlight the fact that the EMP brings a **knowledge management** competitive advantage to buyers by giving them the opportunity to share experience with industrial experts on the implementation and use of IT tools. Finally, Supply On recently offers an **integrated platform** for some tools: between RFQs and eAuctions or between Tier 2 performance monitor and the Problem Solver tool. The supplier perspective is more focused on a sales approach: the EMP brings them a competitive advantage when it effectively offers the opportunity to **get new sales contacts**, and even more increase their market share “a concentrated fair reduced in a monitor” (Tier 2 Supplier). In addition, the EMP is also seen as high value when it brings suppliers **access to the right purchasing contact** in the Tier 1 buyer company. However, only a few suppliers that have privileged relationships with the EMP, recognize these differentiator factors.

Arguments on **costs** are similar between buyers and suppliers. They agree on the fact that Supply On helps them to standardize data (notably the product norms and specifications) and the communication process. It also decreases IT investments by mutualising resources. Finally, suppliers highlight the IT integration as a main advantage in order to directly register customer tenders in their systems, and then, compress the time process execution.

According to buyers, the main **innovation factor** lies in the capacity to easily compare Tier 2 supplier offers thanks to data and process standardization. The EMP also facilitates their decision-making process by providing higher transparency on supplier performance: firstly by being able to share a common Tier 2 supplier database that involves different departments concerned in the Tier 1 company (purchasing, quality and operations) that previously did not open their frontiers from one to another; secondly by linking Tier 2 service commitments to their real performance on time, quality and delivery. Buyers also play on fuzzy supplier relationships by using at the same time eAuction tools and collaborative ones such as VMI. On the other hand, Tier 2 suppliers highlight, as the main advantage, the capacity to technically integrate customer data in their IS. However, they deplore the fuzzy relationships that buyers impose on them, especially when they do not have any unique engineering competence in the market.

Finally, buyers are quite discrete on the capacity of Supply On to create potential **alliances** with other buyers with a view to gaining more knowledge on supplier performance, with shared sourcing policies. Whereas some of them assert this will not happen, others mention the potential emergence of these alliances in the future.

**Growth** is only perceived through the IT tools like eAuctions to buyers, when they get direct returns. Accurate numbers are not communicated by interviewees. Buyers and suppliers both mention the immaterial gains in being able to share knowledge through the marketplace.

How do these strategic factors evolve over time?
The initial strategic advantage of Supply On since the beginning in 2000 rapidly appeared to be sourcing, including eAuctions, but most of all the content of a Business Directory that provides Tier 1 suppliers with unknown low-cost and efficient suppliers. However, if the initial value proposition of Supply On was oriented through engineering (map exchanges and product design), there have been a few advances in this area as Tier 1 suppliers, following the choice of vehicle manufacturers, have decided to exchange through portals with their own manufacturers or Tier 2 suppliers.

Between 2003 and 2004, the business directory and document manager increased the number and quality of its data to become recognized as a unique advantage in the market. Hence, Supply On’s business model strengthened its value proposition by offering long-term benefits to participants. Finally, IT tools aimed at improving the supply chain have been progressively adopted by users since 2004, at the request of buyers, when they sufficiently perceived the process standardization proposed by Supply On. Finally, tools aimed at improving Supplier Relationship Management Quality have revealed their full potential in the field since 2005. By offering the possibility to link supplier assessment to real time supplier performance (based on quality, time and service delivery), and by opening frontiers inside buyer departments (purchasing, quality, operations), they are considered to bring high growth potential in the future.

4.3 Hubwoo: an EMP that deals with MRO goods and services in multiple industries

Hubwoo is a French EMP founded in 1999 with SAP as one of the main shareholders. This EMP offered new opportunities to improve the demand to order delivery cycle, and then managed in paper and fax mode. Hubwoo used to put together suppliers providing electro components, electrical devices, stationery, all goods and services that generally support buyers’ activity known as MRO. The industries that manufacture MRO are characterized by small to middle-sized companies around the world. These companies deliver a wide variety of goods and services where the main value lies in linked services such as low prices, reduced time delivery, reactivity to customer demand and stock capacity.

This industry contains several intermediaries as these suppliers generally sell their products to distributors who concentrate their offer on paper or in eCatalogues, offering joint promotions to large customers in all types of industries. Hence, distributors have been direct partners to EMPs who have helped them to go through the digital economy. As an example, an international distributor and supplier in the EMP generated 30% of its total revenue in Europe through e-Commerce in 2007 and over 60% in Asia. The growing presence of e-Commerce in these exchanges implies a heavy dependency of the industry on data processing and communication systems.

Hubwoo proposed to coordinate the implementation of eCatalogues with the list of major suppliers of indirect goods and services of each buyer, by choosing the same data standards and process agreements. After a technological learning period to create and implement eCatalogues, buyers developed new relational modes with their suppliers and generally reduced their supply base: they first incite them to increase their price transparency; standardize their offer and sometimes enlarge it to answer the needs of their customers. In exchange, buyers offered them mid-to-long-term contracts (2-3 years). In 2005, Hubwoo introduced electronic billing in exchanges.

Let us compare the evolution of the strategic potential of EMPs perceived by buyers and suppliers over time.

In the buyer perspective, the main differentiation factor deals with sourcing. EMPs provide a set of IT tools and services that first allow access to international supplier data bases such as Global Sources,
and then facilitate **supplier comparisons**. A consequence of the use of EMPs and e-procurement is a “consolidation of negotiated contracts on a few suppliers able to serve global market” (Buyer). In the supplier perspective, the main differentiation factor deals with the technological edge provided by IT with the capacity to offer high quality content to customers, such as a unique and up-to-date view of the product offer.

Arguments on **costs** are similar for buyers and suppliers. They agree on the fact that EMPs contribute to reduced costs due to data standardization and quality, mutualisation of technological investments and maintenance, and finally standardization of the communication process around the purchasing and supply chain.

According to buyers, the **innovation factor** deals with the capacity of IT to gain visibility on internal expenses before the accountability stage and finally to improve supplier IS in customer expectations. Innovation also accelerates the process order execution and initiates the opportunity to pilot purchasing and supply flows. Relational modes with suppliers are also impacted with the introduction of fuzzy relationships from collaboration to hard competition. This leads to source large-sized suppliers able to serve international markets. IT then makes it possible to apply these negotiated contracts internally - owing to the structuring power of technology that restricts the choice of suppliers to order. IS act as a Big Brother tool able to control buyer behavior so that only contracts that have been previously negotiated via eCatalogues are ordered, this is the “structuring power of eCatalogues easily exclude suppliers from markets, in the day-to-day habits of operational buyers, when they are not selected in long-term contract partners” (Buyer). By doing so, the EMP contributes to aligning the whole organization on the best price negotiated: “when the group has decided to buy Dell, all the entire company effectively makes its orders to Dell” (Buyer).

On the other side, suppliers highlight the fact that participation in EMPs forces them to orientate their IS towards customers’ needs. As an example, they have to synchronize their electronic catalogues according to the EMP standard format or process communication. These modifications help them to accelerate their capacity to integrate customer data in their back ends. Finally, by being the ones able to exchange electronically, for instance by mastering electronic payments, suppliers become pioneers in IT; that leads them to lock in the marketplace. Process standardization helps them to reduce data errors, accelerate product time delivery and hence to improve their cash flows. Finally, eCatalogues are considered as positive innovations due to their capacity to be an up-to-date show window for their products, and accelerate the pace of time delivery. Despite these benefits, suppliers deplore the fuzzy relationships enhanced by suppliers with eCatalogues.

**Growth** is recognized both by buyers and suppliers. For buyers, the main benefit lies in prices, due to the capacity of the EMP to leverage negotiated volumes by bringing together the purchasing needs of different subsidiaries of the same company. Suppliers just begin to see return on investments by recovering, as incumbents, the market shares of their competitors that were excluded from eCatalogues. Most of them see a growing part of their business through EMPs and electronic channels in the future. The main benefit of these exchanges is to come with the IS integration between the EMP and their supplier IS.

How do these strategic factors evolve over time?

Here again, it is interesting to notice that the strategic advantage factors highlighted by suppliers and buyers do not correspond to the initial announced EMP’s value proposition: “we did not get return on investments where we initially expected to” (Buyer).

When the EMP started in 1999, Hubwoo basically communicated on the following three main advantages: the improvement of the communication process with emphasis on administrative tasks (organizational cost reduction in processing orders), content (data standardization and improvement of
the quality of data) and economic gains. We qualify these factors of improvements as “transactional” as they refer to reduction costs on technology or processes. Actually, buyers do not have an accurate view on cost cuts due to the EMP. However, they all recognize eValues such as differentiation factors as sources of competitive advantages, innovation, alliances/power gains or global growth.

Since 2004, the strategic eValue of Hubwoo appears to be the great visibility that eCatalogues bring in internal purchasing expenses, the control on purchaser behavior as well as power gains for buyers. For suppliers, there is still uncertainty over whether EMPs will help them to differentiate from their competitors with the lock in effect and the technology control. It is too early to see what is coming next: some suppliers think EMPs will bring them a strategic advantage whereas others do not. EMPs will have a contrary effect, by avoiding their differentiation.
<table>
<thead>
<tr>
<th>WWRE</th>
<th>SUPPLY ON</th>
<th>HUBWO</th>
<th>ALLIANCE AND POWER GAINS</th>
<th>ALLIANCE AND POWER GAINS</th>
<th>ALLIANCE AND POWER GAINS</th>
</tr>
</thead>
<tbody>
<tr>
<td>The buyer perspective</td>
<td>The buyer perspective</td>
<td>The buyer perspective</td>
<td>The buyer perspective</td>
<td>The buyer perspective</td>
<td>The buyer perspective</td>
</tr>
<tr>
<td><strong>DIFFERENTIATION</strong></td>
<td><strong>DIFFERENTIATION</strong></td>
<td><strong>DIFFERENTIATION</strong></td>
<td><strong>INNOVATION</strong></td>
<td><strong>INNOVATION</strong></td>
<td><strong>INNOVATION</strong></td>
</tr>
<tr>
<td>- Global data - Synchronisation - Sourcing - Knowledge Management</td>
<td>- Sourcing - Knowledge Management - IT integration</td>
<td>- Sourcing - Change management - Supplier size selection - Reporting tools to follow up supplier performance - Supplier selection</td>
<td>- Technology - Communication process - Supplier relational modes - Reorganization of the industry</td>
<td>- Technology - Communication process - Supplier relational modes - Product: visibility and pace of product introduction into markets</td>
<td>- Not mentioned - Not mentioned</td>
</tr>
<tr>
<td>- A customer-oriented IS - Active supplier participation in EMPs leads to supplier lock-in - Global data synchronisation - Knowledge Management</td>
<td>- Market opportunities for new customers - Access the right purchasing contact in each buyer company</td>
<td>- A customer-oriented IS - Data shared with buyer - Pioneer in IT such as punch out - Customer relationship management</td>
<td>- Alliances with other suppliers or retailers to serve suppliers’ interests (lobbying) - Fuzzy supplier relationship management, empowerment</td>
<td>- Not mentioned</td>
<td>- Not mentioned</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>COST</strong></td>
<td><strong>COST</strong></td>
<td><strong>COST</strong></td>
<td><strong>ALLIANCE AND POWER GAINS</strong></td>
<td><strong>ALLIANCE AND POWER GAINS</strong></td>
<td><strong>ALLIANCE AND POWER GAINS</strong></td>
</tr>
<tr>
<td>- Data standardization and quality - Mutualise IT tools - Standardization of supplier communication process</td>
<td>- Data standardization and quality - Standardization of supplier communication process (workflow) - Mutualise IT tools - IT integration</td>
<td>- Data standardization and quality - Mutualise IT tools</td>
<td>- Potential alliances in the future with other buyers to build a common sourcing policy, empowerment</td>
<td>- Not mentioned</td>
<td>- Not mentioned</td>
</tr>
<tr>
<td>- Data quality through synchronization - Relational modes with buyers</td>
<td>- Decision-making tools - Relational modes with buyers - Supplier performance monitoring</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>INNOVATION</strong></td>
<td><strong>INNOVATION</strong></td>
<td><strong>INNOVATION</strong></td>
<td><strong>GROWTH</strong></td>
<td><strong>GROWTH</strong></td>
<td><strong>GROWTH</strong></td>
</tr>
<tr>
<td>- Comparisons of suppliers’ offers - Supplier relational modes - Sourcing</td>
<td>- Sourcing - Supplier relational modes</td>
<td>- Technology - Communication process</td>
<td>- Price cuts owing to eAuctions - Data synchronization in a long-term perspective</td>
<td>- Not mentioned</td>
<td>- Not mentioned</td>
</tr>
<tr>
<td>- Supplier relational modes - Sourcing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ALLIANCE AND POWER GAINS</strong></td>
<td><strong>ALLIANCE AND POWER GAINS</strong></td>
<td><strong>ALLIANCE AND POWER GAINS</strong></td>
<td><strong>GROWTH</strong></td>
<td><strong>GROWTH</strong></td>
<td><strong>GROWTH</strong></td>
</tr>
<tr>
<td>- Alliances with retailers - International dimension for middle-sized retailers - Fuzzy supplier relationship management, empowerment</td>
<td>- Alliances with other suppliers or retailers to serve suppliers’ interests (lobbying)</td>
<td>- Technology - Communication process - Supplier relational modes - Product: visibility and pace of product introduction into markets</td>
<td>- Price gains due to the negotiations on higher volumes - 50% of supplier business in B2B in 2010 - Market share recovery - IT integration in supplier internal IS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5 DISCUSSION:

In this discussion, we will first compare the main competitive advantages that emerge in the different industries (See Table 1) and complement it with an overall vision of the evolution across industries of the following criteria: the dynamic evolution of IT tools, the eValue with higher frequency of items, a summary of the specific needs of each industry and finally an evaluation of the strategic potential of these eValues using Barney (1991)’s framework. Second, we will highlight the specific competitive advantages that EMPs bring in each industry. We will then conclude on the impact of EMPs and electronic exchanges on the structure of each industry.

Some factors are common to all EMPs. Others are specific to the needs of each industry.

There are three main differentiating factors in all EMPs: sourcing and knowledge management are common to all of them whereas data synchronization is specific to the retail industry. In terms of frequency of items, sourcing has the most impact in the automotive industry as this industry was one of the first to use supplier content to evaluate supplier performance through metrics. MRO industrial suppliers and buyers hold before all the capacity of the EMP to help them to provide returns on experience (implementation and adoption of best practices). The retail industry is the only one that has already started to build a global standard for product and supplier databases thereby creating a worldwide network. This was also possible because WWRE managed to link the main retailers with the major suppliers, contrary to Supply On which only connects smaller players. It is a long-term project that will connect the key players of the industry through a standard and future interoperable IOIS.

All industries benefit from cost reductions owing to data and process standardization, resource mutualisation, EMP change management and the objective to tend towards fully-integrated applications. However, this advantage appears to be short-term and is not going to differentiate one competitor from another, but to improve the whole collective marketplace.

The main innovation factor lies in new ways of managing the supplier relationship in the exchange, combining aggressive IT tools such as eAuctions with more collaborative ones such as CPFR, with the same supplier. This combination of IT tools introduces fuzzy relationships that renew the way purchasers deal with the supplier portfolio. Bakos (1997) prediction that IT will bring higher price transparency is actually limited to buyers through eAuctions. Transparency in buyer-supplier relationships has to be considered as a complex variable used to choose the suppliers allowed to enter
into the EMP (Soh et al. 2006). Hence, IT will not automatically facilitate access to the market (Kogut, 1998), especially for little-sized companies. In addition, the visibility provided by EMPs on supplier performance worldwide will enlarge the way purchasers assess their supplier from a local territory to the globe.

In all case studies and knowing that the EMPs studied are buyer-controlled, power gains clearly appear on buyer sides, in both the purchasing and supply processes. Buyers gain visibility on supplier practices, data and performance.

Finally, several factors of *growth* are mentioned: some of them can be qualified as short-term gains (price gains with eAuctions) whereas others are described as long-term gains. We further discuss the way we should consider long-term competitive advantage in EMPs with the help of RBV.

**What are the specific competitive advantages brought by EMPs in each industry, according to their structure and needs?**

The initial benefits for the *automotive industry* lie in price gains with the diffusion of eAuctions as well as cost reductions due to the improvement of the exchange process (Howard et al., 2006). These gains were short-term as all automotive players, irrespective of whether or not they belong to an EMP, now include eAuctions and process reengineering in their negotiation, whether they belong to an EMP or not. From 2000-2004, the focus was on improving the supplier and product business directory. According to interviewees, this sourcing tool is relatively rare, difficult to imitate or substitute. In addition, the supplier performance metrics appears to be relatively advanced considering that the retail industry, for example, has not yet obtained value in this area. Thus sourcing appears to be a long-term competitive advantage provided by EMPs in the automotive industry by implementing global sourcing systems (Mol, Koppius, 2002), including low cost countries, some of them offering rating functionalities based on performance metrics (Carter et al., 2000; Kleijnen J., Smits M., 2003). This leads us to conclude that EMPs bring an aggregation value (Kaplan, Sawhney, 2000) by reducing the fragmentation in the base of suppliers. That will probably bridge the gap between the oligopolistic structure of the retail industry to the one of the automotive industry and hence orientating the market to a “move to the middle” (Clemons, Row, 1992). In addition, since 2004, buyers have asserted that Supply On also provides value with IT tools in delivery and stock management processes. This drives cost reduction in the short to middle-term perspective. The next step will be to improve coordination and cost cuts derived from the IS integration of the EMP with its users. IS integration between Tier 1 and Tier 2 suppliers has always been a source of value in automotive supply chains (Clemons, Row, 1993).
In the retail industry, initial benefits were similar to those in the automotive industry: cost reductions due to process reengineering, improvement of communication, gains in prices owing to eAuctions. The sourcing value extents matching (Kaplan, Sawhney, 2000) not only on price gains but also on the capacity to deal globally, being able to exploit the specificity of local markets and cultural needs. However, these eValues have been progressively imitated by all retailers and suppliers, even if they do not belong to an EMP. The main competitive advantage brought by the EMP is global data synchronization with an improvement on product and supplier data quality and accuracy. The capacity to have high data quality in their own IS will soon allow buyers and suppliers to develop different collaborative tools such as promotions management, shared planning, stock inventory based on stores’ sales. Besides, buyers and suppliers will be able to better control the data that is sent outside their IS. This improvement in data quality and purchasing and supply IS control appears as a rare and difficult to imitate resource. Hence, suppliers and buyers who benefit from this competence will also benefit from a long-term competitive advantage. Considering the yet oligopolistic structure of this industry, we assert that electronic catalogues are not used to gain aggregation as in the automotive or MRO industries but to improve the communication and coordination all along the supply chain.

Finally, the acceleration of the restructuring of the configuration around a few large-scale retailers with electronic exchanges questions the evolution of power gains in the chain. The power gains from which buyers presently benefit as a middle-term advantage will depend on the capacity of suppliers to take advantage of IOIS to strengthen their positions. In the same connection, network alliances (Gulati et al. 2000) initiated between buyers in 2005 may bring long-term advantages if they manage to use these coalitions to their advantage without being accused of collusion.

For purchasing and supply, the final step will be to reach fully-integrated and agile IS able to coordinate and communicate in the whole supply chain, from the upper food supplier in the chain, to the different warehouses and finally, stores that directly connect to end customers. An increased visibility in the supply chain is also urgently needed to guarantee healthy products and customer protection in the case of food crisis. However, before 2000, IT linkage was limited to demand to order processes and began to reach the purchasing process with eAuctions.

The case of MRO industry is similar to the automotive and retail industries as initial gains (2000-2002) were linked to cost reduction. The EMP brought high value in accelerating change management and supplier roll out by trying to concentrate widespread suppliers in the same database format. From the start, Hubwoo has also greatly contributed to improving supplier and product data content as content was a condition to exchanges through eCatalogues. Here again, and according to interviewees, this content is relatively rare, difficult to imitate or substitute. If the supplier performance metrics is not so advanced in MRO industries as in the automotive industry, the latter will probably greatly improve the optimization of stock inventory which are notably sent to automotive customers. Hence, supplier and product content can be considered as a long-term competitive advantage. Finally, the
main long-term competitive advantage that is recognized to the EMP is the capacity to offer a **fully-integrated platform** able to make links between tenders to order delivery. **Integration** is also expected to reach supplier and buyer IS back ends. Such integration is considered to be relatively difficult to imitate or substitute and hence, considered as a long-term competitive advantage for users that will help to better manage supplier stocks. Finally and in the near future, these tools will probably be used to assess suppliers and compare their offers apple-to-apple. The EMP will then become a rating intermediary and this questions the evolution of power gain in the chain (Webster, 1993) as well as the evolution of the structure of this fragmented industry. As the MRO industry is part of the suppliers of the automotive industry, and as it is still more fragmented than the automotive industry, we anticipate that the MRO industry will follow the same path of the automotive one and will totally benefit from an aggregation value (Kaplan, Sawhney, 2000).

In conclusion, some eValues are common to all of them: cost reduction, standardization, communication process reengineering, price gains, mutualisation and power gains.

In addition, EMPs drive different types of competitive advantage according to the needs and structure of each industry: sourcing, supplier performance metrics and IT integration in the automotive industry; global data synchronisation, internal increased control of purchasing and supply IS and alliances in the retail industry; supplier and product content and a fully-integrated platform in MRO industries.

Looking at the dynamic evolution of EMP eValues, we can highlight **three main conclusions**.

First, EMPs should not only be considered as transactional IOIS aimed at driving cost reductions. They also reveal other types of long-term competitive advantages (Ordanini, 2005) that are larger than an improved collaboration in the chain. Whereas collaboration is generally presented as the main advantage of EMPs (Mahadevan, 2003), our cases show that collaboration is actually a long-term perspective that will be possible with the integration of IOIS from buyers to suppliers. However, the differentiator factors highlighted bring further questions such as the evolution of alliances and power gains in a fully-integrated supply chain between suppliers and buyers, but also in supplier-buyer relationships.

Second, following the example of EDI (Webster, 1995), EMPs have initially brought value to supply processes with the development of eCatalogues and, since the beginning of the 1990’s, EMPs brought value in the purchasing process with eAuctions and eTenders. We do consider this as an historical evolution of the value brought by IOIS from supply to purchasing processes.

Third, the eValue of EMPs are directly connected with the needs and structure of each industry. EMPs help us to imagine the evolution of global electronic markets connected through global IOIS. The example of the retail industry may foreshadow the evolution of global electronic markets around an
oligopolistic configuration that excludes little-sized companies from global and finished products of inter-organizational exchanges. Are the automotive and MRO industries going to follow the same path or are they going to create different industrial configurations and benefits from eCommerce?

6 CONCLUSION:

In this paper, we describe the dynamic evolution of eValues provided by EMPs to buyers and suppliers. We analyze these eValues according to the needs of the automotive, retail and MRO industries. Using both Wiseman (1985) and RBV theoretical frameworks, we discuss what types of eValues could be considered as short to long-term competitive advantages.

Our results first highlight that EMPs were initially introduced to gain transactional value such as cost reduction, mutualisation, standardization, process reengineering. Then, they progressively revealed their strategic potential in several directions: buyer-seller relational modes, alliances, sourcing, product and supplier content and global data synchronization. Hence, we highlight similarities and contrasts on how suppliers and buyers play with the strategic opportunities of EMPs in their exchanges. Finally, we show that this strategic potential of EMPs differs according to the strategic positioning of EMPs (on direct or indirect goods and services) and according to the specific needs of each industry. The dynamic description of three case studies helps us to foresee the evolution of the structure of these industries through the development of eCommerce.

As a continuation of this research, we could consider analyzing the risks perceived by users (Kumar, Van Dissel, 1996; White et al. 2007) and see whether they endanger the strategic factors highlighted in this study. In addition, we could more thoroughly explore how the position of buyers and suppliers in the chain and their stakeholder salience (Howard et al., 2003) may explain their capacity to take advantage of the strategic advantages highlighted.
References


Simmelian ties, organizational justice, and knowledge sharing in virtual workgroups

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0086.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Social Network Analysis, Information sharing, Perceptions, Distributed organization / teams</td>
</tr>
</tbody>
</table>
SIMMELIAN TIES, ORGANIZATIONAL JUSTICE, AND KNOWLEDGE SHARING IN VIRTUAL WORKGROUPS
Research-in-Progress

Ho, Zhi Wei, National University of Singapore, Lower Kent Ridge Road, Singapore, hozhiwei@nus.edu.sg
Chang, Klarissa Ting-Ting, National University of Singapore, Lower Kent Ridge Road, Singapore, changtt@comp.nus.edu.sg

Abstract

We argue that types of Simmelian-tied employee dyads (dyads embedded in three-person cliques) influence organizational justice perceptions, and knowledge sharing within and across organizational boundaries through virtual workgroups. We study the interaction between employees’ advice and friendship ties, shared interpersonal, interactional, procedural and distributive justice perceptions, and the types of knowledge shared from a social network perspective. We predict that Simmelian-tied advice and friendship dyads influence justice perceptions, and in turn knowledge sharing. Compared to Simmelian-tied advice dyads, we suggest that Simmelian-tied friendship dyads were hypothesized to be strongly associated with congruent distributive, interpersonal, and interpersonal justice perceptions. Congruent procedural justice perceptions were likely to be associated with both Simmelian-tied advice and friendship ties. We hypothesized that distributive, procedural, and informational justice perceptions were likely to be shared across formal organizational boundaries through strong friendship ties. We also predicted that positive congruent procedural, interpersonal and informational justice perceptions influenced expert knowledge sharing while congruent distributive justice perceptions influenced product knowledge sharing.

Keywords: Simmelian Ties, Advice Ties, Friendship Ties, Social Network Analysis, Organizational Justice; Knowledge Sharing.
1 INTRODUCTION

Research shows that organizational justice perception, or how employees perceive fairness, influence performance-related work outcomes such as job satisfaction, organizational commitment (Aryee et al., 2002), performance (Kim & Mauborgne, 1997; Colquitt et al., 2002), and work motivation (Colquitt, 2002). Justice research in recent years has shifted from the individual perspective to the congruent perceptions of groups using social structure, influence, and interaction (Lind et al., 1998; Van Den Bos & Lind, 2001; Colquitt et al., 2002; Colquitt, 2004; Roberson, 2006a; Roberson, 2006b).

Several studies point to the impact of formal (Schminke et al., 2002; Ambrose & Schminke, 2003) and informal (Chia et al., 2006; Shapiro et al., 2008; Fang & Shaw, 2008) organization structures upon justice perceptions. Field network studies further confirm that justice perceptions not only transmit via social structure but also possess the ability to influence peers (Lamertz, 2002; Umphress et al., 2003).

This research is grounded on the notion of justice perceptions as a contagion (Degoey, 2000), such that individuals seek and provide justice perceptions based on uncertainties, past experiences and social structure. Our study adds to the current literature in three ways. First, we seek to extend and clarify existing studies to address the influence of strong social ties on congruent justice perceptions between employees. Second, we examine the ability of justice perceptions to endure across organizational and geographical boundaries. Finally, we consider the impact of congruent justice perceptions on knowledge sharing, an outcome of work performance.

2 THEORY AND HYPOTHESES

The manner in which employees seek advice or friendship leads to the emergence of informal ties (as opposed to formal ties, e.g. departments). These ties carry a myriad of organizational implications, such as knowledge sharing (Krackhardt & Stern, 1988; Tortoriello & Krackhardt, 2008) and social influence (Gibbons, 2004). As organizations adopt virtual workgroups, ties possess the ability to span and influence perceptions across local and geographical boundaries (Krackhardt & Stern, 1988).

Justice perceptions are not formed in isolation; rather, they are subject to the influence of social interaction. Through existing ties and referents of injustice (Kray & Allan Lind, 2002), employees exchange and process justice information (Chia & Fang, 2005; Chia et al., 2006; Fang & Shaw, 2008) to determine whether they were subjected to unfair treatment or as a means to avoid potential injustice. However, research involving the influence of ties on justice perceptions had yielded inconsistent results (Lamertz, 2002; Umphress et al., 2003; Roberson, 2006b).

![Figure 1: Proposed Research Model](image-url)
This study aims in part to integrate and empirically examine these issues in the virtual workgroups of a knowledge-intensive organization. We also aim to study the knowledge sharing outcomes of such congruent justice perceptions, an area current literature is relatively silent about (Kim & Mauborgne, 1997; Lin, 2007). Therefore, we attempt to bridge these gaps by formulating a research model (Figure 1) and subsequently grounding the hypotheses in the literature.

2.1 Organizational Justice Perceptions

Justice perceptions of employees hold crucial implications for managers and organizations (Simons & Roberson, 2003). In our study on knowledge sharing, related outcomes include citizenship behavior (Nichoff & Moorman, 1993; Lavelle et al., 2007), job satisfaction (Wesolowski & Mossholder, 1997), rule compliance, commitment, and helping behavior (Colquitt et al., 2001). In order to distinguish the myriad of outcomes associated with justice perceptions, these perceptions have been distinguished as distributive, procedural, interpersonal, and informational (Colquitt, 2001).

Distributive justice perceptions refer to the fairness and equality of outcomes relative to one's contribution (Colquitt et al., 2001). On the other hand, procedural justice perceptions consider the process fairness or the ability to voice one's opinions during the decision process (Lind & Tyler, 1988) represent the cornerstone of procedural justice (Folger & Cropanzano, 1998). Interpersonal justice refers to whether individuals are treated with dignity and respect, while informational justice considers the completeness of processes and outcomes explanations (Colquitt et al., 2001).

2.2 Social Influence and Justice Perceptions

Sensemaking (sharing and clarifying perceptions) within groups serves as a source of social influence, leading to congruent perceptions of behavior and group norms (Salancik & Pfeffer, 1978). In the organization, justice perceptions are shared among colleagues to voice concerns and clarify ambiguities to determine fair outcomes and treatments (Degoeij, 2000; Folger & Kass, 2000; Van Den Bos & Lind, 2001; Lamertz, 2002). For example, group ratings of justice perceptions were more significant compared to the individual (Lind et al., 1998; Van Den Bos & Lind, 2001; Roberson, 2006b), especially in highly interdependent groups (Colquitt, 2004).

We draw from social network analysis to examine this social heuristic phenomenon. Social network analysis provides insights into informal structures such as group cohesiveness (Gruenfeld et al., 1996) and relationships (Krackhardt & Kilduff, 2002). Early applications of social network analysis found evidence for congruent procedural and interactional (comprising of interpersonal and informational) justice perceptions between dyadic (2-person) and triadic (3-person) relations (Lamertz, 2002; Umphress et al., 2003; Chia & Fang, 2005).

These early network studies suggest that unambiguous work outcomes of distributive justice are not subject to social influence. However, this is contrary to later laboratory studies (Roberson, 2006a; Roberson, 2006b). We attempt to clarify social influence on justice perceptions within groups by utilizing the concept of Simmelian ties and distinguishing between different types of relationships.

2.3 Simmelian Ties

A Simmelian-tied dyad (Krackhardt, 1995) is a regular dyad embedded in a triad. While a strong dyad may foster congruent perceptions, a Simmelian-tied dyads enforces such perceptions by introducing three properties: (1) enforcing group interests through outvoting (2) reducing individual bargaining power (3) encouraging cooperation through mediation by a third party.

By utilizing the Simmelian-tied dyad, the complexities of studying groups are stripped away while retaining essence of the group by considering the common third-party common. This allows us to examine the associated influence, constraints and interactions. Simmelian-tied dyads offer opportunities for sensemaking and clarifying issues with similar others (Festinger, 1954; Krackhardt & Kilduff, 2002), forging shared perceptions and agreement. Related outcomes include increased
cognition of workplace relationships (Krackhardt & Kilduff, 2002), knowledge access (Tortoriello & Krackhardt, 2008), and social support spanning organizational boundaries (Krackhardt & Stern, 1988).

2.4 Advice and Friendship Ties

Informal relationships play an important role in the organization (Krackhardt & Hansen, 1993). For example, strong workplace friendship influences shared values (Gibbons, 2004). This study concentrates on two types of social ties: advice and friendship. Advice ties concerns with which individuals approach for work-related advice (Podolny & Baron, 1997), while friendship ties consider with whom individuals share perceptions, experiences and rely for social and political support (Coleman, 1990; Krackhardt & Kilduff, 1990).

Strong advice ties lead to organizational learning (Biele et al., 2008), information sharing, and performance (Verbeke & Wuyts, 2007), while friendship ties transmit affect and emotion, transcending boundaries such as demographics (Plickert et al., 2007), fostering congruent perceptions (Gibbons, 2004). Therefore, information transmitted over advice and friendship ties have the ability to reinforce justice perceptions in varying ways (Umphress et al., 2003; Chia & Fang, 2005).

2.5 Simmelian Ties, Social Influence, and Justice Perceptions

We have addressed the social influence effects of Simmelian ties, the characteristics of advice and friendship ties, and the existing justice research. We now attempt to integrate these factors.

Distributive justice perceptions are associated with objective, work-related outcomes (e.g. pay). While laboratories studies report congruent distributive justice perception through sensemaking (Roberson, 2006a; Roberson, 2006b), earlier network studies suggest the contrary. The unambiguous nature of distributive justice perception do not seem to provide avenues for sensemaking and social influence, and the objective work outcomes led earlier studies to associate this perception with work-related ties (Lamertz, 2002; Umphress et al., 2003; Chia & Fang, 2005).

We argue that distributive justice perceptions are susceptible to social influence through friendship ties rather than advice ties, particularly through Simmelian-tied friendship ties. While distributive justice perceptions work outcomes are objective and unambiguous, individual favorability perceptions of these outcomes remain highly personal and subjective (Sweeney & Sweeney, 1992; Sweeney & McFarlin, 1993; Colquitt et al., 2001). Sensitive and subjective perceptions subjected to sensemaking and social influence (Klein et al., 2001), particularly between strong friendship ties associated with trust and concern (Verbeke & Wuyts, 2007). Thus, we propose that compared to regular friendship dyads, Simmelian-tied friendship dyads are likelier to form congruent distributive justice perceptions.

Hypothesis H1. Compared to regular friendship dyads, Simmelian-tied friendship dyads are more likely to share congruent distributive justice perception.

Procedural justice perception is associated with the ability to voice concerns and appeal against work decisions. Compared to distributive justice perception, procedural justice perception concerns subordinate-leader exchanges and supervisor evaluation (Colquitt, 2001), which are of a less personal nature. Such perceptions are likely to be exchanged among friends in the workplace and workgroup members who may not necessarily be personal friends.

Recent studies found that work and friendship dyads exchanged and accepted procedural justice perceptions amongst themselves (Chia et al., 2006), while Simmelian-tied work dyads and groups were found to share congruent procedural justice perception (Van Den Bos & Lind, 2001; Lamertz, 2002; Colquitt et al., 2002). However, these findings have been met with inconsistency. In another study, friendship but not advice dyads shared procedural justice perceptions (Umphress et al., 2003).

We suggest that while regular advice and friendship dyads exchange and accept procedural justice perceptions, these do not necessarily mean that each individual receive other's justice reports as their own personal views. Reports of secondhand justice information may be seen as less reliable and biased compared to their own experiences (Van Den Bos & Lind, 2001). This issue may be mitigated in a
Simmelian-tied dyad, as justice perceptions may be reinforced with the presence of the third party.

Interpersonal justice perception is often associated with the subordinate’s evaluation of the supervisor (Colquitt, 2001). The sensitive nature and consequences associated with the interpersonal perceptions of supervisors means that close friends who share empathic concerns, social support, and trust are likelier to exchange such information compared to regular colleagues. Informational justice perception addresses the transparency work-related processes and outcomes (Greenberg, 1993). Considering that informational justice perception also deal with sensitive information such as treatment bias, we expect that like interpersonal justice perception, informational justice information is exchanged among close, trustworthy friends (Chia et al., 2006). Similarly, we predict that these notions will be further reinforced within Simmelian-tied friendship dyads.

In the context of our study, we aim to investigate if positive perceptions of procedural, interpersonal, and informational justice perceptions are likelier to promote greater levels of work performance and in turn motivate expert knowledge sharing. Explicit expressions of injustice have been proposed to be more susceptible to social influence and sensemaking (Shapiro et al., 2008). Therefore, in the absence of explicit sharing of injustice among Simmelian-tied dyads, we propose that it is also likely for Simmelian-tied individuals to share positive levels of these justice perceptions.

**Hypothesis H2a.** Compared to regular advice dyads, Simmelian-tied advice dyads are more likely to share positive congruent procedural justice perception.

**Hypothesis H2b.** Compared to regular friendship dyads, Simmelian-tied friendship dyads are more likely to share positive congruent procedural justice perception.

**Hypothesis H3.** Compared to regular friendship dyads, Simmelian-tied friendship dyads are more likely to share positive congruent interpersonal justice perception.

**Hypothesis H4.** Compared to regular advice dyads, Simmelian-tied friendship dyads are more likely to share positive congruent informational justice perception.

### 2.6 Workgroups and Justice Perceptions

Formal organization structure comprise of boundaries such as different workgroups, departments, and spatial locations. Depending on the type of informal ties, the increased use of virtual workgroups allows ties to span across boundaries, opening access to new information and perspectives (McEvily & Zaheer, 1999; Burt, 2004). For example, strong friendship ties spanning boundaries encourage cooperation and remove hindrances (Krackhardt & Stern, 1988).

Employees a part of virtual workgroups may have ties that span multiple countries and departments. This increases opportunities for employees to access and compare justice perceptions. We suggest that while colleagues may seek justice perceptions across boundaries (Chia et al., 2006), close friends forge congruent perceptions. Close friends demonstrate emphatic concern and support for unfair treatment in each other (Coleman, 1990; Krackhardt & Kilduff, 1990), fostering similar attitudes (Gibbons, 2004). However, colleagues place more importance on self-interests (Verbeke & Wuyts, 2007), particularly if they are from different workgroups.

Not all perceptions endure across physical boundaries. Interpersonal justice perceptions are oriented toward individual supervisors and not the organization as a whole. Perceived interpersonal justice by an employee does not apply to his/her friend with a different supervisor, and therefore do not necessarily lead to congruent interpersonal justice perceptions. By contrast, distributive, procedural and interpersonal justice perceptions may propagate across boundaries.

Close friends are likely to empathize and reinforce organization-oriented justice perceptions, such as distributive justice (Cropanzano & Ambrose, 2001). For example, an employee in location A expresses pay dissatisfaction to a colleague-friend in location B. While both employees may have different initial perceptions, both may conclude that the organization did not treat employees as equals and thus unfair. Alternatively, the employee in location B may convince the friend in location A
otherwise, or not at all. The presence of a third friend will aid in reinforcing a particular position.

Procedural and informational justice perceptions involve the ability to appeal injustice and the degree of transparency across the organization. The relative ambiguity of these justice items also prompt employees in virtual workgroups to seek out perceptions to reduce treatment uncertainties, such as the case of procedural justice (Hakonen & Lipponen, 2008). Like distributive justice, procedural and information justice perceptions are also organization-oriented rather than being supervisor-specific. Thus, it is possible that organization-oriented perceptions will endure across boundaries.

Hypothesis H5a. Simmelian-tied friendship dyads are likely to share distributive justice perceptions across distributed workgroups.

Hypothesis H5b. Simmelian-tied friendship dyads are likely to share procedural justice perceptions across distributed workgroups.

Hypothesis H5c. Simmelian-tied friendship dyads are likely to share informational justice perceptions across distributed workgroups.

2.7 Justice Perceptions and Knowledge Sharing

Knowledge sharing within the organization may be defined as the exchange of information, expertise or feedback (Cummings, 2004). While knowledge sharing has been recognized as crucial to organizational success, equally important is the type of knowledge shared. Product knowledge (e.g. documentation) saves time while expert knowledge (e.g. advice and experience) improves work quality and performance (Haas & Hansen, 2007). Employees were found to regard expert knowledge as their own and product knowledge as organizational property (Constant et al., 1994). Thus employees are obligated to share product and not expert knowledge (Jarvenpaa & Staples, 2000). Distributive justice perception is associated with commitment, citizenship (Aryee et al., 2002), and compliance behavior (Kim & Mauborgne, 1997). Unlike other perceptions, distributive justice perceptions do not induce counterproductive work behavior (Jones, 2008). Also, individuals sharing similar perceptions (such as of the organization) are likely to be closer and familiar with each other, sharing resources to solve problems (McPherson & Smith-Lovin, 1987; Gruenfeld et al., 1996; Umphress et al., 2003). Therefore, we suggest that individuals who share congruent distributive justice perceptions will be inclined to share product knowledge in order to get the job done.

Hypothesis H6. Congruent distributive justice perceptions will be positively related to product knowledge sharing.

Positive procedural justice perceptions allow employees the ability to voice concerns and influence decisions about their work. This encourages volunteering, helping, and discretionary service behavior (Kim & Mauborgne, 1997; Colquitt, 2001; Simons & Roberson, 2003; Spitzmüller et al., 2006). This motivate individuals fulfill more than the basic job requirements to share expert knowledge (Kim & Mauborgne, 1997; Spitzmüller et al., 2006), and contribute to the community (Wasko & Faraj, 2000). Interpersonal justice perception concerns the supervisor (Colquitt, 2001), and also encourages work performance akin to procedural justice perception (Simons & Roberson, 2003; Roch & Shanock, 2006; Jones, 2008). Similarly, informational justice perception instills voluntary performance-oriented behavior through trust and transparency (Colquitt, 2001; Turel et al., 2008; Ellis et al., 2009).

As procedural, interpersonal, and informational perceptions are related to employees are treated, recognized (Constant et al., 1994), we predict that positive perceptions will allow employees to share expert knowledge. Unlike distributive justice however, negative perceptions of procedural, interpersonal, and informational justice lead to counterproductive work behavior (Jones, 2008). We suggest that employees are likely to withhold utilizing expert knowledge in their work to sabotage work quality if they harbor negative perceptions of procedural, interpersonal, and informational justice items. Therefore, such individuals would be in a position to share and reciprocate expert knowledge only if they hold to congruent positive justice perceptions.
Hypothesis H7. Positive congruent procedural justice perceptions will be positively related to the expert knowledge sharing.

Hypothesis H8. Positive congruent interpersonal justice perceptions will be positively related to expert knowledge sharing.

Hypothesis H9. Positive congruent informational justice perceptions will be positively related to expert knowledge sharing.

We have discussed the effects of Simmelian-tied dyads on congruent justice perceptions, and the effects of congruent justice perceptions upon knowledge exchange. We noted that with Simmelian-tied dyads, there are greater trust, shared cognition and congruent perceptions. Considering that such strong ties between employees foster like-mindedness and congruent perceptions, these similarities in turn encourage work performance (Phillips et al., 2004) and knowledge exchange (Levin et al., 2002). Thus, we hypothesize that congruent justice perceptions mediate the relationship between Simmelian-tied dyads and knowledge sharing.

Hypothesis H10. Congruent justice perceptions mediate the relationship between Simmelian-tied dyads and knowledge sharing.

3 PROPOSED METHOD

3.1 Research Participants

We will conduct this research within virtual workgroups in a global knowledge intensive organization. Within each workgroup, members are both co-located and distributed across different continents and time zones. Communication occurs over mediums such as face-to-face meetings, and emails. The collaborative nature of these workgroups allows us an appropriate setting to execute our study.

3.2 Measures

Advice and Friendship Dyads. To capture reciprocated advice dyads, respondents will be asked to whom they provide and seek work-advice. To capture reciprocated friendship dyads, respondents will be asked whom they consider a personal friend (Krackhardt & Kilduff, 1990; Shah, 1998). A reciprocated advice dyad existed only if person \( i \) gives to or seeks advice from person \( j \), and vice versa. The same applies to a friendship dyad. This gives us the raw advice and friendship dyads.

Simmelian-tied Advice and Friendship Dyads. A hypergraph matrix illustrating every instance in which a respondent is tied to every other respondent is derived from the raw dyadic matrix obtained from the survey. From this hypergraph, the matrix of Simmelian ties is then derived to determine which respondents are Simmelian-tied (Krackhardt & Kilduff, 2002).

Workgroup Membership. To study whether Simmelian-tied dyads transmit justice perceptions within and across organizational and geographical boundaries, workgroup membership are used. We consider workgroup boundaries of members who span geographical boundaries.

Justice Perceptions. Shared justice perceptions refer to the congruence in distributive, procedural, interpersonal and informational perceptions within a Simmelian-tied dyad. Congruent perceptions within a dyad are measured by taking the absolute difference in the rated perception. The measures are adapted from Colquitt’s (2001) four-factor justice model to fit the organization. All of the items were measured using a 5-point Likert-type scale (1 = strongly disagree).

Knowledge Sharing. The degree of knowledge shared by each respondent to another respondent will be measured by indicating the frequency and type of knowledge shared over different mediums. Frequency is measured over a Likert-type scale (1 = never and 5 = several times a day).
Control Variables. We include three other variables to eliminate alternative explanations for our hypotheses. Different job hierarchy may alter justice perceptions, e.g. managers may not be open in communicating or expressing fairness perceptions to subordinates. We also control for organizational tenure. New employees may not attune to the organizational dynamics, while long-staying employees may have accepted certain practices as a norm. Finally, we control for individual centrality, which have been shown to increase the social influence of interactional justice (Umphress et al., 2003).

3.3 Study Design and Procedures

The independent variables, mediating variables, dependent variables and control variables will be collected through an online survey. The moderating variables will be gathered by the organization. The network survey will contain a list of study participants to capture the respondents’ social and communication structure. As we are examining the influence of congruent justice perceptions and reciprocal knowledge sharing within dyads, we adopt a dyadic level of measurement and analysis.

The network survey will request each respondent to note to whom they turn to for work-related advice and whom they consider friends. From this set of responses, raw advice and friendship matrices will be generated. We will utilize a social network analysis package, UCINET 6 (Borgatti et al., 2002), to calculate advice and friendship Simmelian-tied dyads. A similar procedure will be used to calculate reciprocated knowledge-sharing dyads.

Distributive justice perception congruence will be calculated by taking the negation of the absolute difference of each rated response pair. This generates a matrix where a higher value between each pair signifies greater congruence. This procedure will be repeated to generate matrices of the differences in job hierarchy and tenure for each pair. Positive congruent procedural, interpersonal, and procedural justice perceptions will be calculated by generating a matrix for each justice. Positive justice perceptions between each pair will be marked with a 1, or otherwise a 0 will be given.

In line with existing dyadic studies (Umphress et al., 2003), we will adopt quadratic assignment procedure analyses to perform bivariate correlations as this procedure demonstrated the ability to remain unbiased despite the autocorrelation of network data (Krackhardt, 1988). Hypotheses testing will be conducted with a multiple regression quadratic assignment procedure analysis.

4 CONCLUSION

Our research proposed a theoretical model utilizing congruent justice perceptions as a mediator between Simmelian-tied dyads and knowledge sharing. This study sought to address the limitations and conflicts found in existing studies to examine how social structure may potentially influence justice perceptions between employees.

The effects of congruent justice perceptions on organizational outcomes were also examined, which may potentially differ from the effects of individual justice perceptions. For example, in situations of perceived injustice, a dyadic perspective demonstrated that helping behavior might still occur between close, like-minded individuals, whereas a purely individual perspective may not yield contribution behavior toward the organization. We also attempted to explain between fulfilling minimum job requirements and discretionary behavior through congruent justice perceptions. This is particularly crucial for organizations attempting to motivate their employees into sharing and utilizing their expertise and to achieve greater levels of work performance.

This study also attempted to explain how justice perceptions might propagate across local and distal boundaries within the organization. The ability for justice perceptions to carry throughout the organization potentially holds important implications for managers and decision-makers. For example, there may be a greater need for uniformity in making and executing decisions across the organization regardless of location to eliminate perceptions of bias.
There is much room for future research. Future studies might consider the effects of other factors that may potentially affect the flow of justice information such as network size (Fang & Shaw, 2008), small world networks (Shapiro et al., 2008), and external ties (Umphress et al., 2003). Issues such as culture may also need to be readdressed in the context of justice perceptions in virtual workgroups. The interactional effects between different justice perceptions through the use of fairness theory (Colquitt et al., 2005) or fairness heuristics theory (Van Den Bos & Lind, 2001) may also reveal greater insights into how congruent justice perceptions are formed and influence outcomes.

As organizations increasingly rely on useful knowledge and adopts greater use of distributed workgroups, there is a greater need to understand the concepts and mechanisms underlying the interactional effects between informal structure and justice perceptions upon the individual contribution behavior.

References


Expressive Social Ties in Employees' Perceptions of Organizational Justice. *Organization Science, 14.*


WEB-ENABLED BOUNDARY SPANNERS AND THEIR ROLE IN THE KNOWLEDGE FLOW NETWORK

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0187.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Knowledge-based community, Knowledge Management Systems, Social Network Analysis, Case Study</td>
</tr>
</tbody>
</table>
WEB-ENABLED BOUNDARY SPANNERS AND THEIR ROLE IN THE KNOWLEDGE FLOW NETWORK

Abstract

No organisation, no matter how large, can remain totally reliant on the stocks of new knowledge it generates itself. In order to keep abreast of the latest scientific and technological developments, R&D organisations must continuously import knowledge from beyond the organisations boundaries. How does this external knowledge which is critical to success then become absorbed and integrated into the firm? Our paper addresses this question through the lens of the influential technological gatekeeper theory. Drawing on social network analysis (SNA) and interview evidence from a medical devices R&D group, we find that the gatekeeper role is still vital, but no longer needs to be performed by a single individual. Instead, the modern R&D group can keep abreast of the latest technological advances through a combination of Web-enabled internal and external communication specialists. A unique contribution this paper makes to the IT-enable social network literature is the development of an updated conceptual framework of how the gatekeeper role is performed in the modern R&D group.

Keywords: Web Technologies, Socio-Technical, Technological Gatekeeper, Boundary Spanners, R&D.
1. INTRODUCTION

Research and development (R&D) intensive firms must keep abreast of and adapt to the latest scientific and technological developments in order to remain competitive. Increasingly, however, knowledge vital to the firm’s competitive advantage is located outside the firm’s boundaries (Chesbrough 2003). How does this external knowledge which is critical to success then become absorbed and integrated into the firm? A large body of literature exists which highlights the importance that both formal and informal social networks play in transferring knowledge within and between organisations (Allen and Cohen 1969; Allen 1977; Hansen 1999; Wenger, McDermott et al. 2002; Assimakopoulos and Macdonald 2003; Wasko and Faraj 2005; Su, Mark et al. 2007). However, the rapid adoption of Web technologies in recent years is challenging our extant understanding of what constitutes a social network. Today, not only can a person be a member of close, face-to-face social networks, but he or she may also participate in numerous Web-based collaborations comprising thousands of globally dispersed unknown and anonymous individuals. Research is needed to inform organisations of how business value can be generated from the interplay between the social and technical aspects of these socio-technical systems (Parameswaran and Whinston 2007). Thus, the purpose of this exploratory paper is to address this research gap by examining how external knowledge is accessed through Web-based channels and personal contacts and then disseminated throughout the R&D group.

To fulfil our purpose, we turn to the concept of the technological gatekeeper (Allen and Cohen 1969; Allen 1971; Allen 1977; Tushman and Katz 1980; Katz and Tushman 1981; Tushman and Scanlan 1981). Through decades of innovation research, the role of the gatekeeper has proven to be a critical factor in understanding the performance of R&D organisations as gatekeepers have served as key nodes in the innovation process – acquiring, translating, and disseminating external knowledge throughout the R&D organisation. The question arises, however, as to how the role and tasks of the gatekeeper are changing due to the ability of every professional in an R&D organisation to quickly and easily access external knowledge through Web-based channels. Thus, after a review of the seminal literature on gatekeepers as well as of the relevant IS literature, we develop the research question: how have Web technologies impacted the technological gatekeeper’s tasks of acquiring, translating, and disseminating external knowledge? We then present our results from a case study of a medical device company in Ireland in which we collected social network and semi-structured interview data. Finally, the paper presents an updated gatekeeping conceptual framework and concludes with the implications of our findings on research and practice relating to organisational knowledge and learning.

2. THE TECHNOLOGICAL GATEKEEPER

R&D groups are the drivers of innovation in high-technology firms. In order for the group to sustain itself, the literature on R&D innovation emphasises the importance of acquiring a diverse and novel body of knowledge from beyond the organisation’s boundaries (Allen 1977; Tushman 1977; Aldrich and Herker 1997). The acquisition of external R&D knowledge helps the firm to build its ‘absorptive capacity’ (Cohen and Levinthal 1990) and will serve as the seeds for future technological developments (March and Simon 1958; Leonard-Barton 1992). A rich stream of research throughout the 1970s and early 1980s examined the processes through which scientific and technological knowledge enters the R&D group. This particular stream was headed by MIT’s Thomas Allen and his seminal book Managing the Flow of Technology (Allen 1977) documents over a decade’s worth of studies with some of the largest American R&D corporations. Using social network analysis techniques, Allen discovered that knowledge of the latest scientific and technological developments entered the R&D group through a two-step process. Not every R&D professional was directly connected with external sources of knowledge. Instead, a small minority had rather extensive contacts and served as sources of knowledge for their colleagues. These individuals were termed ‘technological gatekeepers’ (Allen and Cohen 1969; Allen 1971; Allen 1977; Tushman 1977; Allen, Tushman et al. 1979; Katz and Tushman 1981; Tushman and Scanlan 1981) as they act as the conduit through which knowledge of external technology flows into the R&D group. A more formal
definition explains that technological gatekeepers are those key individual technologists who are strongly connected to both internal colleagues and external sources of information (Allen and Cohen 1969; Allen 1977; Tushman and Scanlan 1981).

It is logical to assume that direct communication between R&D professionals and external sources of knowledge would be a more efficient knowledge integration mechanism than a two-step process. However, studies have found that widespread direct contact by all project members is not an effective method for transferring technical knowledge into a project from external sources (Katz and Kahn 1966; Allen 1977; Tushman 1977). The reason for this phenomenon relates to task specialisation and the evolution of local norms, values and languages that emerge as a result (Tushman 1977). Not every individual has the ability to understand contrasting coding schemes and misinterpretations are likely to occur if one communicator is without knowledge of the others local coding scheme (Cherry 1965).

Thus, scholars have argued that specialised boundary spanners are required to facilitate the transfer of knowledge across intra and extra-organisational boundaries (Tushman 1977; Tushman and Scanlan 1981; Bouty 2000; Teigland and Wasko 2003; Cross and Parker 2004). The technological gatekeeper is one such boundary-spanner who mediates between the local R&D Group and the world beyond the firm’s boundaries. Allen and Cohen (1969) noted when studying gatekeepers in the R&D division of a large aerospace firm that "...if one were to sit down and attempt to design an optimal system for bringing in new technological information and disseminating it within the organisation, it would be difficult to produce a better one than that which exists". The inference is that there is an association between gatekeepers and higher performance, however, no empirical data existed until the early 1980s when a number of studies advanced Allen’s original gatekeeper concept by examining the relationships between the existence of gatekeepers and project performance for different types of tasks (i.e. research work vs. development work). Tushman and Katz (1980) and Katz and Tushman (1981) found that development projects with gatekeepers were significantly higher performing than those without gatekeepers. Thus, development projects are higher performing when external communications are monopolised by a small number of individuals.

Given the vital role which gatekeepers perform in development projects, it would be useful for R&D managers to be able to identify these individuals. While there is no pre-requisite checklist that an individual has to conform to, the literature does provide some clues to recognising those performing the gatekeeping role. The original studies of Allen and Cohen (1969) and Allen (1971;1977) suggest that the gatekeeper is a highly competent technical performer who is likely to be a first line supervisor. Seldom were gatekeepers found with fewer than five years organisational experience as it takes time to develop one’s communication network. They are genuinely interested in keeping abreast of developments in their technology domain and knowledge of their specialty is deep as opposed to wide-ranging. They tend to read the harder-literature (e.g. scientific journals), present more papers at technical conferences, and maintain long-term relationships with colleagues outside their own organisation. The gatekeeper’s principle contribution comes by way of the translation that they can perform (Allen 1977). The gatekeeper can convert knowledge gained from journal papers and personal contacts into terms that are understandable by members of the local R&D group. It is because of this ability and their technical competence that they are frequently sought out by their colleagues.

The gatekeeper concept has received modest attention since the early 1980s, presumably because the likes of Allen, Katz, and Tushman ploughed the field so thoroughly and left little for other scholars to explore. In recent years however, interest in the concept has been reignited, particularly in the IS field. The gatekeeper existed in a time when it was a difficult and time consuming process for the average R&D professional to acquire knowledge from beyond the company’s boundaries. The past decade has borne witness to major advances in ICT and particularly Web technologies. What these advances have changed is the ease and speed with which employees at all organisational levels can access and disseminate knowledge (Teigland and Wasko 2003; Whelan 2007). With a PC and an internet connection, a knowledge worker can join computer-supported social networks to seek solutions, share expertise, and discuss ideas with like-minded individuals far beyond the reach of their
local social network of friends, contacts, and colleagues (Wasko, Faraj et al. 2004). Recent ethnographic research by Su, Mark et al. (2007) has found that throughout the working day knowledge workers constantly switch between multiple social networks, all of which are a complex mixture of formal and informal, face-to-face and computer-mediated, intra-organisational and extra-organisational, and work-related and private interactions. Yet, we have a limited understanding of how the interaction between the social and technical aspects of practice-based networks impact the knowledge flow network. We address this gap by examining how Web technologies have impacted the gatekeeper processes in development focused R&D:

**RQ:** How have Web technologies impacted the gatekeeper’s tasks of acquiring, translating, and disseminating external knowledge?

We have specifically chosen to examine the impact of Web technologies as an earlier pilot study by some of the authors (Whelan and Donnellan, 2008) found that the key communication technologies used by R&D engineers to acquire and distribute technological information were websites, search engines, and email. Hence, our definition of Web technologies centres on these applications. For the purposes of this study, we define Web technologies as “Web-based communication technologies, such as websites, search engines, and email that enable the easy exchange and retrieval of digitized content.”

### 3. METHODS

For the purpose of our research, case study methods are appropriate as the objective of the study is theory building (Eisenhardt and Graebner 2007), there is a need to focus on contemporary events (Benbasat, Goldstein et al. 1987; Yin 1994), and the phenomenon of interest cannot be studied outside its natural setting (Yin 1994). In order to compare with the original high-technology engineering gatekeeper studies, we have collected data from MediA, a medical device firm who have requested to remain anonymous. The case study setting is further described below.

#### 3.1 Case Setting - MediA

MediA is an American multinational that has been in the medical device business for over 25 years with an annual turnover of $8.3 billion. MediA employs approximately 4,200 people in Ireland. The company has advanced the practice of minimal-invasive medicine by providing a broad and deep portfolio of innovative products, technologies and services across a wide range of medical specialties. The company employs approximately 3,000 R&D engineers, scientists, and technicians worldwide. While the majority of these are based in the US, an R&D group comprising 76 professionals are co-located in MediA’s Irish subsidiary (referred to as Group A in the rest of this paper). While a high level of collaboration exists between the Irish and US R&D bases, Group A is largely a stand alone entity. Both the Irish and US groups are design owners of certain products, and it is the responsibility of each group to advance those designs. Group A is organised on a functional basis into four specialist subgroups – Drug Eluting Stents, Test Method Development, Vascular, and Stent Delivery Systems. Each of the four subgroups has a technology brief which relates to a specific part of the product design.

#### 3.2 Data Collection and Analysis

Data were gathered from Group A between the months of February and March 2008. The data collection methods are summarised in table 1.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1 – Social network analysis</td>
<td>Online survey issued to group members, n=76, 92% response rate</td>
</tr>
<tr>
<td>Phase 2 – Semi-structured interviews</td>
<td>n = 11, recorded and transcribed Interviewees = 3 gatekeepers, 2 external stars, 4 internal stars, 2 non-stars</td>
</tr>
</tbody>
</table>

*Table 1 Data Collection Methods*
Phase 1 involved analysing the flows of knowledge into and around Group A using social network analysis (SNA) techniques. SNA or sociometry is an established social science approach of studying human relations and social structures by “disclosing the affinities, attractions and repulsions between people and objects” (Moreno 1937). SNA views social relationships as nodes and ties which can be illustrated visually and mathematically. As such, it can provide an x-ray into the inner workings of a particular network. With this tool, important patterns become visible, the relationships between people can be better understood, the health of a group can be assessed and, the people playing key roles within the group can be identified (Cross and Parker 2004). In recent years, SNA has found increasing use as a structured way to analyze the extent of informal relationships that exist within various formally defined groups (Cross, Nohria et al. 2002). However, despite the knowledge intensive nature of R&D, SNA of the R&D function remain relatively rare (Allen, James et al. 2007).

The purpose of phase 1 was to identify the ‘stars’ of the knowledge flow network. To collect these data, all group members were asked to complete a short online questionnaire on their internal and external communications. To measure internal communications, we used the question asked by the original gatekeeper scholars – ‘Please identify which work colleagues you discuss technical issues with at least once a week?’ The choice of once-a-week frequency is purely arbitrary although it does represent a fairly heavy degree of consistent communication (Allen 1977). To measure external communications, respondents were asked to indicate how often they used three sources of external knowledge: personal contacts, internet, and academic publications. An earlier pilot study by one of the authors at MediA indicated that these three knowledge sources were the most frequently used by R&D professionals when acquiring knowledge from outside the company. We used the SNA software package UCINET (Borgatti, Everett et al. 2002) to illustrate the knowledge flow network in Group A.

In phase 2, we conducted semi-structured interviews with selected members of Group A. Interviewees were selected based on the SNA results from phase 1. The objective of the interviews was to explore how the use of Web technologies impacts the acquisition, translation, and dissemination functions of the technological gatekeeper. Group A members were categorised as being a gatekeeper, an internal communication star, an external communication star, or a non-star. Following the approach of (Allen 1977; Tushman and Katz 1980; Katz and Tushman 1981), this study operationalised gatekeepers as those individuals who were in the top fifth of both the internal and external communication distributions. Internal stars were operationalised as those individuals in the top fifth of the internal communication distribution but outside the top fifth of the external communication distribution. The reverse applies for external stars. To get a non-biased view of how knowledge flows around the R&D group, we interviewed a sample of gatekeepers, external stars, internal stars, and non-stars. Care was also taken to ensure that all levels of the formal group hierarchy were represented in the interviewee sample. All interviews were conducted face-to-face and ranged in length from 30 minutes to 75 minutes. In addition, all interviewees gave permission for the interview to be recorded. The procedures outlined in the dramaturgical model (Myers and Newman 2007) were adopted in order to ensure that high-quality interviews were conducted. Interview data analysis was performed using the NVivo software package and followed established inductive qualitative methods: coding, data categorization, and pattern identification (Miles and Huberman 1984; Eisenhardt 1989; Yin 1994).

4 FINDINGS

4.1 Social Network Analysis

Figure 1 illustrates the flow of technical knowledge into and around Group A. The nodes in the diagram are the individual members of the R&D group and the lines represent the flow of technical knowledge between them. The external stars are represented as triangles. The size of the triangle is reflective of how well connected that individual is to external knowledge sources. For example, Node 52 is the biggest triangle as this individual is the most frequent user of external knowledge sources.
Nodes 23, 33, 42, 46, 60 and 75 did not complete the questionnaire hence the reason they are isolated on the left. Nodes 2, 18, 21, 41, 56 and 69 completed the questionnaire but are also isolates because they have no reciprocated interactions with another group member. The overall reciprocation rate in Group A was 64%.

![Figure 1: Group A’s Knowledge Flow Network](image)

The SNA data reveals that only 4 members (or 6%) of the group can be classified as technological gatekeepers. The gatekeepers of the group are nodes 5, 9, 11, and 54. Rather than relying on single individuals to both acquire and disseminate external knowledge, the SNA evidence shows that one set of boundary spanning individuals acquire external knowledge, and a largely different set of individuals disseminate this knowledge around the group. The relationship between node 62 and node 66 can be used to demonstrate this process. Node 62 is an external communication star. This individual is well connected to external knowledge sources but is not very well connected internally. Node 62 acquires external knowledge and communicates this to node 66. Node 66, on the other hand, is well connected internally and can distribute this knowledge around the group through his or her many connections. In fact, many of Group A’s external communication stars have low levels of internal communication, hence the reason why they are located on the periphery of figure 2. The average number of reciprocated internal interactions per week in Group A is 4.343. Nine of the 14 external communication stars fall below this level.

Where do the external stars acquire their knowledge from? The Web was by far the most widely used source with 79% of external stars using this knowledge source at least once a day. 29% reported consulting academic publication while only 21% would consult an external colleague on a daily basis. Thus, knowledge from beyond the company’s boundaries is acquired by the external communication
stars who predominately use the Web to acquire this knowledge. This knowledge is then passed to the internal communication stars who distribute that knowledge around the R&D group.

4.2 Semi-structured Interviews

External Knowledge Acquisition

The SNA of Group A reveals that external knowledge flows into the group via external communication stars who predominately use the Web to acquire this knowledge. There are a number of specific medical technology websites which these external stars access in order to keep up-to-date with developments in the field. For example, cvpipeline.com is one website that many of the external stars identified as being a good source of external knowledge. Cypipeline.com is a subscription based service that promotes itself as “a new online database solution that keeps you up to date with emerging companies, products, technologies, people, and clinical studies in the fast-changing cardiovascular market” (www.cvpipeline.com). One theme in the interviews with Group A’s external stars focused on the benefits that websites such as cvpipeline.com offered over traditional sources of external information e.g. conferences and journal articles. As is reflected in the following quotation, the prime advantage of the Web relates to the ease with which technical professionals can keep abreast of the latest developments in the industry:

I would use the internet quite a bit. For my own development I use it to keep up-to-date with new technologies, new medical device developments. Recently I subscribed to a [trade] magazine…which I think is very good for providing information on new technologies and new medical devices outside. I think another good source of information would be attending conferences…but I think that can be got through the internet. You’d get current information, very up-to-date. I think the internet is a great source of information in that way…it’s there at your fingertips and it’s just a matter of using Google. But subscriptions to magazines, attending conferences, attending procedures over in the hospital and watching the ‘docs’ do their stuff…they can all be good sources of information too.

The external communication stars interviewed not only scan their own industry for the latest developments, but also monitor advances in related industries like electronics and pharmaceuticals. Some of the best innovations in the medical devices field have actually come from other industries. This is explained by one of the gatekeepers. He highlights in the following quotation that the Web provides him with access to a broader range of external information and this is used to stimulate more and better ideas:

There would be a lot of overlap between what we do, and say, the electronics industry. The physicists and electronic engineers in companies like Intel have been coming up with ideas and solving problems for years and the medical device industry has said “Oh, hold on a second now, that could be very useful in this application”. Inkjet technology for example…Hewlett Packard has developed that technology down to such precise detail…inkjet technology is actually being used now for injections in low dosages onto the tiniest medical devices that you can think about. The internet keeps you in touch with those industries. You might see some new drug delivery system treating some obscure disease that had nothing to do with our industry. You find out that they took X and Y and sorted the problem out. Now we can try a similar approach with our problem. [The internet] stimulates a thought process rather than sorts something out for you there and then.

External Knowledge Translation

A number of interviewees commented that it took a number of months working at the company for them to become familiar with the technical jargon and abbreviations unique to Group A. In order to be usable by Group A, knowledge acquired from the external environment needs to be translated into a form consistent with these local norms. However, as is evidenced in the following quotation from an
internal star, these contrasting coding schemes create a problem when communicating with external contacts:

*It’s mind blowing how much jargon we have – not jargon – more abbreviations. We have abbreviations for everything. I’d say within the [this] R&D group, there is probably no real misinterpretation of abbreviations. People usually understand what you mean when you say something like that. However, when you go outside R&D – if I was talking to customers, which I do sporadically, probably about once or twice or three times a year…it’s like I’m talking a completely different language. The jargon and abbreviations are needed but they can be a barrier.*

The analysis of the interview data reveals the existence of a small number of individuals who perform the knowledge translating function for Group A. These individuals are frequently sought out by their colleagues, hence the reason they are also likely to be internal communication stars. One interviewee, Chris1, acknowledged that many of his colleagues often consult with him when they have discovered novel external information. The analysis of the SNA questionnaire reveals that while Chris is one of the most connected people internally, he has very low exposure to external sources of knowledge. Chris is a senior person in the Group A and has 13 years experience in the medical device field, eight of those with MediA. The knowledge translation discussions he has with his colleagues are almost always conducted face-to-face and focus on figuring out if and how outside knowledge can be used by Group A. While he believes that his colleagues consult with him because of the formal reporting structure, he also suggests that he has certain skills which are useful for translating external information. In the following quote, Chris explains these skills. As he is well connected internally, he has the ability to see the bigger picture within the wider R&D group, and he understands how external knowledge needs to be modified in order to fit into that bigger picture:

*So anything mechanical related...people would probably run it by me just to make sure it makes sense. The reason for that is probably – I wouldn’t say it’s my technical expertise – there are a lot of people in the group that would burn me in terms of pure technical expertise. My skill sets would lie in that I know a little about a lot of different things, and I probably have a good appreciation for how they all fit together into the overall picture. We do have people who are bond experts, who are crimping experts, who are balloon experts, [but] they probably wouldn’t have as good an appreciation for the impact that something new would have on other people...whereas I probably would have that visibility. The skill set I have – other people probably don’t have that.*

While the Web is the most widely used source of external information, there is a realisation within the group of the need to be selective when gathering web-based information. There are no guarantees that information sourced from the Web is truly accurate. The medical device industry is highly regulated and the information used to produce these products has to be documented for FDA and EU inspection. Popular websites like wikipedia are extremely convenient for explaining a particular topic however, anyone in the world has the potential to edit a wikipedia article. Thus, the reliability of this information is always open to question. One internal communication star acknowledges that while wikipedia is frequently used as an information source, the validation of this source information is an important process. Group members cross-reference web-based information against other data sources to check its validity:

*I’ve heard comments where people talk about something like wikipedia [but] you have to be careful with it. I suppose I’m guilty of it myself – it’s just convenient, you just pull the information. If it’s just for illustration purposes it’s not a problem. But if it’s something where you’re probably going to rely on this as a source to make a decision or to go and use it in support of a submission to a Regulator, then yes – clearly you have to go and check the source of the information. And we do enough cross referencing, reviewing or peer reviewing*

---

1 Fictitious name

Proceedings ECIS 2009
of our internal documents, and that in itself is the catch for it. We go look for a source document.

External Knowledge Dissemination

Through a combination of social and technical means, the internal communication stars disseminate and integrate novel external knowledge into Group A. Email is the bedrock of the internal communication system. Many of the interviewees suggest that they easily receive over 50 emails per day. Group A has a very clear hierarchical structure and the flow of knowledge tends to follow the chain of command. The following quotation from an internal communication star provides evidence that the pattern of email traffic conforms to the formal organisational chart:

We have a lot of distribution lists. We all have functional managers and they have managers. So I report to my boss. He may have 5 or 6 engineers and a number of technicians in his group – he distributes to us. I’d have a couple of technicians reporting to me and I’d keep them on a little distribution list. I’d have 3 or 4 people I work with outside the area – my peers – on another list. It just goes on and on. So my manager’s boss, he would have all of us in stent delivery on his distribution. Everyone has their own little tiers of it.

Email is the primary system used to alert colleagues to new information from outside the company. However, there is an expectation that only the more senior people in the group should be sending out these emails. There is a feeling among the younger and less experienced members of the group that any external knowledge they discover would not be accepted by their colleagues if they tried to distribute it themselves. Instead, they usually ask an internal communication star to distribute that knowledge around the group on their behalf. The process of disseminating novel external knowledge usually begins with the internal star sending an email with the attached information to the group members they know would be interested in that information. The email will include one or two sentences explaining why the sender believes the attached information is relevant to the receiver. If the information is of interest to that individual, they then return to the internal communication star and have a face-to-face discussion about how that information can be used by the group. An example of this process is provided in the following quote from a non-star:

The information would go up the chain and then fed back down. During the week, one of the members of our group found an interesting external training course, sent an email to the functional manager who was my boss as well, and the manager sent it out to all our extended teams saying, “We’ve just found this, if we’re interested, please come back to me”. The same applies with new recruits from universities…they may have spent a lot of their last year or two in conferences. They may have exposure to new developments that people here may not have known about. They would communicate that to the functional manager, and the manager will then feed it out to everyone.

While email is used to distribute external knowledge, face-to-face discussions are needed in order to figure out how to make use of that knowledge. The consensus among the interviewees is that it would be virtually impossible to perform the work of the R&D group without face-to-face interaction among peers. R&D work in the medical devices sector is very technical in nature. Email is useful for alerting people to external developments but a discussion about that knowledge through email is cumbersome. For this very reason, group members frequently travel to their sister site in the US. In fact, when engineers are being interviewed for positions in R&D, it is explained to them that travel is part of the job. Face-to-face time is a vital component of the group’s work. This is highlighted in the following quotation from an internal communication star. He explains that integrating new knowledge into the group requires face-to-face discussion:

You can only truly understand something new if the other person asks questions and you reply straight away, so that you can address their needs straight away...whereas in e-mail you can’t do that. You do need face-to-face time. You can do a certain amount over email and the phone but you have to build up that face-to-face rapport. What happens is once you build up
that face-to-face rapport, people get the measure of you. They understand what your convictions are, where your strengths are, how you behave – or misbehave – and how to manage that.

5. DISCUSSION

This paper asked the question “How have Web technologies impacted the technological gatekeeper’s tasks of acquiring, translating, and disseminating external knowledge?” While we find that the gatekeeping tasks are integral to the R&D operation, we also find that these tasks no longer need to be performed by a single individual. Gatekeepers do exist, but they are rare. When Allen (1977) first formulated the theory, the gatekeeping role could only be performed by a single individual because technical communications were predominately oral based. Among other skills, the traditional gatekeeper needed excellent social networking abilities in order to effectively acquire and disseminate knowledge orally. While other R&D engineers may have wanted to perform the gatekeeping role, the lack of these social networking skills possibly impeded them. Combining the results of this study with the literature, an updated conceptual framework of the gatekeeper concept is illustrated in figure 2. We acknowledge however that the framework is a simplistic representation of an extremely complex process.

![Figure 2 An Updated Gatekeeping Conceptual Framework](image)

The framework explains that outside knowledge is largely brought into the R&D group by external communication stars. External stars primarily use the Web to scan and acquire their knowledge. External knowledge is then presented to the ‘go-to’ people of the R&D group – the internal communication stars. As well as translating the knowledge into a form that can be used by the group, the internal star also validates that the outside knowledge is accurate and reliable. Knowledge validation is an important step in the knowledge integration process for the modern R&D group but it is not discussed in the original gatekeeper studies. The process of disseminating novel external knowledge usually begins with the internal star sending an email with the attached information to the group members they know would be interested in that information. The email will include one or two sentences explaining why the sender believes the attached information is relevant to the receiver. If
the information is of interest to the receiver, they then return to the internal communication star and have a face-to-face discussion about how that information can be used by the group. It is through this discussion that learning occurs and the knowledge becomes internalised (Nonaka 1994).

6. CONCLUSION

The findings of this paper are of benefit to both theory and practice. We contribute to the advancement of the gatekeeper theory into the 21st century. We show that the gatekeeper role has fragmented, enabling it to be performed by Web-enabled boundary spanners and internal communication specialists. This study should be of particular interest to the IS community. Practitioners are increasingly aware that innovative knowledge is located beyond the boundaries of their firm. This study finds that the Web is a vital tool for accessing this knowledge and that certain people exist who have the innate ability to find relevant knowledge on the Web. It will be increasingly important for R&D firms to find people with the right blend of social and analytical skills. We have also identified some negative aspects to the Web-enabled R&D group that managers will need to be aware of. Firstly, with so much information freely available on the Web, verifying the accuracy and reliability of this information is becoming a critical step in the knowledge integration process. Managers will need to ensure that proper verification procedures are in place. Secondly, an interesting finding was the existence of certain individuals who constantly send FYI emails to their R&D colleagues. Rather than enhancing knowledge flows, the actions of these individuals are seen as a hindrance. The FYI phenomenon could become a significant problem if group members become overloaded with irrelevant messages and end up not reading the important messages.

We see two additional areas for future research. Firstly, while our findings make a unique contribution, they are based on only a single case study. For the purposes of generalisability, future research studies should examine multiple R&D groups in differing industries. Secondly, our findings show that the gatekeeping role can be performed by a single individual or by a combination of internal and external communication specialists. Future research needs to examine which of these routes is most effective for R&D project performance.

References


ARE “DESIGN NETWORKS” SHAPED BY THEIR OWN OUTCOMES? COORDINATION PROCESSES BETWEEN ACTORS AND ARTEFACTS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0227.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Coordination, Collaboration, Network analysis, Web services</td>
</tr>
</tbody>
</table>
ARE "DESIGN NETWORKS" SHAPED BY THEIR OWN OUTCOMES? COORDINATION PROCESSES BETWEEN ACTORS AND ARTEFACTS

Francesco Bolici, OrgLab-DIAM, Cassino University, via Sant'Angelo, 03043 Cassino, Italy, francesco.bolici@eco.unicas.it
Francesco Virili, OrgLab-DIAM, Cassino University, via Sant'Angelo, 03043 Cassino, Italy, francesco.virili@eco.unicas.it

Abstract

This paper sets the basis for a research project focused on collaborative social network’s genesis and dynamics. It introduces a research framework for the empirical investigation of a network focused on the design of a shared artefact, the so-called "Web services Architecture". Our hypothesis is that network artefact’s characteristics, seen as the final outcome of a collaborative process, influence and drive the genesis and the structure of the social network that is designing it. We embraced this view in order to avoid a limitation of the traditional perspectives that consider the network structure as exogenous and stable. We consider the reciprocal influence between the artefact and the social network structure, with a phase in which the desired artefact may shape the network genesis and a phase in which the emergent network’s structure may drive the artefact design.

Keywords: social network genesis and dynamics, design network, Web services, coordination, boundary object
1 INTRODUCTION

This paper sets the basis for a research project focused on collaborative social network’s genesis and dynamics. This work aims at the empirical investigation of a social network collaborating in the design of a shared artefact, the so-called "Web services Architecture" (WsA).

We named "design networks" those social networks that exist in order to design an artefact, like technical standards or open source software. We are interested in verifying whether the characteristics of the network artefact, seen as the final outcome of a collaborative process, may influence the genesis and the structure of the design network that is producing it. The traditional causal relationship between network structure (often seen as the independent variable) and network outcome (often seen as dependent variable) may be reversed in design networks: in our case, the (desired) characteristics of the artefact (WsA) may influence the genesis, the structure and the activities of the collaborative network. We therefore hypothesize that one important factor in the genesis of a design network is its fit with the designed outcome (i.e. the network artefact). If our intuition is correct, there could be situations in which the process of design a specific object (with certain characteristics) rather than another one (with different characteristics) shapes and influences the genesis and the dynamics of the social network committed to design the artefact.

The network artefact here investigated is the WsA, a rapidly emerging standard architecture recently defined by the W3C standardization consortium for distributed componentized software applications over the Web (Alonso et al. 2003). The WsA is a set of guidelines enabling any system to remotely access and use software components as “Web services”. The “Web services” concept is a relevant innovation, attracting significant investments in the software industry. The use of the Web services makes it possible to assemble a software application with several independent parts (software components), each accessible over the Web. Using Web services, programmers can compose a "virtual" software application like a puzzle, by accessing the different pieces (Web services) over the Web, independently of their physical location. In principle, many different (and incompatible) implementations of Web services may be possible: standardized architectural guidelines are therefore crucial to ensure true compatibility and interoperability of systems based on Web services by different vendors.

Within the W3C consortium, more than 40 different companies all over the world collaborated for a two-year period toward the definition of this artefact, with weekly teleconference meetings, mailing lists and conferences. The company experts collaborated not only to harmonize different technical specifications, but also to actually face difficult technical issues including not only compatibility, but also security, privacy and reliability of the overall system. The entire standardization process is therefore in the same time a negotiation process (aiming at finding a common definition) and a design process (aiming at solving complex technical issues and devising innovative use scenarios).

With the help of social network analysis techniques and tools, we are here interested in exploring the complex web of network relationships between the actors during negotiation and design of the artefact "WsA". More specifically, we are interested in verifying whether the characteristics of the final outcome of the collaborative process, may actually influence the structure of the network that is producing it.

In comparison with the mainstream literature on social network analysis, this study has two main peculiarities: 1) it is focused on design networks, i.e. social networks aimed at the production of an artefact like open source software or standard technical specifications; 2) it is investigating the influence that the network outcome could have on the network genesis and structure (and not only the opposite relationship as traditionally analyzed by literature).
2 THEORETICAL BACKGROUND AND CONTRIBUTION

Traditional social network analysis (SNA) was originated in Social Psychology, and then applied to several different fields and levels of analysis, including for example inter-firm relationships, using categories and concepts like centrality, structural equivalence and clique analysis (Grandori & Soda 1995). Typically, SNA investigates how the characteristics and the structure of the network may influence the organizational or economic outcomes. For example, some organizational studies discuss the role of relationships in generating innovation (Ahuja 2000) and competitive advantage in markets (Gulati, Nohria, & Zaheer 2000); others focus on the influence of social networks on knowledge creators and spreaders (Singh 2005). In Granovetter’s works (Granovetter 2005;Granovetter 1973) a thorough analysis is proposed on social networks, and why and how they may influence economic outcomes like hiring, prices, productivity and innovation. The same interest emerges from several economic studies, both theoretical, like (Rees 1966) and (Montgomery 1991), and empirical, like (Calvo-Armengol & Jackson 2004), showing the effects of social networks on employment and wage inequality or (Rauch 2001), discussing the role of social and economic networks in international trade. We recognize that the traditional perspective network structure drives outcome is crucial for the understanding of the network relevance and effects, but we also suggest that – particularly in design networks - the outcome itself (i.e. the artefact) may play a key role for the emergence and the genesis of the network structure of a successful collaborative project.

The debate about the endogenous or exogenous nature of the changes in the structural network’s characteristics is still open. The attention of the researchers have been captured by the intangible aspect of the relationships between nodes and had not investigated any influence that the object itself, often only considered as the results of the network activities, could have on the network dynamics. Instead, our perspective, though recognizing the influence of the network structure on the outcomes, is based on the idea that the network artefact can influence the network genesis and dynamics, especially for creative design activities. This approach not only aims at throwing new light on antecedents of social network emergence and dynamics (Carugati & Bolici, 2006; Virili 2003); it is also an answer to a call for more research focusing on IT artefacts and on their direct influence on organizational processes (Orlikowski & Iacono 2001).

In order to give a contribution to this emergent research area, we will address the following research question: does the desired outcome of a specific design project (i.e. the artefact) influence the genesis and the structural characteristics of the network? In particular, we are asking: how do the characteristics of the artefact influence the design process? How do they eventually shape the actors positions, roles and relationships in the network? Answering to these questions would mean not only showing how the artefact could shape network structure and dynamics, but also covering some ground towards showing how the artefact itself may play a crucial and central role in the overall coordination of the design process.

The contribution of this research work is thus twofold: it shows the role of the working standard definitions (the social network outcome) in influencing the genesis and structure of a design network; it sets an empirical ground for further analysis of artefact-centered coordination processes in design networks.

3 METHODOLOGICAL SET-UP

In the W3C consortium, there is an “activity” for each technology being standardized. For each activity there may be several workings groups, each one focused on the production of a technical specification. Our analysis is focused on the WsA Working Group, which has the objective to define the technical specifications of a standard architecture for the Web services activity. Every working group has one or two leaders. There are two modalities of interaction: mailing list or group meetings – they can be either face to face (F2F) or distributed (technology mediated)--. The charter document
specifies that "to be effective every working group should have from 10 to 15 active participants". The F2F meetings don't have a predetermined frequency; they are settled by the group leader according to the matters to be treated, to the deadlines and the opportunities of co-location of side events (e.g. conferences, other W3C meetings). The distributed meetings are scheduled at least once a week (twice when required by the deadlines). The participation to the F2F meetings is limited, on invitation of the group leader. Guests or external experts may be occasionally invited.

4 DATA SET DESCRIPTION

All these forms of discussion and negotiation (mailing list, F2F meetings and teleconferences) are recorded, and the scripts are publicly accessible via the working group Web site. This collection of documents, eventually integrated with external information sources like specialized press articles and news, is illustrated in Fig. 1 here below. All the teleconference scripts and the public mailing lists are available for analysis on the W3C web site.

Fig. 1. Structure of the design process and – on the right – the empirical traces left.

The available data sources for the WSA activity are: deliverables, meeting records, mailing lists and the external sources (like technical magazines and newspapers). The published deliverables include: the WS Architecture and the supporting documents (requirements, glossary, usage scenarios, service life cycle, and ontology). For each document all the previously published releases are accessible, together with the list of the modification intervened. For example, the first official version of the Web services architecture was published in November 2002, followed by two major updates in May 2003 and February 2004. The editor copies are also available: the first editor copy of the Web services architecture is dated June 2002, and was followed by updates in August 2002, March 2003 and July 2003. There are about fifty meeting records per year (on average the meetings were held once a week); they are usually written or edited by a scribe nominated for each session. The working group mailing list, during the period from February 2002 to February 2004 (25 months), collected 6895 messages,
with an average of about 276 messages per month. The two “hottest” months were July 2002 (546 messages) and January 2003 (516 messages); the “coldest” were November and December 2003 (32 and 86 messages, respectively). Moreover, the working group created two temporary task forces that produced proposals on specific issues. The Management task forces produced four deliverables; the Security task force produced one deliverable. The task forces mailing lists are also available for consultation.

4.1 Exploratory data analysis

The strategy for initial data analysis was the following: 1) selection of an appropriate data sample for text analysis; 2) exploratory analysis of the relationship network-artefact, based on initial text labeling.

1) Sampling - The document sample initially selected for this preliminary analysis, was formed by the weekly teleconference scripts for the first year of activity of the working group. It is about 45 scripts, dating from February to December 2002, representing about half of the whole standardization period, focusing on the initial phases design giving birth to the first four drafts of the standard architecture. At the end of this year the standard was already at an advanced stage of design.

2) Labelling and first exploration of the meeting notes - The very first information we looked for was about text size and text authorship for the discussions documented in the scripts. To this aim we had to label more than 4,500 text passages, assigning an author name to each of them. The textual labelling was done with the software “NVivo” version 2.0 (Gibbs 2002).

Text size can be measured as number of text passages, or as number of characters, words, etc. The distribution along time of text sizes, measured as number of text passages is shown in Fig. 2 below.

The graph illustrates the evolution over time of the textual discussion reported in the scripts of the weekly teleconferences and of the four face to face (F2F) workshops in April, June, September and November. The four peaks of the F2F meetings are due not only to the higher number of participants, but also to the length of the conferences (2-3 days of a F2F conference against 1-2 hours of a teleconference). Moreover, considering the distribution of the authorship in the sample data set, grouped per company, it is quite evident that some companies (and some subject) like Sun (26% of total communication) or WW Grainger (11%), participated in the process much more intensively than others, like Oracle (4%) or Chevron (5%).
3) **Mail selection** - In order to have different data sources we also investigated the mailing lists coordinated by the W3C consortium and open to discussion with the public. Thus we analyzed all the emails exchanged between Feb. and Dec. 2002. Moreover we decided to group by the emails along the same time scale adopted for the meeting notes. Thus, our first analysis was to count the number of email exchanged between two direct communication events (F2F or teleconference).

As shown in Fig. 3, our preliminary evidences indicate periods of intense email exchange followed by period of relatively low interaction (min. 22, max. 182 emails, for periods of around 7 days).
Comparing the email volumes and the number of text passages in teleconferences or in F2F meetings (Fig. 4) we can have a detailed representation of the communications between the different actors along the design phase. 

A preliminary finding that emerges from this activity is that the two channels of communication (emails and teleconference/F2F) seem to be complementary. When direct communication is maximum (e.g. April 10th, June 13th, September 12th, November 14th), the email exchange drops off. Similarly, the picks in the email exchanges (7-14 March, 2-9 May, July, etc) correspond to period without any meeting or teleconference.

4) Artefact analysis - The subsequent analytical step has been the selection and the analysis of pilot episodes related to a specific part of the artefact. We selected the first five most relevant parts of the artefact in terms of volume of communications. The analysis about these components of the artefact is presented in section 5 and the design activities related to a specific part of the artefact (the “security” requirement) is described in section 5.1 based both on the F2F and teleconference meetings and on the exchanged emails.

4.2 Strategy and issues for further analysis

The author-based text labelling phase discussed above was useful to trace an overall picture of the data set. The subsequent steps would be: 1) drawing the social network; 2) exploring the relationship between the social network structure and the characteristics of the artefact. Both these steps present crucial methodological issue to be faced.

Social network analysis is often based on detecting and representing the number of direct communications (links) between actors (nodes). However, in our research domain, also the simple task of defining and representing direct communication presents intriguing problems. How do you consider a teleconference? Is it a discussion one-to-many or many-to-many? Is it enough to take part to the
meeting to be considered a node of the network (any actor is exposed to all the messages exchanged in the teleconference) or is it better to require some form of active response to draw a network node? A discussion is often triggered by the intervention of one actor, with responses by one or several actors that may not be immediately and automatically connected to the original one. Methodologies for automatic detection of the network are under study (XXX 2009, removed for anonymity); in the meanwhile, a sensible option in the exploratory phase is to manually trace down the flux of conversation for selected episodes deciding on a case-by-case basis.

Another relevant empirical issue is how to investigate the relationship between the characteristics of the artefact to the structure of the network. Which characteristics might be relevant? How to detect and measure them?

<table>
<thead>
<tr>
<th>Design activities: (3 2 1) - (3 2 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3 2 1) WSA goals-requirements</td>
</tr>
<tr>
<td>(3 2 1 1) 03 goals 12-14 inclusion</td>
</tr>
<tr>
<td>(3 2 1 2) 04 first requirments document</td>
</tr>
<tr>
<td>(3 2 1 3) 03 goal 1 ensure vs enable</td>
</tr>
<tr>
<td>(3 2 1 4) 02 initial WSA goals gathering</td>
</tr>
<tr>
<td>(3 2 1 5) 03 goals 1-14 initial list</td>
</tr>
<tr>
<td>(3 2 1 6) 06 assigning goal champions</td>
</tr>
<tr>
<td>(3 2 1 7) 06 goal 1 assure too strong</td>
</tr>
<tr>
<td>(3 2 1 8) 06 goal 5 simplicity</td>
</tr>
<tr>
<td>(3 2 1 9) 06 goal 6 security</td>
</tr>
<tr>
<td>(3 2 1 10) two req doc updates before april F2F</td>
</tr>
<tr>
<td>(3 2 1 11) 07 goals 1-3</td>
</tr>
<tr>
<td>(3 2 1 12) 07 goal 4</td>
</tr>
<tr>
<td>(3 2 1 13) 07 goal 5 simpicity</td>
</tr>
<tr>
<td>(3 2 1 14) 07 goal 6 security</td>
</tr>
<tr>
<td>(3 2 1 15) 07 goals 7-16</td>
</tr>
<tr>
<td>(3 2 1 16) 07 add new goals 17-19</td>
</tr>
<tr>
<td>(3 2 1 17) 08 goal 7</td>
</tr>
<tr>
<td>(3 2 1 18) 08 goal 3</td>
</tr>
<tr>
<td>(3 2 1 19) 08 goal 8</td>
</tr>
<tr>
<td>...</td>
</tr>
<tr>
<td>(3 2 1 176) 32 final F2F cleanup</td>
</tr>
<tr>
<td>(3 2 1 177) 32 AC001</td>
</tr>
<tr>
<td>(3 2 1 178) 32 AC2 6 11 16 17 19 AR 19 23</td>
</tr>
<tr>
<td>(3 2 1 179) 32 new choreography goal</td>
</tr>
<tr>
<td>(3 2 1 80) 32 AR33</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Fig. 5. Hierarchical classification of design activities as related to the artefact – part 1.*
Fig. 6. Hierarchical classification of design activities as related to the artefact – part 2.

Figures 5 and 6 above show a classification of some of the main activities of the design network as detected by the grounded theory analysis of the Telcon and F2F scripts.

We started our exploration by classifying the artefact (the WsA) and the related activities according to the most important architectural documents, one for each different type of deliverable to be issued by the design group.

The following parts were evidenced:

- The goals/requirements list (activities under code 3.2.1)
- The architectural document (activities under codes 3.2.3 and 3.2.9)
- The issues list (activities under code 3.2.4)
- The glossary document (activities under code 3.2.5)
- The usage scenarios (use case) document (activities under code 3.2.7).

Within each of these parts the hierarchy of sub-entities with the related activities is partially shown in the figures. The activities under code (3.2.2: Discussion on how to design WsA) are about the design method and can be referred to in general to all the parts above.

We are now investigating the differences and similarities on network structures that emerge around each object. We may expect to identify differences in the network structures according to differences in the designed outcome/artefact. At the same time we should notice a similar network structure for those objects that present similar characteristics.
5 DATA ANALYSIS AND DISCUSSION

The WsA design network under analysis in the selected sample period (between February and December 2002) consists of 66 persons belonging to 46 companies, collaborating for the design of the Web service Architecture standard in the W3C environment. According to our hypothesis we expect that (sub)networks with different structures may emerge according to the specific characteristics of the (sub)artefact under development.

Therefore, the attention is here focused on the (eventually) different sub-networks of WsA, each one working on a specific part of the artefact (i.e. architectural document, use cases, glossary, architectural requirements). We expect that similar artefacts will drive the emergence of similar network (both in its genesis and structure) and that objects that strongly differ for their characteristics will lead to different network dynamics and structures.

A preliminary analysis of the data set shows that the most active artefacts (those with a high number of communications between actors) present different network structures. We are able to report at this stage an exploratory matrix, accounting for the number of text passages attributed to each actor and relative to a specific category (Fig. 7). We select the first five artefacts for number of communications: artefact A (3.2.9.8 “ARCH document/F2F40 arch disc-revis”; 164 messages); B (3.2.152 “goals-requirements/20 DAR 6_1-13 security”; 160); C (3.2.9.9 “ARCH document/Management Task Force”; 147); D (3.2.1.43 “goals-requirements/discussing reqrmnts doc voting res”; 119); E (3.2.9.2 “ARCH document/F2F outline of arch”; 115).

<table>
<thead>
<tr>
<th>Artefact/Actor</th>
<th>3</th>
<th>5</th>
<th>8</th>
<th>11</th>
<th>13</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>21</th>
<th>22</th>
<th>30</th>
<th>32</th>
<th>33</th>
<th>41</th>
<th>43</th>
<th>51</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (3.2.9.8)</td>
<td>7</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>23</td>
<td>1</td>
<td>20</td>
<td>0</td>
<td>3</td>
<td>13</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>B (3.2.1.52)</td>
<td>45</td>
<td>14</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>11</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>C (3.2.9.9)</td>
<td>18</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>9</td>
<td>12</td>
<td>0</td>
<td>7</td>
<td>18</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>3</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>D (3.2.1.43)</td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>57</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>E (3.2.9.2)</td>
<td>27</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>17</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Fig. 7. Communication matrix of teleconference and F2F meetings. Columns report the actor codes; rows the artefact codes illustrated above in Figures 5-6. In each cell the number of text passages for each actor/artefact combination is shown.

The sub-artefact codes (A, C, E), evidenced in grey, are part of the architectural document (code 3 2 9). The codes (B and D), evidenced in orange, are part of goals/requirements list (code 3 2 1).

The matrix cells report the text passages produced by each actor in reference with each object. The cell colours are lighter for lower volumes of communication, darker for higher volumes.

Actor #16 was very active for the artefacts B (28% of all the communication about the artefact) and D (48%) but he was less active for the other artefact (only 15% of his communication has focused on artefact A, C and E). Artefacts B and D seem to be driven by the activity of few persons compared with artefacts A, C and E. In details, if we consider the amount of communication between the most active actor for every artefact we have that they represent the 26% for A; 51% for B; 27% for C; 59% for D and 40% for E. So, the communication for artefacts B and D result to be much more concentrated than for the others.
5.1 Artefact analysis. 3.2.1.52 design task: security

As a further exploratory step, in this section we focus on a specific (sub)artefact. We selected the most relevant artefact for volume of communication (3.2.1.52 - the "D-AG004 Security" requirement) and then we investigated all the direct communications (F2F, teleconference) or emails related to it. We will then try to replicate this analysis to different artefacts and than to compare these results in order to test our original hypothesis about the role of the network outcome in network genesis. Our hypothesis will be confirmed if similar artefacts present similar network structures and if different artefacts present different structures. Thus, this section will focus on D-AG004 Security artefact as a debated artefact inside the WsA distributed group.

*D-AG004 Security* is a specific part of the Web services Architecture Requirements (3.2.1.52), it "addresses the security of Web services across distributed domains and platforms” and it “enables privacy protection for the consumer of a Web service across multiple domains and services”. During the sample period there have been 164 text passages in the minutes regarding this specific topic and several threads.

*Minutes* - Analyzing in details the minutes of the (physical and virtual) meetings, it emerges that few actors lead the discussion on the security requirement. In details, actor 13 is the most active (23 txt passages) followed by actor 91 (22), 51 (21), 17 (20) and 21 (13). Thus more than the half of the communications on the security requirement (86/164) is conducted by four actors. After them, there are three actors (8, 11 and 87), each with 10 passages and others with few communications. In total, only 16 actors participate in an active way to this discussion (around 24% of the group members). Considering the organization as level of analysis, W3C is leading the discussion with around 20% of the communication, followed by Sun (around 15%) and Thompson Corporation (14%). The group of interest composed by Microsoft, IBM and BEA does not often comment on this requirement (17 passages; around 10%). The security requirement appears to be immediately interesting for the group. In the first meeting (06 February 2002) there are 10 different members pointing at the security issue as a crucial requirement that should be designed. The second meeting (14 February 2002) has already the security as an issue in the agenda (point “e” and initially AG006) and in the first WsA Requirements Document (29 April 2002) the security is already well identified as D-AG004 Security and decomposed in two specific sub-parts (D-AC006 and D-AC020); this classification is the same of the final release of the WsA Requirements (11 February 2004).

*eMails* - In order to identify the relevant emails for this specific topic, we have investigated the mailing list repository for the period considered in this analysis (2002). As result for a searching with the keyword AG004 we have identified 69 mails. However in the initial period the “security” requirement have been identified with the tag AG006. Thus we have analyzed all the emails with either the keyword AG004 or AG006 in order to identify those in which the security topic was debated. Non pertinent emails selected by automatic search were manually spotted and removed.

At a preliminary view, the tendency to centralize the discussion of this requirement among few actors seems to emerge also in the mailing list. This finding is in line with similar results in different distributed context, as the open source development teams that seem to be founded on solo-work into a shared and collaborative environment.

6 PRELIMINARY CONCLUSIONS

Summing up the contribution, we can consider this paper as a first attempt to analyze a well known phenomenon (social networks) from a different perspective. We embraced this challenging path in order to avoid a limitation of the traditional perspectives that consider the network structure as exogenous and stable. Instead, in our perspective, we consider the reciprocal influence between the artefact and the social network structure, with a phase in which the desired artefact may shape the network genesis and a phase in which the emergent network’s structure may drive the artefact design.
These two phases may be iterative along a collaborative project life. Further investigation may uncover relevant patterns and factors of the process, advancing our knowledge of design networks and contributing to new methods and perspectives in social network analysis.

In conclusion, this article is illustrative, rather than confirmatory. It is desirable to continue the study, moving forward to a deeper data analysis in which structural measures of the social network (cohesion, homogeneity, connectedness, centrality) will be conducted. Moreover, the study should also take into consideration a more detailed literature review and the possible overlapping with other research streams, as collective actions or common resources.

References


A UTILITY-BASED MODEL TO DEFINE THE OPTIMAL DATA QUALITY LEVEL IN IT SERVICE OFFERINGS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0508.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Data quality, Decision making, E-Service, Information</td>
</tr>
</tbody>
</table>
A UTILITY-BASED MODEL TO DEFINE THE OPTIMAL DATA QUALITY LEVEL IN IT SERVICE OFFERINGS

Cinzia Cappiello, Politecnico di Milano, Dipartimento di Elettronica e Informazione, Via Ponzio 34/5, 20133 Milano, Italy, cappiell@elet.polimi.it

Marco Comuzzi, City University London, Department of Computing, School of Informatics, Northampton Square, EC1V 0HB London, UK, sbbd286@soi.city.ac.uk

Abstract

In the information age, enterprises base or enrich their core business activities with the provision of informative services. For this reason, organizations are becoming increasingly aware of data quality issues, which concern the evaluation of the ability of a data collection to meet users’ needs. Data quality is a multidimensional and subjective issue, since it is defined by a variety of criteria, whose definition and evaluation is strictly dependent on the context and users involved. Thus, when considering data quality, the users’ perspective should always be considered fundamental. Authors in data quality literature agree that providers should adapt, and consequently improve, their service offerings in order to completely satisfy users’ demands. However, we argue that, in service provisioning, providers are subject to restrictions stemming, for instance, from costs and benefits assessments. Therefore, we identify the need for a conciliation of providers’ and users’ quality targets in defining the optimal data quality level of an informative service. The definition of such equilibrium is a complex issue since each type of user accessing the service may define different utilities regarding the provided information. Considering this scenario, the paper presents a utility-based model of the providers’ and customers’ interests developed on the basis of multi-class offerings. The model is exploited to analyze the optimal service offerings that allow the efficient allocation of quality improvements activities for the provider.

Keywords: Data Quality, Efficiency, Service Offerings.

1 INTRODUCTION

Enterprises increasingly focus their business strategy on information management, since an effective use of organizational data can have a considerable impact on business decisions and provide high benefits. Furthermore, information can be often exploited to offer additional and valuable services to external customers. Informative services manipulate raw data and produce information products (Ballou et al. 1998). In order to evaluate the service effectiveness and thus the ability of the provided information to meet both the organizational and users’ (customers) requirements, it is then possible to consider data quality theories. Data quality is a multi-dimensional concept that evaluates the suitability of data to the tasks for which they are required, and thus to the users that access them. Data quality can be assessed by means of different dimensions, whose definition and evaluation are strictly dependent on the context and users involved. For this reason, the users’ perspective has been always considered fundamental in data quality. Consequently, literature contributions have always focused their attention on the definition of methodologies and methods that support providers in the achievements of data quality targets that would completely meet users’ needs. Quality management mainly suggests the adoption of the Zero Defects approach that consists in setting targets to the highest quality values (English 1999). However, if the organization follows a zero defects approach in areas which do not need it, resources may be wasted. Moreover, reaching the highest quality values might lead to quality improvement that the organization may not be able to afford. In fact, it is necessary to consider that
providers have their own requirements in provisioning services and many times the complete satisfaction of users’ requirements is not convenient since costs are greater than benefits. Hence, the Zero-Defects approach to data quality management is often excessive, since it does not consider that the data quality improvement is not a trivial task and in some cases it requires very expensive projects, which are not always feasible for the service providers. Generally, we argue that it would be better to adopt an approach that fixes data quality targets on the basis of a conciliation of providers’ and users’ needs. The definition of this equilibrium is a complex issue since, for each provided information, we can have different utilities depending on the type of user that accesses it. Considering this scenario, the paper presents a utility-based model of the providers’ and customers’ interests developed on the basis of multi-class offerings. The model is exploited to analyze the optimal service offerings that allow the efficient allocation of quality improvements activities for the provider.

The paper is organized as follows. Section 2 reviews the literature on similar contributions. Section 3 presents the main useful concepts for data quality management and shows the model for the definition of the users and providers quality targets. Section 4 presents the model of the provider and users’ utility functions in a data service scenario and discusses the issue of optimal data quality level definition in informative service offerings.

2 RELATED WORK

The identification of service offerings that define the most suitable quality targets that contemporarily satisfy providers and users’ needs is a research issue that can be generally related to the identification of quality level agreements. This is a new open issue in the data quality field, as well as in the broader field of Service Oriented Computing. Here, the Service Level Agreement (SLA) is defined as a binding contract which formally specifies end-user expectations about the solution and tolerances, i.e., it is a collection of service level requirements that have been negotiated and mutually agreed upon by the information providers and the information consumers. In fact, providers define some service levels as a fixed combination of their specific capabilities on a set of quality dimensions that are also considered by the users to define their targets. If providers’ capabilities and users’ needs are not immediately compatible, a negotiation phase is required in order to find the most suitable conciliation between providers’ and users’ quality targets. In this field, there are several languages proposed for the definition and monitoring of the SLA such as WSLA (Keller and Ludwig 2002) or WS-Agreement (Ws-Agreement Framework 2003). WSLA allows providers to define quality dimensions and to describe functions to evaluate them. Furthermore, it provides monitoring of the parameters during operations and invocation of recovery actions when contract violations occur. Similarly, WS-Agreement provides constructs for advertising the capabilities of providers and creating agreements based on creational offers, and for monitoring agreement compliance at runtime. Once the service capabilities description is provided, the selection of the most suitable service is enabled by the definition of the users requirements. The SLA definition starts from provider capabilities and users’ requirements specification and defines all the condition of the service provisioning.

In the data quality field, quality requirements are focused on a set of criteria able to define the suitability of a data set for the process in which it is involved. Data quality is a multi-dimensional concept and the data quality literature provides a thorough classification of data quality dimensions, even if there are discrepancies on the definition of most dimensions due to the contextual nature of quality. The six most important classifications are presented in (Wand and Wang 1996, Wang and Strong 1996, Redman 1996, Jarke et al. 1999, Bovee et al. 2001, Naumann 2002). By analyzing these classifications, it is possible to define a basic set of data quality dimensions including accuracy, completeness, consistency, timeliness, interpretability and, accessibility, which represent the dimensions considered by the majority of the authors (Scannapieco and Catarci 2002). The assessment of these dimensions reveals the ability of a data collection to meet users’ needs.

In the literature, data quality users’ requirements have been mostly used as one of the driver for the identification of the most suitable data source (e.g., Scannapieco et al. 2004). Users’ requirements have been sometimes translated into utility functions. In (Even and Shankaranarayan 2007), utility
functions have been used by supporting multiple assessments of quality, each within a different usage context. Utility functions have been also used to alleviate the problem of data fusion in the presence of inconsistencies, for example in combining different versions of the same data (Motro et al. 2004). As already discussed in the Introduction, in the data quality field, the provider perspective has been scarcely considered. Data quality agreements issue has been only addressed in quality-constrained data provisioning (Missier and Embury 2005). Missier and Embury (2005) propose a framework for the definition of formal agreements between the provider and the customers. Focusing on the completeness dimension, they also provide an algorithm for dealing with constraints on the completeness of a query result with respect to a reference data source.

In our work, the approach can be considered innovative since providers capabilities are not fixed at priori. In fact, we primarily consider the users’ requirements and we assume that the provider capabilities are functions of the current quality level of their IT services and of the costs related to the improvement activities needed to satisfy users requirements.

3 THE SERVICE PROVISION AND DATA QUALITY REQUIREMENTS SPECIFICATION

A business process can be composed and executed by means of IT and physical services. The former are services that are responsible for data manipulation and that aim at generating and providing useful information. The latter are business services that are composed of physical activity that cannot be made automated (e.g. delivery of goods). In this paper we focus on the first type of services and we characterise them by considering functional and non-functional requirements. Since the output provided by IT services is information, the quality of such services can be mainly evaluated by considering data quality dimensions.

In our model, in the data quality assessment phase, we consider the quality dimensions introduced in the previous section and define an aggregate measure of data quality level ($q_c$) by using a weighted average, that is:

$$q_c = \sum_{i=1}^{N} w_i \cdot dq_i$$

(1)

where $w_i$ are the weights that denote the importance of the single dimension $dq_i$ for the user or the provider and $N$ is the total number of the considered criteria. In order to use this model, we make the main assumption to consider the quality dimensions independent of each other.

If the assessment results reveal that the provider sources are characterized by an insufficient data quality level, the adoption of quality improvement techniques should be considered. Improvement methods are distinguishable in data-oriented and process-oriented techniques. The former focus on error detection and correction, whereas the latter aim at identifying and correcting the activity in the process responsible for the error. Therefore, data-oriented techniques are characterized by low investment costs and short term benefits, whereas process-oriented techniques imply a very high investment cost, even though they are likely to provide long-term benefits. Process-oriented techniques are, in general, to prefer, since data-oriented techniques need to be performed periodically to obtain long-term benefits and thus the total cost will be higher than the initial investment of any process-oriented technique.

In the framework proposed in this paper, the providers should evaluate their convenience to improve the data quality level by also considering that low data quality levels raise poor quality costs, mainly due to service failures and consequent repair actions.

A fundamental hurdle is that costs and benefits of attaining a certain data quality level are difficult to estimate ex ante. We consider a distinction between non-quality and quality costs.

- **Non-quality costs:** they are costs associated with poor data quality and, consequently, with all the activities necessary to correct errors and re-execute tasks.
- **Quality costs:** they are associated to the activities and resources necessary in the improvement project.
Improvement interventions may be of variable complexity. They could regard the purchase and implementation of standard software tools (e.g., data cleaning tools), the design and the development of ad-hoc software modules or, in the most complex cases, the re-organization of the whole IT architecture (e.g. for improvement of information availability or security). Therefore, quality costs can include licence costs, hardware costs for the acquisition of new machines and human resources costs for analysis, development or implementation activities.

It is necessary, however, to consider that non-quality costs can be considered as a potential saving, and represent tangible benefits of quality improvement. The benefits of the improvement process are at least equal to the savings from non-quality costs. Additional tangible and intangible benefits can be achieved in higher-performance scenarios. It must be noted that the quality costs depend on the improvement techniques that are implemented and benefits are related to the type of services that are improved through data enhancement.

**Figure 1 - Isolated and interconnected services**

On the basis of the role played by information, it is possible to distinguish different types of IT services. In this paper, we consider a distinction between isolated and interconnected services (see Figure 1):

- **Isolated services**: these are services for which the output data are specially produced for the final user and they are not used in the information system of the organization, which supports the execution of the organization’s business processes.

- **Interconnected Services**: they are services which produce information that is also used in the organization’s daily operational activities.

Systems that provide isolated services can be compared to open loop systems in which the improvements in the quality of output data are totally dedicated to meet the user requirements. In fact, such improvements do not impact on the provider operational processes and benefits from data quality improvements will derive only by the increase of the customer satisfaction. Conversely, systems that provide interconnected services can be compared to closed loop systems in which improvements are likely to influence the organization’s internal business processes and will produce higher benefits for the provider. In fact, improvements of the output data will also impact on the correctness of operational data and thus, on the execution of all the business processes. In this case, improvements decrease the probability that services might fail as well as the poor quality costs.

Real-time data about stock quote rates, for instance, can be provided by either financial brokering institutions or merchant bankers. In the former case, we can label the service as isolated, since brokers simply collect data from different sources in order to satisfy the requirements of their customers. In the latter case, the stock quote provisioning service can be considered interconnected, since financial institutions, besides selling data to customers, are also likely to exploit the same data for their internal activities, e.g., managing customers’ investment portfolios.

Generally, we can introduce a coefficient $\alpha$ to express the degree of interconnection of a service. It defines the impact that a quality improvement intervention undertaken for a service has on the provider operational processes. If $\alpha = 0$, then the service provisioning process is not connected with the organizations business processes, whereas if $\alpha > 0$ the service provisioning process shows an increased degree of interconnection with internal business processes ($0 \leq \alpha \leq 1$).

Furthermore, it is also necessary to clarify that service providers are used to offer services along different quality configurations (i.e., service levels) in order to satisfy different user requirements. As
an example, we can consider the visualization of stock values in a trading activity, a classical financial service for which timeliness is a critical dimension. For the sake of simplicity, let us consider two classes of users: ordinary retail customers and traders. The two classes of users have different requirements on stock information. The ordinary retail users can tolerate a lower quality than the trader, as they usually access the service with a lower frequency. Moreover, errors on stock values have a considerably lower impact, since operations performed by ordinary retail users are usually less risky and involve lower amounts of money. On the other hand, the trader needs accurate and updated information and, therefore, requires high-quality information. The effects of an inaccurate or delayed value can be disastrous, since the trader is often involved in very risky operations involving stocks from a large portfolio of customers. Therefore, banks are likely to provide two different quality profiles for the stock quotes information service. The first profile is characterized by an acceptable value of timeliness while the other one specifies real-time information provisioning and is also associated with a higher cost.

In respect of this example, we consider an information service associated to several classes of users. Users belonging to a class have the same requirements on the quality of the data provided by the service S. In details, our model of service offerings considers a service S and assumes that a user (or customer) \( u \) is assigned to one of the \( K \) user classes \( UC_k \), where \( k = 1, ..., K \). Each class contains users with similar characteristics. The number of users in a given class \( UC_k \) is indicated as \( M_k \). Users belonging to the same class are associated with the same quality requirements for the service S. For each user class \( UC_k \), we define the data quality level \( q_{ck} \) defined in the service offerings \( QC(S) \) for service S. Note that each \( q_{ck} \) is calculated as a weighted average of the requirements specified for the different quality dimensions by using the formula shown in Eq. 1. Hence, the service offering \( QC(S) \) for service S is defined as a set of increasing data quality levels associated to \( K \) classes, that is:

\[
QC(S) = \{q_{c1}, ..., q_{cK}\}.
\]  

From the provider perspective, the aim is to define a service offering \( QC(S) \) that satisfies some optimization criteria. A first criterion can be of defining service offerings on the basis of the fulfillment of the user requirements. Usually, such criteria tend to minimize the specification of subjective quality levels, since service offerings are developed to best fit user requirements. In the next section we introduce a utility model for describing the provider and the customers’ interest, and define a criterion for defining service offerings which jointly considers the interests of both the provider and the customers.

### 4 A MODEL FOR SERVICE OFFERINGS EFFICIENCY ASSESSMENT

In order to define an efficient way for the provider to define service offerings and to decide the quality improvement actions to be performed on data, we first need to introduce a model which defines the provider and the customers’ utility functions in our informative service scenario. The model relies on the definition of utility functions for both the data provider and customers. In our model, we adopt quasi-linear utility functions (Jackson 2003). Quasi-linear utility functions represent an efficient and compact modeling tool for negotiations and bargaining problems in which it is easy to isolate, for every participant, value and payment terms. We argue that the case of data quality and, specifically, information service offerings falls within such category. Sources of benefits and costs related to data service offerings for providers and customers, in fact, have already been analyzed by a large body of academic literature (Batini et al. 2006, Eppler and Helfert 2004, English 1999, Loshin 2001).

Quasi-linear utility functions are such that the utility value for an agent on a given contract \( X \) is defined by two terms, i.e., a value and a payment term. The value term determines the value obtained by an agent from the contract, whereas the payment term refers to the amount of money that an agent is going to receive or pay for the contract. Value and payment terms can be either positive or negative.
For the provider, the payment term is positive and value term is negative, because the provider receives money from customers, but, at the same time, it sustains a cost for providing the negotiated contract, therefore losing utility. Conversely, the payment term is negative for customers, whereas the value term is positive, because the customers pay money for a contract and, at the same time, have a positive evaluation of the contract negotiated with the provider.

We first introduce the definition of quasi-linear utility functions for data providers and customers in the multi-class data service scenario introduced in the previous section. Then, we show how the utility model can be exploited to provide a criterion for the provider to define the optimal service offerings and, consequently, clarify which quality improvements need to be performed. The optimal service offering obtained through our model is then compared against the findings of the zero-defect approach for data quality management, which, generally, implies the complete fulfillment of the customers’ quality requirements.

Generally, a quasi-linear utility function defined for an agent $P$ behaving in a service provider’s perspective is defined as:

$$U_p = \text{Price}(X) - \text{Cost}(X)$$  \hspace{1cm} (3)

where $\text{Price}(X)$ is the price, that is, the amount of money obtained by $P$ for providing the generic contract $X$ (payment term), whereas $\text{Cost}(X)$ represents the cost sustained by $P$ to provide the contract $X$ (negative value term).

Similarly, for the generic agent $C$ behaving as a service customer, the utility of a contract $U_C(X)$ assumes the following form:

$$U_C = \text{Value}(X) - \text{Price}(X)$$  \hspace{1cm} (4)

where $\text{Value}(X)$ is the value generated by the contract $X$ to the customer, whereas $\text{Price}(X)$ is, as in $U_p(X)$, the price that the customer has to pay for receiving the contract $X$.

According to the model presented in Section 3, in our multi-class data service scenario, the contract $X$ assumes the form of the service offering $QC$:

$$X = QC = (qc_1, \ldots, qc_K)$$  \hspace{1cm} (5)

where $qc_k$ is the data quality level provided to customers in class $k$, with $k = 1, \ldots, K$.

In the following, we will consider two agents, i.e. the provider $P$, providing the service $S$, and the collection of customers $C$ of service $S$.

The total amount of money $\text{Price}(QC)$ received by the data provider $P$ for the provisioning of a given service offering $QC$ is given by the sum of the money received from customers in each service class defined in the service offering, that is:

$$P(QC) = \sum_{k=1}^{K} p(qc_k) \cdot M_k ,$$  \hspace{1cm} (6)

where $p(qc_k)$ is the price of data provided for users in class $k$, while $M_k$ is the number of users that belong to class $k$.

The term $\text{Cost}(QC)$ represents the cost sustained by the provider to provide a service offerings $QC$, and can be expressed as:

$$\text{Cost}(QC) = C_p(qc_K) - B_p(qc_K)$$  \hspace{1cm} (7)

where $C_p(qc_K)$ is the cost actually sustained by the provider to provide the maximum quality level $qc_K$ to its customers, whereas $B_p(qc_K)$ is a quantification of the benefits introduced in the provider’s internal business processes by attaining a maximum data quality level $qc_K$. $C_p(qc_K)$, is a function of only the maximum data quality level $qc_K$ in $QC(S)$ since we argue that, once the provider commits to
provide $q_{c_k}$ to its customers, the marginal cost of providing the data service with a lower quality level $q_{c_k}$, with $1 < k < K-1$, is negligible. Similarly, we also argue that it is rational for the provider to use, in its internal business processes, data at the maximum quality level $q_{c_K}$. Therefore, $B_p(q_{c_K})$ is a function of only the maximum data quality level $q_{c_k}$.

The value term $Value(QC)$ in $U_c(QC)$ can be expressed as the sum of values $v_i(q_{c_k})$ generated for each class of customers $k$. Therefore:

$$Value(X) = Value(QC) = \sum_{k=1}^{K} v_k(q_{c_k})$$  \hspace{1cm} (8)

Our objective is to study the optimal service offering $QC^*=(q_{c_1}^*,...,q_{c_K}^*)$ that maximises the sum of the utility for the provider and the service customers. Moreover, we demonstrate that such $QC^*$ differs from the optimal $QC_z$ imposed by the zero-defect approach to data quality management, which implies the service offering to fully satisfy the requirements of the customers. More specifically, in our model, we argue that the full satisfaction of customers’ requirements occurs when a service offering maximises the value $v_i(q_{c_k})$ for each class of customers. Therefore, in the zero-defect approach, the optimal service offering $QC_z$ can be defined as follows:

$$QC_z = (q_{c_1}^z,...,q_{c_K}^z)$$, where $q_{c_i}^z = \text{arg max}_{q_{c_i}} v_i(q_{c_i})$ for $i=1,...,K$.

The sum of the utility of customers and suppliers can be expressed:

$$U_p(QC) + U_c(QC) = Value(QC) - Cost(QC) = \sum_{k=1}^{K} v_k(q_{c_k}) - C_p(q_{c_k}) + B_p(q_{c_K})$$ \hspace{1cm} (9)

Since $U_p + U_c$, as a function of $QC$, is separable in the variables $q_{c_k}$, with $k=1,...,K$, the conditions under which $U_p(QC) + U_c(QC)$ is maximised can be expressed as:

$$\frac{\partial(U_p + U_c)}{\partial q_{c_k}} \bigg|_{k=1...K} = 0$$  \hspace{1cm} (10)

Since both $C_p$ and $B_p$ are a function of only $q_{c_K}$, for the first $K-1$ equations implied by Eq. 10, we can write the following:

$$\frac{\partial(U_p + U_c)}{\partial q_{c_k}} \bigg|_{k=1...K-1} = \frac{dv_k(q_{c_k})}{dq_{c_k}} = 0$$  \hspace{1cm} (11)

While the $K$th equation becomes:

$$\frac{\partial(U_p + U_c)}{\partial q_{c_k}} \bigg|_{k=K} = \frac{dv_k(q_{c_K})}{dq_{c_K}} - \frac{dC_p(q_{c_K})}{dq_{c_K}} + \frac{dB_p(q_{c_K})}{dq_{c_K}} = 0.$$  \hspace{1cm} (12)

In order to go into detail in the analysis of the optimal service offering $QC^*$, we need to characterise the functions $v_i(q_{c_k})$, $C_p(q_{c_K})$, and $B_p(q_{c_K})$ in our information service offering settings.

Considering past contributions (Batini et al. 2008, Eppler and Helfert 2004, English 1999, Loshin 2001) on data quality costs, it is possible to consider the following form of costs $C_p(q_{c_K})$ sustained by a provider:

$$C_p(q_{c_K}) = F + P \cdot q_{c_K} + I \cdot \exp(q_{c_K})$$  \hspace{1cm} (13)
where the $F$, $P$, and $I$ are coefficients that reflect the complexity of a generic data quality project. The coefficient $F$ is related to the fixed costs of the data quality projects required for achieving $qc_K$, such as licence costs or hardware costs for the software and IT infrastructure required by the project. The coefficient $P$ relates to the project development part, in which we have variable costs associated to the analysis and implementation activities that are evaluated by considering their duration and involved human resources. The coefficient $I$ is related to the improvement of some data quality dimensions (e.g., availability, accessibility, security). Such improvements usually require considerable changes in the whole IT architecture and therefore the costs associated to them grows exponentially with the level of data quality $qc_K$ that needs to be achieved.

We make the assumption that the benefits obtained by the service provider to attain a maximum data quality level $qc_K$ are a fraction $\alpha$, with $0 < \alpha < 1$ of the sustained costs. Such a fraction $\alpha$ is determined by the degree of interconnection of the provider’s internal business processes defined in Section 3. Specifically, the higher the degree of interconnection, the higher the benefits that the provider can obtain on its internal business processes. Therefore:

$$B_p(qc_K) = \alpha \cdot C_p(qc_K) = \alpha F + \alpha P \cdot qc_K + \alpha I \cdot \exp(qc_K)$$

We propose two different types of value functions $v_k = v_k(qc_k)$ for the set of customers in the generic $k$-th class.

**Type 1: Gaussian customers value functions.**

In this case (see Figure 2), the optimal data quality value $qc_k^z$ for customers in the $k$th class under the zero-defect principle is $qc_k^z = \mu_k$, that is, the one that maximises the customers’ value.

$$v_k(qc_k) = \frac{1}{\sqrt{2\pi} \sigma_k} e^{\frac{(qc_k - \mu_k)^2}{\sigma_k^2}}$$

**Figure 2 - Gaussian utility function for customers in class UC_k**

**Type 2: Monotonic increasing customers value functions.**

The second type of value functions considers monotonic increasing value that saturates at a certain value $qc^z$ (specifically, we use a sigmoid function to express this second type of value function, see Figure 3). Such level $qc$ is the one identified by the zero-defect approach as the optimal data quality level for customers in the $k$-th class, since it maximises the customers’ value.

$$v_k(qc_k) = \frac{K}{1 + e^{-qc_k}} - 0.5$$

**Figure 3 - Increasing Monotonic utility function for customers in class UC_k**
The Gaussian increasing value function is suitable to model cases in which customers cannot accept higher or lower quality values than the requested ones. Considering the example discussed in Section 3, normal customers can accept out-of-date stock values since they do not have to use these data in critical processes, but they cannot accept higher quality level, since this would require an additional cost that they should pay for information that they do not actually need. The monotonic increasing value function is suitable to model cases in which customer define their acceptable quality level as the minimum acceptable value. For example, traders obviously cannot accept out-of-date stock values and, therefore, they are likely to fix a minimum quality requirement. However, traders are also likely to be equally satisfied with a quality level that exceeds their minimum requirements, since they may need to deal with unpredictable and critical situations which could be benefit from higher quality of data.

In respect of Eq. 11, the optimal data quality level in our model \( q_{c_k}^* \), with \( k=1,\ldots,K-1 \), coincides with the optimal data quality model identified by the zero defect approach, that is, \( q_{c_k}^* = q_{c_k}^{\ast} \).

In fact, for both Gaussian and monotonic increasing value functions, the following condition holds:

\[
q_{c_k}^* = \left\{ q_{c_k} : \frac{d\nu_k(q_{c_k})}{dq_{c_k}} = 0 \right\} = q_{c_k}^{\ast} \text{ for } k=1,\ldots,K-1.
\]

In other words, both our model and the zero-defect approach to data quality management imply that customers who do not ask for the maximum level of data quality should be provided with a level of data quality that fully satisfy their requirements, that is, that maximises their value.

For the maximum level of quality provided to customers, the findings of our model differ from the corresponding findings obtained with the zero-defect approach, i.e. \( q_{c_K}^* \neq q_{c_K}^{\ast} \).

In our model, the data quality level \( q_{c_K} \) is determined by solving Eq. 12, which can be rewritten as:

\[
\frac{\partial\nu(q_{c_K})}{\partial q_{c_K}} - P(1-\alpha) - I \exp(q_{c_K}) + \alpha \cdot \exp(q_{c_K}) = 0;
\]

which then leads to the following equation:

\[
\frac{\partial\nu(q_{c_K})}{\partial q_{c_K}} = P(1-\alpha) + I \cdot (1-\alpha) \cdot \exp(q_{c_K}).
\]

A first consideration that must be made is that the maximum level of data quality \( q_{c_K} \) does not depend on the fixed costs \( F \) of the data quality project. This is consistent with the fact that, by definition, fixed costs must be sustained for any data quality project, the maximum level of attained quality \( q_{c_K} \) notwithstanding.

Eq. 16 can be solved graphically in two cases C1 and C2 that consider, respectively, Gaussian and monotonic increasing functions for modelling the customers’ value. A graphical representation of the solutions of Eq. 16 in case C1 and C2 is given in Figure 4 and Figure 5, respectively. Please note that in the graphical solution, the function \( f(x) \) represents the derivative of the customers’ utility function on \( v(q_{c_K}) \) (i.e. left argument in Eq. 16).

The graphical representation of the solution to Eq. 16 in case C1 is reported in Figure 4.
The first conclusion that can be drawn is that, generally, $q_{c_k}^* < q_{c_k}^\circ$, that is, our model implies a maximum data quality level that is lower than the optimal level predicted adopting the zero-defect approach. This is consistent with the fact that, while the zero-defect approach only considers the maximisation of customer’s requirements, our model considers a trade-off between the costs sustained by the provider to provide a certain maximum data quality level (and the benefits induced by these) and the value produced for the customers requiring maximum data quality.

Second, the maximum data quality level $q_{c_k}^*$ decreases as the costs (both $P$ and $I$) for providing a certain data quality level increase (see $q_{c_k}^*$ low costs and $q_{c_k}^*$ high costs in Figure 4). Moreover, $q_{c_k}^*$ decreases as $\alpha$ decreases, since both terms $P(1-\alpha)$ and $I(1-\alpha)$ increase. A lower value of $\alpha$ implies a lower degree of interconnection, as defined in Section 3, of the provider’s business processes. In our modelling of the informative service provisioning scenario, a lower degree of interconnection implies a decrease of the benefits induced on internal business processes by the attainment of a certain level of data quality. Therefore, our model can be used to identify which is the impact of the degree of interconnection on the decrease of the maximum data quality level $q_{c_k}^*$ that can be provided to customers.

Third, it has to be noticed that the more spike shaped the Gaussian function representing the value created for customers, the closer the quality level $q_{c_k}^*$ to the zero-defect approach level $q_{c_k}^\circ$ (see $q_{c_k}^*$ spike and $q_{c_k}^*$ low costs in Figure 4). This can be interpreted by stressing that when customers’ interested are clearly stated, i.e., the function representing the value created for customers in the Kth class is spiked around $\mu_K$, then it will be easier for the provider to identify the optimal trade-off between the costs sustained for providing quality of data and the value generated for customers (and this will be closer to the optimal satisfaction of customers’ requirements hypothesized by the zero defect approach to data quality management).

Finally, in case C1, it has to be noticed that when costs are too high, Eq. 16 cannot be solved, i.e., there is no intersection between the two functions (dotted and straight lines in Figure 4, respectively)
represented in the Figure 4. In other words, it is impossible for the provider to find the trade-off between the cost of quality of data and the value generated for customers.

\[
g(x) = P \cdot (1 - \alpha) + I \cdot (1 - \alpha) \exp(qc_K)
\]

Figure 5 - Determining the optimal data quality level (Case C2, Monotonic increasing customer utility)

The graphical representation of Eq. 16 in case C2 is shown in Figure 5. The major findings already discussed for case C1 hold also in this case. More in detail, the maximum data quality level \( qc_K^* \) predicted by our model will always be lower than the value \( qc_K^z \) implied by the zero-defect approach and \( qc_K^* \) decreases as (i) the costs sustained for providing data with a certain level of quality increases and (ii) the provider’s business process are less interconnected (i.e., \( \alpha \) decreases).

Similarly to what happens in C1, a solution to Eq. 16 may not be found if costs are very high also in C2. Moreover, it has also to be noticed that the value \( qc_K^* \) drastically decreases as the cost terms \( P(1-a) \) and \( I(1-a) \) increase.

5 CONCLUDING REMARKS AND FUTURE WORK

The paper has presented a novel approach for defining optimal service offerings for information services. In particular, our model defines the optimal service offering as the one that maximises the sum of the service provider and customers’ utility functions. The optimal service offering obtained with our model differs by the one defined by the zero-defect approach in the definition of the maximum quality level. In particular, our model argues for lower maximum quality levels, in order to keep into consideration the trade-off between the costs sustained by the provider for improving the quality of data, the value created by the service offering for customers, and the benefits obtained by the provider on its internal business processes from the improvement in the quality of data.

The limitations of our model imply the need for future work on the model development. First, the model relies on the ability of the service provider to estimate the utility functions of the classes of users for the provided service. Understanding preferences of users requires the development of user profiling and clustering techniques, which should be further investigated. Second, our model defines
optimality of service offerings in terms of the maximisation of the sum of the provider and customers utilities. Optimality may be defined according to other metrics involving utilities of the involved actors. In particular, we want to investigate the notion of equilibrium of service offerings, i.e., studying optimal service offerings that define an equilibrium among the service provider and the customer in the utility space.

References


## UNPACKING THE ERP INVESTMENT DECISION: AN EMPIRICAL ASSESSMENT OF THE BENEFITS AND RISKS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0699.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Adoption, Enterprise resource planning (ERP) (packaged systems), Decision making, Consumer behavior / choice / demand / empowerment / reviews / consumerism</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
UNPACKING THE ERP INVESTMENT DECISION: AN EMPIRICAL ASSESSMENT OF THE BENEFITS AND RISKS

Keating, Byron, University of Wollongong, Northfields Avenue, Wollongong 2522, Australia, bkeating@uow.edu.au

Coltman, Tim, University of Wollongong, Northfields Avenue, Wollongong 2522, Australia, tcoltman@uow.edu.au

Michael, Katina, University of Wollongong, Northfields Avenue, Wollongong 2522, Australia, katina@uow.edu.au

Baker, Valerie, University of Wollongong, Northfields Avenue, Wollongong 2522, Australia, vbaker@uow.edu.au

Abstract

Most leading organizations, in all sectors of industry, commerce and government are dependent upon ERP for their organizational survival. Yet despite the importance of the decision to adopt ERP and its impact on the entire firm’s performance the IT literature has been in the large part silent on the nature of the ERP investment decision. This study is the first of its kind to determine the preference structure of senior managers around the organizational benefits and risks of adopting ERP. We present the results which provide interesting insights into how managers’ perceive the benefit and risk factors salient to the organization’s adoption decision. In line with prior research we found that improved productivity, and information and planning are important drivers of the ERP adoption decision. Moreover our findings reveal that the benefits of ERP are weighted almost twice as important as the risks when making an ERP investment decision. However when it comes to risk, interestingly managers consider issues such as top management commitment and vendor support as more important than financial risks.

Keywords: ERP, Adoption, Decision-making, Discrete Choice Analysis.
1 INTRODUCTION

Studies in the management and IT literatures that have focused on the ERP adoption decision are virtually silent on how managers make their ERP investment decisions, and the strategy making processes that support such decisions. This is despite the fact that ERP investments are generally considered to be a high cost and high risk investment for most firms. Over the past decade, companies have spent over US$300 billion dollars on ERP investments (Carlino and Nelson 2000; Shepherd and Klein 2006), and the failure rate outweighs the success rate (Hong and Kim 2002).

According to a review of published ERP research between 2001-2005, 47 percent of the existing research has focused on the implementation phase (Esteves and Bohorquez, 2007). The critical acquisition phase (or adoption phase) was the second lowest investigated—the lowest being the retirement phase of the ERP lifecycle. Esteves and Bohorquez logically argue that the limited number of studies attempting to investigate how adoption decisions are made in an ERP context is a real problem that needs to be addressed. They agree that the adoption stage is critical because as the stage preceding the implementation phase, it presents the opportunity for both researchers and practitioners to examine the dimensions and implications (benefits, risks, challenges, costs) of buying and implementing ERP software, prior to the commitment of formidable amounts of money, time and resources.

In terms of the strategic decision making process, scholarly effort to measure the process aspects of IT strategy frequently relates to the implementation of strategy or the deployment of resources and capabilities. This orientation tends to emphasize the “doing” rather than the “deciding” aspect of processes. But as Helfat et al. (2008) demonstrate, the processes for making decisions prior to taking action also matter, at least as much as deployment. This point is central to our research approach.

If strategies are made by patterns of decision making as Mintzberg (1973; 1985) suggests then studying the decisions made by senior managers would help to understanding the strategic processes of a firm. This is an idea supported by the strategic process literature which is focused on how firm capabilities can lead to improved organizational performance and competitive advantage. Strategic decisions are viewed as fundamental to firm performance because they can affect the future of the firm through actions taken, resources committed and precedents set (Mintzberg et al. 1976). Priem and Rosenstein (2000) add that “understanding the judgments of strategic leaders is essential to determining (1) how mental processes are manifest in the strategies they develop, and (2) how these processes and strategies affect firm performance” (p2).

Research in strategic management indicates that managers use intuition for key decisions, such as large capital investments (Dane and Pratt 2007), relying on what Hammond (Hammond 1974) describes as the cognitive process of last resort—human judgment. In complex real-world situations, the decision maker often has to rely on something other than facts and a full information when making decisions and resolving problems. Moreover, it is assumed that as the leader of the firm, it is the senior manager’s role to make the right judgment.

When it comes to ERP investments, the mixed results reported in industry research highlights the far reaching impact of poor strategic decision making processes (Shepherd and Klein 2006). For instance, in a survey of 232 managers conducted by Robbins-Gioia (2002), 51% of respondents viewed their ERP implementation as unsuccessful. One of the major problems identified was that organizations lacked a rigorous ERP evaluation and procurement process, which in turn, made it difficult to formulate effective investment decisions and avoid a failed ERP investment (Shang and Seddon 2002).

The purpose of this research is to open up the “black box” of managerial decision-making around ERP by measuring the relative importance of the factors that contribute to the decision to invest in an ERP system. To achieve this aim, we will utilize a novel method that will expose the organizational attributes considered to be most important and least important to the investment decision. We will use a utility based approach based on maximum difference scaling or best-worst experimentation. This
method has been successfully applied to many different contexts in order to identify the efficacy of managerial decision making (Buckley et al. 2007), and to provide insights into the preference structures for products and services (Coltman et al. 2007). To address the research issues and provide the focus for this paper we pose two questions:

1. What is the relative importance of the factors that influence the decision to adopt ERP?

2. How do key decision makers’ trade-off between the perceived benefits and risks when making an ERP investment decision?

This research draws on two streams of IT literature to determine the factors considered to be most relevant to the ERP adoption decision; (1) the research on the impact of IT, and (2) the specific literature on the organizational impact of ERP adoption and implementation. Whilst a review of the literature provides guidance as to those factors that are likely to drive an ERP adoption decision, there is little evidence regarding the relative importance of these factors to decision makers. This observation was also made by Keil and Tiwana (2006), who suggest that caution should be taken when attempting to hypothesize \textit{apriori} about the factors that have the greatest influence on the ERP investment decision, or the relative weighting that should be given to such factors by a manager charged with responsibility for making ERP adoption decisions.

The lack of past research in this area suggests that this issue may not be a theoretical problem, but rather, an empirical question that should be revealed through appropriate investigation. Furthermore, the IT and ERP impact literature acknowledges that ERP can both positively and negatively impact on organization performance. Yet, surprisingly, the literature has also not considered how both benefit and risk factors influence the decision to adopt ERP. In this study we examine how senior managers value the positive and negative aspects of an ERP investment decision.

2 THEORETICAL BACKGROUND

Studies in managerial cognition demonstrate that executives apply their own mental models to simplify the complex strategic problems they often need to solve (Porac et al. 1989). They use these models as templates to explain and interpret information relevant to decisions they need to make (March and Simon 1958; Walsh 1995). In other words, it is evident in the decision making process that decision outcomes are affected by the way a decision problem is framed. The presentation of a problem or how it is perceived will affect the choices made by the decision maker. This has a dramatic implication for executives, because how an issue is structured can directly impact the investment outcome. Whilst this highlights that managers may be biased in their decision making, using their own judgment to resolve a decision problem, it does not explain the process by which managers filter information in the decision-making process.

\textit{Cognitive categorization theory} argues that executives, during the awareness and comprehension (learning) stages of decision-making, will group information into categories in order to deal with the abundance of information available and to help them communicate with others about ambiguous strategic issues (Dutton and Jackson, 1987). Executives will then form utility preferences that ultimately form the basis of their strategic choices, in this case the choice to adopt an ERP system.

The filtering of information, which is part of the perceptual process, means that executives can have different perceptions about a problem and form different mental models around a particular decision. Starbuck and Milliken (1988) have identified two types of filtering that may occur: noticing and sense-making. As Tyler and Steensma (1998) suggest these perceptual differences occur because managers \textit{notice} different stimuli and attribute differing meanings to that information, in turn, this will lead to executive’s having different mental models around the decision to adopt enterprise systems.

The extant strategic management literature also suggests that threats and opportunities are relevant categories that are often used in strategic planning and environmental scanning activities (Christensen
et.al, 1982). Studies in the strategic decision making field have found that the categories of opportunity and threat are relevant and consequential for decision processes (Mintzberg et al. 1976). The literature on managerial decision making identifies opportunities as a “positive situation in which gain is likely” (Dutton and Jackson 1987). Alternatively, threats are seen as “a negative situation in which loss is likely” (Dutton and Jackson, 1987, p80). Following the work of Tyler and Steensma (1998) we can derive that the organizational benefits achievable through ERP adoption present as opportunities and the potential adoption risks are categorized as threats.

2.1 Benefits and Risks to ERP Adoption.

In an overview of the literature on the organizational benefits of IT, Mirani and Lederer (1998) begin to operationalize and capture the distinctions between benefits. They make an important point regarding the benefits of IT in organizations, which is that any instrument applied to capture organizational level benefits of IT must be tailored to the project being assessed. This fits with Pettigrew’s (2003) assertion that strategy research needs to address both the content and context of the phenomena under study.

The framework used for this study to assess the organizational benefits of ERP adoption is derived from the extant literature (Weill and Olson 1989; Sethi and Carraher 1993; Shang and Seddon 2002). The following dimensions, as presented in Table 1, are deemed as important antecedents to the ERP adoption decision.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Productivity and operations</th>
<th>Competitive advantage</th>
<th>Information and planning</th>
<th>IT Infrastructure</th>
<th>Compliance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Bakos and Treacy 1986)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Broadbent and Weill 1993)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(DeLone and McLean 2004)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>(Gattiker and Goodhue 2004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Hitt et al. 2002)</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Mahmood and Soon 1991; Sethi and Carraher 1993)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Mirani and Lederer 1998)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>(Porter 1985; Miller 1988)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Shang and Seddon 2002)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>(Tallon et al. 2000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Key organizational benefit factors

In light of the problems organizations have had with ERP initiatives, it would seem reasonable to assume that risk factors would generate significant research attention. However, the risks around enterprise system investment are not well developed in the literature with most of the scholarly attention being directed towards the benefits. Esteves and Bohorquez (2007) concur and state that most of the studies that do focus on risk are case-based, exploratory studies and lack a strong theoretical basis.

The risk dimensions used in this study represent a synthesis of the major risk factors from the IT and ERP impact literature. This also includes work that has focused on critical success factors, because they are dimensions that if ignored can be detrimental to the success of the investment. These dimensions, along with the associated research are presented in Table 2.

Whilst prior research outlines the importance of both benefit and risks in terms of their impact on an organization there is a lack of research that considers exactly how these factors interact to influence the decision to adopt an ERP system. For example no research could be found to determine whether
cost is more important to a manager deciding to invest in ERP than vendor support or whether
productivity and operations is more important to a manager deciding to invest in ERP than
management commitment.

<table>
<thead>
<tr>
<th>Internal Productivity</th>
<th>Cost</th>
<th>Vendor support</th>
<th>Management commitment</th>
<th>Human resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Eisenhardt and Zbaracki 1992; Besson and Rowe 2001)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Chau 1995)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Gupta 2000)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Somers and Nelson 2001)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Themistocleous and Irani 2001)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Stratman and Roth 2002)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Umble et al. 2003)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Beard and Sumner 2004)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Gattiker and Goodhue 2004)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Gargeya and Brady 2005)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Keil and Tiwana 2006)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 2. Key organizational risk factors

3 METHODOLOGY

An effective method for evaluating the issues impacting on the adoption of a new information system
(such as those related to an ERP investment decision) is to model preferences as a response to
experimentally designed decision profiles. This approach, commonly known as probabilistic discrete
choice analysis (DCA), has been used to model choice preferences of decision makers in a variety of
organizational areas spanning marketing, operations management, transportation and economics (e.g.,
Verma et al. 2006).

The statistical model underpinning DCA draws on Thurstone’s (1927) original propositions in
Random Utility Theory to provide a well-tested theory of human decision making that has been
generalized by McFadden (1974). This theory allows scholars to conceptualize individual choice as a
process of decision rules (Louviere et al. 2000). When selecting any product, service, or combination
of both, a customer will consciously or unconsciously compare alternatives and make a decision that
involves tradeoffs of the components of those choices. The result of this process is a choice outcome
that can be decomposed based on the options available within some underlying experimental design.

3.1 Best-worst Scaling

There are a number of different DCA methods that allow a researcher to elicit stated preferences that
can be used as a basis of understanding and predicting actual behavior in the marketplace. One
relatively simple method particularly useful in narrowing down and getting a quick snapshot of
preferences is best-worst scaling. The formal statistical proofs and the measurement properties for
best-worst scaling can be found in Marley and Louviere (2005). Fundamentally, best-worst scaling is
an ordering task that requires respondents to make a selection from a group of items and choose the
‘best’ (most preferred) and ‘worst’ (least preferred) items in a series of blocks of N>2 items. The items
could be attributes of a product, factors influencing a decision, or bundles of services and products.
The approach is particularly effective in creating a preference ordering for the items when the number
of items is large, as individuals are better able to determine which 2 of group of items are ‘best’ and
‘worst’ than they are the specific ordering of 1, 2, …, 12, 13. Best-worst scaling has the added benefit
that it is quick and simple to execute, provides results that are empirically consistent with more complex ordering tasks and theoretically in line with the precepts of random utility theory.

The cognitive process undertaken in the selection of the best-worst or least-most important items is statistically equivalent to:

- Identifying every possible pair of items available;
- Calculating the difference in utility between the two items in every pair; and
- Choosing the pair that maximizes the difference in utility between them.

Thus, the pair of items chosen maximizes the difference in the marginal utilities on offer between each of the various items in each block of items presented to the decision maker. Empirically, the distance between items is modeled as a difference where the relative ordering of the items is proportional to the number of times it is mentioned best less the number of times it is mentioned worst (Szeinbach et al. 1999).

In this study, we use best-worst scaling to determine the relative importance of the factors influencing the ERP adoption decision. This allows us to reduce a list of factors associated with the ERP adoption decision down to a manageable number of important components that can be scrutinized in more detail. In particular, we can use the resulting scale to make direct comparisons based on the preference for one factor versus any other factor in the original list.

### 3.2 Operational measures and survey construction

Preliminary research captured a wide range of factors that are important to the decision to adopt ERP. The factors selected in our study were based on an extensive review of the academic literature and two rounds of qualitative fieldwork. The literature review resulted in the identification of 10 factors that influence the ERP investment decision (see Tables 1 and 2). This list was then validated with a sample of IT managers using semi-structured interviews based on the laddered technique (see Reynolds and Gutman, 1988). This qualitative technique used means-end theory and probing to reveal a salient list of factors that either positively or negatively influence the ERP investment decision. The findings of these interviews helped us to recognize that business growth was an omitted driver of ERP investment, and that the cost factor needed to be separated to recognize the impact of actual versus opportunity costs, where the actual costs were further dissected to reflect differences between the initial investment and ongoing running costs. This increased our original list to 13 (see Appendix 1). While we acknowledge that this list is not exhaustive, we do believe the rigour of the above process does ensure that the list is representative.

As a final stage, and in line with the recommendations of Rossiter (2002), a second round of interviews was conducted with a sample of academic experts to validate the definitions and classifications of the 13 factors. The goal of this second qualitative phase was to ensure that each factor was concrete-singular. That is, we wanted to make sure that each factor had a common meaning across the respondent group. To do this, each expert provided qualitative feedback on the focus and purpose of the factor definitions. The construct definitions and labels were subsequently revised until all experts agreed on the definitions and the classification of the factor as either positively or negatively influencing the ERP investment decision.

Respondents were required to examine 13 sets of four factors. Within each set they were asked to indicate which factor they considered to be the most important, the next most important or the least important when investing in ERP (see Table 3). A balanced and incomplete block design was used to determine which factors appeared in which set, and to ensure that each factor appeared an equal number of times and at least once with all other factors (Street and Burgess 2004). The experimental designs that support this analysis also mean that we can obtain more data from each respondent. This increases the effective sample size, and allows us to obtain reliable estimates of demand preferences with smaller samples. This is a key advantage of DCA methods that is derived from assumptions regarding the independence of individual choices and the distribution and variance of measurement errors (for a more detailed explanation, see Louviere et al. 2000).
Which factor matters
MOST?

Set of factors for you to consider…

Which factor matters
LEAST?

- Low maintenance cost
- Improved IT infrastructure
- Adequate human resources
- Better compliance

Table 3. Example best-worst question.

4 ANALYSIS AND RESULTS

Fifty-seven middle-to-senior managers completed the best-worst experiment. The distribution of respondents by industry and occupation is shown in Table 4. In all cases the respondents were key decision makers that were involved in the evaluation of an ERP investment. Two industries dominated our sample, manufacturing and services, with the respondent roles including directors (CEO, GM) and IT executives (CIO, CTO and IS managers).

<table>
<thead>
<tr>
<th>Respondent Industry</th>
<th>%</th>
<th>Respondent Occupation</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>32</td>
<td>Director</td>
<td>32</td>
</tr>
<tr>
<td>Service</td>
<td>28</td>
<td>IT Executive</td>
<td>24</td>
</tr>
<tr>
<td>Public Administration</td>
<td>9</td>
<td>Finance Executive</td>
<td>12</td>
</tr>
<tr>
<td>Construction</td>
<td>9</td>
<td>Vice President</td>
<td>9</td>
</tr>
<tr>
<td>Wholesale</td>
<td>5</td>
<td>Sales</td>
<td>3</td>
</tr>
<tr>
<td>Mining</td>
<td>3</td>
<td>Services</td>
<td>2</td>
</tr>
<tr>
<td>Retail</td>
<td>3</td>
<td>Consultant</td>
<td>2</td>
</tr>
<tr>
<td>Transport</td>
<td>2</td>
<td>Business Analyst</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
<td>Other</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 4. Sample profile

<table>
<thead>
<tr>
<th>Factors</th>
<th>Best</th>
<th>Worst</th>
<th>B-W</th>
<th>Weighted Best</th>
<th>Weighted Worst</th>
<th>Ratio scale</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Productivity &amp; operations</td>
<td>155</td>
<td>12</td>
<td>143</td>
<td>1252</td>
<td>251</td>
<td>2.23</td>
<td>1</td>
</tr>
<tr>
<td>Information &amp; planning quality</td>
<td>106</td>
<td>15</td>
<td>91</td>
<td>863</td>
<td>226</td>
<td>1.95</td>
<td>2</td>
</tr>
<tr>
<td>Business growth</td>
<td>93</td>
<td>43</td>
<td>50</td>
<td>787</td>
<td>437</td>
<td>1.34</td>
<td>3</td>
</tr>
<tr>
<td>Management commitment</td>
<td>57</td>
<td>34</td>
<td>23</td>
<td>490</td>
<td>329</td>
<td>1.22</td>
<td>4</td>
</tr>
<tr>
<td>Competitive advantage</td>
<td>72</td>
<td>76</td>
<td>-4</td>
<td>652</td>
<td>680</td>
<td>0.98</td>
<td>5</td>
</tr>
<tr>
<td>Vendor support</td>
<td>31</td>
<td>44</td>
<td>-13</td>
<td>292</td>
<td>383</td>
<td>0.87</td>
<td>6</td>
</tr>
<tr>
<td>IT infrastructure</td>
<td>30</td>
<td>53</td>
<td>-23</td>
<td>293</td>
<td>454</td>
<td>0.80</td>
<td>7</td>
</tr>
<tr>
<td>Human resources</td>
<td>33</td>
<td>59</td>
<td>-26</td>
<td>323</td>
<td>505</td>
<td>0.80</td>
<td>8</td>
</tr>
<tr>
<td>Compliance</td>
<td>42</td>
<td>78</td>
<td>-36</td>
<td>414</td>
<td>666</td>
<td>0.79</td>
<td>9</td>
</tr>
<tr>
<td>Acquisition cost</td>
<td>35</td>
<td>76</td>
<td>-41</td>
<td>356</td>
<td>643</td>
<td>0.74</td>
<td>10</td>
</tr>
<tr>
<td>Maintenance cost</td>
<td>34</td>
<td>76</td>
<td>-42</td>
<td>348</td>
<td>642</td>
<td>0.73</td>
<td>11</td>
</tr>
<tr>
<td>Internal productivity</td>
<td>20</td>
<td>71</td>
<td>-51</td>
<td>231</td>
<td>588</td>
<td>0.63</td>
<td>12</td>
</tr>
<tr>
<td>Opportunity cost</td>
<td>19</td>
<td>89</td>
<td>-70</td>
<td>241</td>
<td>731</td>
<td>0.57</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 5. Ranking of ERP decision factors

The best and worst frequency score were calculated for each of the 13 attributes according to the number of times the attribute was selected by respondents. The simple rank ordering process creates individual-level scales for each attribute that are easily comparable across the entire sample (see Table 5). The “best” column illustrates the frequency that the particular attribute was ranked “best” or matters “most” to respondents from the group of factors. For example, the top-scoring attribute was

Proceedings ECIS 2009
productivity and operations (selected 155 times), followed by information and planning quality (selected 106 times) through to opportunity cost (selected only 19 times).

The “worst” column shows the frequency with which respondents selected an attribute as the “least” important feature. To determine the rank ordering for the factors we created a best-worst ratio scale. This scale provides a relative rank for each factor compared to every other factor. To develop the ratio scale of “best” we calculated the square root (SQRT) of the weighted best/worst based on Luce and Suppes (1965) ranking theorem and the mathematical proofs provided by Marley and Louviere (2005).

The interpretation of Figure 1 requires some discussion because the scores are on a relative scale. This means that information and planning (2.23) is two times more important than competitive advantage (0.97) and four times more important than opportunity cost (0.57). Likewise, information and planning (1.95) is more than twice as important as human resources (0.80), compliance (0.79) and acquisition cost (0.74). In general, we see that ERP benefits outweighed the risks when it comes to making an ERP investment, with four of the top five ranked factors drawn from the list of potential investment benefits. Another effective way to evaluate how key decision makers trade-off between the perceived benefits and risks when making an ERP investment decision, is to compare the average B-W ratio score for the two groups of factors. Using the classifications shown in Appendix 1, we see that the benefit factors had an average B-W ratio score of 1.35 compared to just 0.80 for the risk factors. Stated simply, this finding reveals that benefits of ERP are weighted more than one and a half times more important as the risks when making an ERP investment decision.

Figure 1. Plot of B-W Ratio Scale

5 DISCUSSION

The results of this study highlight some interesting findings. Our study supports prior research on the impact of IT in the organization which suggests productivity and operational improvements remain important to organizations investing in ERP technologies. Numerous studies on IT impact and adoption report that operational and productivity improvements are a key driver of ERP adoption decisions (Mirani and Lederer 1998; Tallon et al. 2000; Shang and Seddon 2002; Gattiker and Goodhue 2004). Based on the theory of production, ERP adds value by increasing output whilst reducing costs, through the automation of production processes (Shang and Seddon 2002). For example, several previous studies have found that organisations that have invested in ERP lowered
inventories, shortened delivery cycles and had shorter financial closing cycles, which lead to an overall reduction in costs (Beard and Sumner 2004). In fact, productivity and operational improvements have long been considered as a central source of IT value.

We also report that improved information and planning has a positive influence on ERP adoption, given its ranking as the second most important factor. Access to accurate and timely information is a major source of value stemming from the adoption of ERP. The improvement of coordination amongst functional areas is central in supporting organisational processes and decision making (Sethi and Cárriher 1993; Tallon et al. 2000). The reporting functions of ERP enable faster retrieval of information formatted in a concise manner, thus allowing more control over organisational performance. This means that ERP is ideally positioned to provide decision and planning benefits through the use of centralised databases with built in analysis and business intelligence tools (Shang and Seddon, 2002).

Perhaps surprisingly, financial risks (acquisition cost, maintenance cost and opportunity cost), whilst being dominant factors in the literature around technology adoption and implementation, make up three of the four least important factors to the ERP adoption decision. When considered against the numerous other decision factors, cost based risks appear less important than prior research has led us to believe. For example, it has been observed that organisations cited such risks as a major deciding factor in their ERP purchase decision (Chau 1995; Gupta 2000; Keil and Tiwana 2006).

It is also interesting to note that the first risk factor deemed most important is management commitment, not financial risk. This finding could partly be attributed to the industry press around large ERP failures and past business experience. These reports have made managers increasingly aware of performance around ERP initiatives. We found that having the commitment of management was seen as almost two times more important than, gaining competitive advantage and was on par with creating business growth and information and planning improvements. Without commitment from managers to support strategic IT projects; in terms of resources, including buy-in from the business; the prospect of an ERP project is much less certain. In terms of the strategic management literature, the deterministic perspective suggests that managers are the key drivers of successful strategic initiatives.

Furthermore, the emphasis on benefits vis-à-vis risks in the ERP evaluation process could also provide valuable insight into why so many ERP investments fail. Drawing on the work of Kahneman and Tversky (1972), the findings of this study suggest that managers may be over optimistic about the performance and productivity rewards that may flow from an ERP investment, and subsequently marginalising or even ignoring the inherent risks. Interestingly, this would make ERP consistent with other past emergent technologies, where history has consistently demonstrated that even experts suffer from optimism bias when espousing the future benefits of new technologies (Avison and Nettler 1976).

5.1 Limitations and Future work

Much of our understanding of ERP has been shaped by interviews, case studies and industry surveys (Besson and Rowe 2001; Shang and Seddon 2002; Beard and Sumner 2004; Gargeya and Brady 2005). The few quantitative studies undertaken have essentially focused on single cases or limited sets of adoption antecedents (Hitt et al. 2002; Keil and Tiwana 2006). Whilst prior research and industry evidence has been able to tell us which factors are important, it has not determined the order of importance of these factors to the ERP adoption decision, or the relative importance of benefits vis-à-vis risks. This kind of information would surely be invaluable not only for academics studying enterprise system adoption, and for practitioners trying to make effective investment decisions, but also for ERP vendors trying to develop products that better match the preferences of their customers.

This research is the first to consider the relative importance of the organizational impact factors on the ERP adoption decision of senior managers. Moreover, the study has combined both benefit and risk factors to obtain a more realistic view of how senior managers evaluate the factors that influence an
ERP investment decision. By applying the best-worst method, we add to a growing body of research that suggests that such methods are less cognitively demanding and more accurate in their estimation of preference. Furthermore, the method applied in this study overcomes the inherent problems of scale bias and other scale issues (such as how to interpret individual ratings on a traditional scale). Future work can extend on this study by examining the impact of external factors on the preference structure and ERP adoption choice, or by replicating this research in different contexts and with different decision making factors.

6 REFERENCES


Appendix 1.  Factor definition and classification

<table>
<thead>
<tr>
<th>Factor</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acquisition cost (R)</td>
<td>The total cost of acquiring the initial package (site license).</td>
</tr>
<tr>
<td>Business growth (B)</td>
<td>Increased business growth- the technology provides support for increased business growth</td>
</tr>
<tr>
<td>Competitive advantage (B)</td>
<td>Competitive advantage is created by differentiating core competencies from competitors e.g. Creating a differentiated product or service offering</td>
</tr>
<tr>
<td>Compliance (B)</td>
<td>The technology increases the organisation’s ability to ‘stay on track’ by creating internal controls and structures to ensure conformance to regulations, legal procedures and standards.</td>
</tr>
<tr>
<td>Human resources (R)</td>
<td>The organisation has an adequate level of expertise needed to implement and maintain ERP in-house or has access to external expertise via a consultant or third party service provider.</td>
</tr>
<tr>
<td>Internal productivity (B)</td>
<td>There will be no loss of productivity when the system goes live.</td>
</tr>
<tr>
<td>Information and planning quality (B)</td>
<td>Quality of information (accuracy, reliability and timeliness) is improved resulting in efficient and effective decision making and planning across the enterprise.</td>
</tr>
<tr>
<td>IT infrastructure (B)</td>
<td>Improved base infrastructure which will allow future application investments to be less costly and more efficient.</td>
</tr>
<tr>
<td>Maintenance cost (R)</td>
<td>The annual cost associated with the operation of the system.</td>
</tr>
<tr>
<td>Management commitment (R)</td>
<td>Complete commitment exists from managers, to provide organizational resources to support the investment.</td>
</tr>
<tr>
<td>Opportunity cost (R)</td>
<td>The cost of ERP in terms of the opportunity forgone to invest in other capital.</td>
</tr>
<tr>
<td>Productivity and operations (B)</td>
<td>Productivity is increased and operational costs decreased, through automation and standardisation of business processes</td>
</tr>
<tr>
<td>Vendor support (R)</td>
<td>The technology is supported for at least 5 years and the level of support provided by the vendor is substantial.</td>
</tr>
</tbody>
</table>

Legend: B = benefit factor, R = risk factor.
TOWARDS UNFOLDING CRM IMPLEMENTATION CHALLENGES IN PAKISTAN: A CASE STUDY

Journal: 17th European Conference on Information Systems
Manuscript ID: ECIS2009-0749.R1
Submission Type: Research Paper
Keyword: Stakeholder theory, IT/IS management, Customer relationship management (CRM), Case Study
TOWARDS UNFOLDING CRM IMPLEMENTATION CHALLENGES IN PAKISTAN: A CASE STUDY

Hamid, Huma, Brunel University, UK, currently working as Software Testing Engineer at Bentley Systems, Suite # 307, ETC, F-5\1, Islamabad, Pakistan, huma.hamid@bentley.com

Abstract

In the recent times, Customer Relationship Management (CRM) has become one of the most dynamic topics of Information and Communication Technologies (ICT), both in the academia and the market. This popularity is indeed a result of the promising features offered by CRM. All the studies to date have highlighted elements associated with CRM success by studying its adoption and implementation in the organizations, with a focus on developed countries. Despite the usefulness of these contributions, today, very limited information is available, where implementation of CRM systems has actually been evaluated in a developing country like Pakistan, employing an interpretive approach. Hence, this current research aims to address this gap in literature, not only by taking a non-traditional approach of success evaluation; using system’s stakeholders’ expectations as an evaluation criteria but also focusing a case study from Pakistan, where ICT industry and specifically CRM is in the initial stages of adoption. This research effort has not only revealed the social aspects of CRM implementation but also unfolded some cultural aspects associated with the CRM in Pakistan. In order to guide the research, theory of stakeholders’ expectation failures turned out to be the most suitable option.

Keywords: CRM, success evaluation, stakeholders’ expectations.
1. INTRODUCTION

The concept of maintaining effective relationship with customers has been there for the last two decades, but rapid advancements in ICT and data warehousing concepts, increased awareness of customers and the intense market competition has now given new dimensions to this domain. The reason for increased popularity of CRM is mostly due to the promising benefits it offers in the form of improved company’s performance and long term customer retention for obtaining healthier financial payoffs etc. (Chen and Popvich, 2003). In order to enjoy these promised benefits, several companies have made heavy investments in implementing this technology. However, unfortunately, reaping them does not appear to be an easy task and only 35% (Davids, 1999) to 45% (by Gartner Inc. mentioned in Yoon et. al, 2003 cited in Paulissen et. al, 2007) of the companies have been successful at harnessing them. As CRM implementation greatly affects a company’s external and internal operations, consequently, its successful implementation also becomes critical. Hence, CRM stands out in the pool of other ICT innovations because of its novelty, promising benefits and company wide scope which affects a major part of a company’s stakeholders e.g. employees, suppliers and customers etc.

This increased fame of CRM has also gained immense attention from the Information System (IS) research communities. Numbers of research studies are conducted in order to reveal different aspects of CRM implementation where a mainstream of studies (Bull, 2003; Kim and Pan, 2006; Corner and Hinton, 2002) revolves around identifying the Critical Success Factors (CSF), largely covering the socio-technical aspects. A, widely accepted view is that majority of the CRM related issues revolves around the human factor and particularly the stakeholders involved. Nevertheless, very few have actually identified the role and importance of stakeholders’ expectations associated with the successful implementation of this novel and fast growing technological innovation. Additionally, a strong need also exists in order to explain that how people behave in different social contexts with a particular focus on developing country, while implementing and operating CRM systems.

Therefore, keeping in view the above mentioned gap in literature, specialized role of the research presented here is to expose the practical world of CRM and gain a richer insight for a CRM project success evaluation, by taking an interpretive approach with a focus on the developing country of Pakistan. This has been done by using Theory of Expectation Failure (Lyytinen and Hirschheim, 1987) with a focus on its special instances (correspondence, process and interaction). Also, the study has considered two major types of stakeholders’ which were external (consultants) and internal (management and employees) to the organization. Additionally, an effort has been made to expose the interplay of these stakeholders’ expectations associated while CRM project implementation. The findings of this exploratory study will be able to help the relevant stakeholders to develop a better understanding of the potential problems associated with CRM systems implementation in accordance with the expectations developed at different levels. This, in return, will help them make informed decisions about implementing and managing CRM systems and improve their chances to get promised benefits with minimum losses. Also, the study has unfolded project implementation issues for the emerging IS industry of Pakistan.

Ensuing sections of this report covers different stages of this research study by giving a brief background of existing perspectives on both IS and CRM evaluation and a reflection on the selected theoretical perspective. This has finally aimed to evaluate a CRM implementation in the emerging ICT industry of Pakistan and along with exploring the role of expectations in successful CRM implementation. Finally, research findings and future research directions are highlighted along with some concluding remarks at the end.
2. BACKGROUND

This section presents a review of the state-of-the-art literature in the areas of CRM and IS success/failure studies with a focus on Theory of Expectation Failure done so far. The chapter also presents a reflective illustration of the theory chosen to carry out this research.

2.1. Customer relationship management

Despite the popularity of CRM, still no single agreed definition exists and varying perspectives are found in literature to define CRM. Some researchers view it as a management approach to maintain relationships with customers using technology (Schellong 2005) for others it stands as an IS which enables an organisation to realise a customer focus (Bull, 2003). Another view in the research considers CRM as an essential strategy that integrates knowledge management, data mining, and data warehousing concepts to support an organisation’s decision-making process of maintaining long-term relationship with its customers (Cunningham et al, 2004). In simple words, CRM has established itself as an integrated synergy of people, processes and technology to maintain long-term customer relationship (Chen and Popovich, 2003). The current research study has also revealed that even within a single organization, different stakeholders associated different meaning to this concept depending on their vision, knowledge, role and scope of work. Majority of the client organizations in Pakistan, view CRM as an automation of call centre operations rather a company wide solution.

2.2. Existing perspectives on CRM success evaluation

The surveyed literature has exposed varying perspectives related to CRM evaluation. One view is that, CRM is an integration of technologies, business processes and people; therefore, the perceived success of CRM is subjected to the successful integration of these three entities. Whereas, a dominant stream of research has been devoted to discover the CSF where majority of the CRM practitioners believed that success of any CRM project is dependent on the successful management of these CSFs. Few of such CSFs identified and emphasized in CRM projects evaluation case studies include: need for a strategic vision of CRM (Chen and Popovich, 2003), top management support (Chen and Popovich, 2003), effective leadership (Bull, 2003), interdepartmental communication and coordination (Parvatiyar and Sheth, 2001), employees engagement and motivation (Chen and Popovich, 2003), integration with existing applications (Pan and Lee, 2003) and alignment with existing business processes (Chen and Popovich, 2003). In a similar study, Bull (2003) has investigated the issues ELMS Ltd. faced while CRM adoption and discovered that CRM project turned out to be a failed project due to lack of management commitment, customer driven culture, in-house expertise and resources and a bad outsourcing experience, thus emphasizing on the fact that successful implementation of CRM is subjected to the successful adoption of the mentioned CSFs.

Yet, another view over the matter holds the position that majority of the issues related to CRM implementation are not technical in nature but they revolve around the human factors. It is believed that CRM won’t be successful without managing its impact upon the people who will interact with this system. Although, this factor is predominantly viewed as the most critical factor which decides the fate of any CRM project, however a gap in the available literature and implementation practices has been also highlighted which fails to explain that how people and specifically different stakeholders behave in different social contexts while implementing and operating CRM systems. Even though, the above mentioned range of studies have successfully identified the areas which can cause problems during a CRM project implementation, but further information available on the root cause analysis of such
problems is very limited. Therefore, a strong need has been raised in order to find out the reasons of ignoring such CSF’s by analyzing the CRM project’s implementation in different social contexts.

2.3. Existing perspectives on IS evaluation

Now coming towards the pool of literature which covers IS evaluation studies over the last few years. Numerous studies have been conducted to evaluate and analyze the success/failure of IS from varying perspectives which gave rise to a hot debate between different schools of thoughts. Adding towards the debate around evaluation studies, DeLone and McLean (1992) have fairly argued that out of hundreds of studies conducted for evaluating success/failure of IS, there are as many measures as there are individuals who presented these studies. Largely accepted IS project evaluation includes the measures for success as e.g. use of system by its intended users, user satisfaction, improved decision making, better job performance and financial payoff etc.

Thus, there are various ways in which a system can be analyzed for success in different phases of its implementation and use. But again, like CRM literature, IS success evaluation literature is concentrated around factor-based studies and lacks in the identifying role of organisation in influencing the whole range of activities that can affect project outcomes (Sauer et al, 1997).

A number of researchers have taken into account the renowned perspective of project management factors (resources, time, team and cost) and technology factors (hardware, software, system crashes etc) for IS evaluation. Glass (1999) has criticized the traditional project management constructs which has failed to give a complete picture and reasons for project failures. Third category consists of organisational factor (efficiency and effectiveness) and fourth one has covered the environmental factors (social/economic/political) of IS success/failure (Drummond, 1996). Recent studies have also reported that most of the issues related to IS implementation are either social or political in nature. Therefore, IS evaluation studies ignoring these critical dimensions, present a very limited aspect of any IS failure, hence suffering from conceptual weaknesses.

To address the gap of limited evaluation aspects, a new notion of expectation failure of stakeholders has gained attention of different researchers in the recent years. This theory strongly holds the view that for different people (known as stakeholders), success holds different meaning. It also challenges the fact of measuring success on generalized basis and strongly argues that success is a rather subjective measure. In the next section, theory of Expectation Failure will be discussed in more detail.

2.4. Theory of stakeholders expectations failure

According to the Theory of Expectation Failure, a system is perceived as a failed system when it fails to meet the expectations and satisfaction levels of the stakeholders involved with it (Lyytinen and Hirschheim, 1987). Though this theory has been presented in late eighty’s by Lyytinen and Hirschheim, still it has formulated the basis of the studies conducted for evaluation of stakeholders’ expectations for the IS projects at different stages primarily because of its inherent understanding of strong relationship between IS and its stakeholders’ perception of success.

Lyytinen and Hirschheim (1987) argued that evaluation of success presented in different literature studies for IS projects, lacked in defining IS failures and success. Most of the studies identify only one or limited aspects of the IS failures. Therefore, according to Lyytinen (1988) problems with IS systems and projects are too many and just few features are not sufficient to describe this rich phenomenon. To answer the limitations and gaps in the literature available for IS failures, Theory of Expectation Failure
has been presented. Theory of Expectation Failure not only takes into account the multidimensional view of project failure but also represents the pluralistic and political account. Moreover, it doesn’t neglect the human factor and give immense importance to the stakeholder’s associated with any IS project. It also covers three other notions of failure (correspondence, process, interaction) which are special instances of this bigger theme as shown in Figure 1.

![Figure 1: Special Instances of Expectation Failure](image1)

Considering Davies et al’s (2004) view that failure is not an objective concept and it depends on the position and perspective of the definer, Lyytinen and Hirschheim's (1987) concept of the expectation failure holds a special position in the IS evaluation literature as it refers to the inability of an IS to meet a specific stakeholder group’s expectations having varying perceptions of IS success. These people, which are known as system stakeholders, possess some distinct characteristics and they affect and are affected by IS implementation and its usage, which purely depends on their role, relationship and level of impact with respect to the system. Failure for one group or an individual might be success for another group(s) or individual(s). Therefore, no generic or absolute definition of success/failure exists and it is relative to different stakeholders.

2.5. Expectation failure and CRM systems

Current literature has revealed a strong relationship between stakeholders’ expectations and an IS by using expectation failure notion. But it is also found that, CRM being a novel and most popular IS innovation of the recent years has remained deprived of such a rich evaluation phenomenon and CRM based studies suffered from the similar conceptual weaknesses as identified by the supporters of expectation failure theory for other IS projects. It is evident from the earlier discussion, that CRM holds a very special place in the IS domain due to its globally wide scope where CRM may affect much larger number of stakeholders (internal or external) on national/ international level which have already implemented it or looking at its worldwide trends to make decision for its adoption. Therefore, this attractive, popular, global and specialized system itself raises the need to explore it using rich phenomenon of success evaluation in order to understand its failure reasons in a better way. Based on the above discussion, a strong need has been raised to evaluate a CRM project implementation in a developing country like Pakistan by using the above mentioned rich phenomena of success assessment. Moreover, system implementation challenges and role of different CRM stakeholders in the particular social and cultural context of Pakistani ICT market can also be explored and analyzed.

In order to answer the above research queries, a case study of Pakistani bank has been used to evaluate the success/failure of a CRM implementation project. As Theory of Expectation Failure was used to guide the research therefore, certain aspects are taken into account for the analysis of a CRM system being implemented i.e. i) what are stakeholders’ expectations or common pool of values associated with CRM at different levels? ii) How far these expectations are translated to meet the final objectives of CRM project? iii) How CRM project survived the correspondence, process and interaction
evaluation (this includes the relevant measurement criteria against all these special instances e.g. time, cost and resources overrun, number of change requests and level of acceptance of design objectives, employees training and response to system etc)? It is also worth mentioning here, that in Theory of Expectation Failure, the nature of assessment procedure is informal (no well defined technique is applied), dynamic and against a continuous measurement scale. Other than that, study also revealed aspects of expectation development and affect on final outcome.

3. RESEARCH METHODOLOGY

To achieve the guiding principles of interpretive research, a combination of different research methods were selected. For data collection, a specialized form of interviews (i.e. telephonic interview), open ended email communication and media reports were used. Interviews are generally considered as one of the most strong and popular methodological approaches, whereas telephonic interview is also a well supported practise which can be considered due to limitations of resources (distance, time, cost and participants availability (Saunders et. al, 2003). All the interviews were conducted using telephone from UK to the participants in Pakistan and to ensure that maximum advantage, all the interviews were tape-recorded with the permission of the participants. Moreover, open ended email communication during interaction with the participants also greatly helped to seek richer insight about the research questions by engaging research participants into open-ended discussions about their personal experiences. Some press releases available on the vendor and client companies’ websites have also been used as research data. In these press releases, both the client and consultants have shared their expectations they have associated with that particular CRM implementation project. Total 9 participants were invited to take part in the research study and data was collected over a time period of almost 6 months.

The researcher aimed at analysing the data collected from the study by identifying relevant themes from the data relating to the research question by using thematic analysis approach (Boyatzis, 1998). Boyatzis (1998) has defined thematic analysis as a process which assists in the encoding of qualitative data by identification of specific themes. The data was analyzed using combination of both deductive and inductive approach for research. The reason for adopting these two approaches for data analysis further supports the idea of using initial combination of hybrid methods (semi-structured and in-depth interviews, open-ended communication and media reports) used for data collection mentioned in the above section.

4. CASE STUDY: IMPLEMENTATION OF CRM IN BANK A

This section unfolds the current practices of Pakistani CRM by using case study of a bank. Before going into the details of the selected case study, the rationale behind choosing this particular case study from the banking sector is also worth mentioning here. Currently there are two sectors in Pakistan which are highly influenced by the latest advancements in ICT to stay abreast in the competition i.e. banking and telecom sectors. One possible reason of this could be that, even internationally these two sectors are found to be the most adaptable ones in embracing the technological changes. Secondly, in the recent years major foreign investments have been made in Pakistan, particularly in these two sectors. A number of telecom companies and banks have started their operations locally, where they not only brought their business setups to the local market but, also carried along the latest technological systems already popular and proven successful in these companies in their respective (developed) countries. This shift has given a new challenge to the local players, but also opened new horizons of technological advancements for them to level the competition brought by the foreign
companies. Hence, out of the limited CRM implementations in the local industry, majority CRM case studies can be found in these sectors. Now selecting bank’s case study out of that was a limitation, because the data provided by the resources from telecom companies was not sufficient enough to build a case. The researcher strongly felt that due to immense competition, major players of telecom industry were also reluctant to share their true insight experience of CRM implementation.

4.1. Rationale behind CRM adoption

Bank A refers to a newly established bank which aimed to rapidly expand its network of branches both locally and internationally. It planned to capture the consumer market by offering a wide range of financial products and services along with the adoption of state-of-the-art technology to stand out in the market. The top management of Bank A took a major decision on CRM adoption and this motive was considerably seemed to be inspired by the popularity of CRM’s success in the international financial sector. Therefore, a strong commitment and support to exploit the existing opportunity in the form of CRM technology was witnessed within the top management of Bank A. This fits in what Corner and Roger (2005, page. 268) highlighted as “It is reasonable to assume that the software systems used as the technology basis for CRM are reasonably competent because many are sold, and in similar environments will have been successful”.

4.2. CRM system selection and requirement gathering phase

Just like majority of the companies in the market, Banks A had no option but to render services of a CRM vendor/solution providing companies (referred as consultants) due to insufficient in-house expertise (even if it was a primary choice or not). After a detailed market research, the Bank A’s management finalized the solution which not only suited their existing needs but also offered much more than they actually required, it was rather 100 years ahead of the Bank A’s capabilities and expectations (as claimed by the consultants). The world wide popularity of that particular product played a major role in its selection. It would be interesting to mention here that, the product was selected on the basis of popularity, rather than the actual need, therefore critical phase of requirement gathering started after the system’s selection. But this order is also justified as at project selection level, only abstract level requirements are shared and detailed business process re-engineering normally starts after system selection. In the requirement gathering phase of Bank A, business professionals, executive level management and managerial heads within their own respective domains coordinated with the consultants. Till that point, the middle management and lower level staff was not even aware of any such initiative. Though rumours were circulating about an upcoming system, but no official information was provided to take employees into confidence. This greatly challenges the need raised by Corner and Rogers (2005) that, employees needs to be informed and engaged beforehand.

This lack of communication brought what numerous studies have indicated i.e. employees’ resistance due to lack of information. Therefore, an interesting situation was raised when a middle manager was introduced to explain his core business operation and the consultant tried to impress him with an efficient technical functionality for opening a new bank account by reducing time from one hour to two minutes. It was expected by the consultants, that the manager will be agreed at once. However, the manager insisted on doing it the way it was previously done, where a new customer had to fill-in a lengthy paper application form of up to 15 pages. Upon enquiring about the reason for this resistance, two different view points were found at both the ends. The consultant responded that it was just lack of awareness at their client’s end who didn’t know what they were missing. While on the other hand, the manager explained that his resistance was merely because of data validity and security concerns which are more important than a remarkable technical functionality. However, at the end of the day the
manger had to convince himself and his resistance was over shadowed by the functionality being offered. Upon agreeing on the system’s requirements product development phase started.

4.3. First demonstration phase

Upon completion of the product development, first demonstration phase was held after few months. During first demonstration of the application, normally termed as the testing phase, the Bank A’s management was highly impressed with what the system could perform. Different responses were observed at this particular stage:

- The top management’s initial response exhibited that the system lived up to their initial expectations and they were glad that new systems will help them to monitor staff level activities.
- Call centre agents were particularly excited as in they view CRM is to handle call centre operations.
- IT support department team was puzzled and was in a dilemma whether they actually needed this system? Because once the new system would be in place, they would need to adapt to it, otherwise their jobs would be at an obvious stake.
- Sales and marketing managers were particularly excited and looking forward to see that how their reporting and analysis will become easier and better.

At this stage, a major concern identified by the consultants was lack of awareness at employees’ level which caused job insecurity and low motivation among them. Whereas, the management didn’t find these concerns of much importance, therefore did no effort to eliminate them.

4.4. Final rollout phase

After the satisfactory results of the initial testing phase, a staged implementation was adopted where CRM application was implemented in one department after another. This implementation went smoothly with no delays, concern or systems integration issues as previous application was replaced by new one. After completion a follow-up visit was conducted by the consultants in order to analyze the system’s performance. Some interesting observations were made by them, which are not new to CRM literature.

- The call centre staff particularly complained about the complexity of user interface having too many screens. According to them, it has actually increased the response time to customers instead of making the tasks easier (as they were told).
- Most of the staff particularly avoided talking to the consultants about the system as they feared that speaking negatively about the newly deployed solution early in its stages will pose threat to their jobs. The consultants observed scarce use of the system. The obvious reason for this was inadequate training and also less knowledge sharing culture between employees which also affected the expected outcomes and also the possible systems benefits.
- Some unrealistic assumptions on the part of higher management regarding the response of employees also contributed to the situation. They had the assumption that the employees will get used to the new system with the passage of time especially when they will realize that it is an ultimate solution to their day to day problems. It was also assumed that since the employees have been briefed positively about the functions of the system, therefore, they would definitely be happy to use it and explore it themselves rather than providing formal training sessions.
- The management was also reluctant to provide access to critical data to the lower-level staff. On the other hand, the consultants held the view that to get better results from the system it was an important working practice.
Throughout the implementation, consultants kept on reviewing the status and recapturing different responses to revisit their implementation strategies. With that, CRM implementation phase ended without any significant time, cost and resources overrun issues (as mentioned by research participants, no research data/project reports have been shared with the researcher in support of this argument). Later on, during another follow-up visit, some better response from employees was observed. Also, within 2 months of system deployment, no major change requests were made by the client and system use was also improving gradually.

5. **KEY FINDINGS AND ANALYSIS**

In this section, the researcher has discussed the research findings, where different instances of expectation failure and varying expectations of different stakeholders during the multiple phases of the CRM project are revealed. Study has also exposed the dynamism of expectations which affected the perspectives about project success. Some traits associated with the Pakistani working culture were also unfolded. In order to find the answers, theory has been mapped against practices, where close similarities between theory and practise are found and some already known issues are witnessed.

5.1. **Project failure modes in light of expectation failure**

In order to use the Theory of Expectation Failure, special instances were used for assessment, where measurements were taken during different phases of implementation. A summary of different project phases and special failure instances of expectation theory are presented in the Table 1.

<table>
<thead>
<tr>
<th>Project Phase</th>
<th>Analysis</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Failure Mode : Correspondence</strong></td>
<td>• Successful achievement of design objectives is correspondence success. In case of Bank A, design objectives translated in terms of requirement specification were seemed to meet. Possible reason was not competency of consultants or managers; rather system was selected on the basis of its already renowned worldwide features. Therefore, Bank A tried to adapt itself according to already available functionality and process redesigned by the consultants. Also, no major change requests were made within 3 month of project launch, which strongly reinforce the above mentioned point.</td>
<td>Quite Successful</td>
</tr>
<tr>
<td>System Selection &amp; Requirements Gathering</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td>• Exact cost and time was not shared with the researcher even on enquiring but both the parties claimed that there were no significant delays cost or time overruns. Therefore the project was assumed to be finished within the planned time and resources. • No major technical, system or process integration issues were reported.</td>
<td>Quite Successful</td>
</tr>
<tr>
<td>Over all Project Implementation</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interaction</strong></td>
<td>• Negative attitude towards system, scarce use of it and complaints about the complex user interface were reported and also used as assessment for interaction failure mode. Possible reasons of this could be inadequate training and management’s resistance to allow access to critical data to lower level staff. • System use was increased with time but it is really hard to say if that improvement came due to increase in satisfaction level or system use was enforced.</td>
<td>It was overall together a win-win situation. Though, data was not sufficient to support complete success of complete failure</td>
</tr>
<tr>
<td>Project Launch Phase</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
If we analyze the CRM implementation of Bank A within pre and post-implementation phases, it is found to be fairly successful against correspondence and process instances. Design objectives which management expressed in the form of system’s scalability, automation of major operations (especially call centres) and monitoring of staff activities were fairly met by the system. Due to less rigidity in the business process being a young bank, also played a major role in having flexible design objectives. But, it was strongly felt that measuring design objectives is a hard task, also mentioned by Theory of Expectation Failure. These objectives are not always expressed and requirement specification is just a too little aspect which can cover the overall design objectives. Also, validation of mentioned objectives is challenging. Therefore, this dimension of project evaluation provided very limited information to decide a project’s success. The project performed well against “process instance” which was basically due to the fact that top management being major project’s sponsor, provided sufficient resources and priority to the project as needed. On the other hand, if user attitude and system use is considered as assessment criteria, the system significantly failed from user interaction point of view. Yet again, these two measurement criteria were not sufficient for a fair assessment as in case of Bank A, where system use was finally enforced by the higher management and employees had no option but to use the system.

Not many studies have highlighted the silent but significant role of consultants for the overall project’s success during implementation phases. In this case, consultants being a major stakeholder contributed towards the overall project success. They remained optimistic throughout the project and guided their client on certain occasions. Oppositely, consultants were also fortunate that Bank A was a committed and flexible client. It was also found that “expert power” has great role in developing and managing expectations. The consultants equipped with power of expert technical knowledge, throughout lead the whole project and their hold on the project planning and layout seemed very strong. With that, they were able to mould their client as per their project planning and project outcomes. This depicts that consultants and their client shares a very strong and dynamic relationship. This relationship affects the overall expectations about the project and by shaping them and achieving them at later stages.

5.2. Additional themes

Apart from analyzing the project implementation in a particular social context for a particular case study, few additional themes were also identified during the research study. One prominent theme was cultural difference of Pakistani working practices from the western world from where majority of the CRM solutions are being “imported”. Few observations which were made during the research study particularly for Bank A and generally for Pakistani CRM industry are as follow:

- Even though number of companies within Pakistan have already adopted, in process of adoption or making up their mind for adopting CRM, still there are certain cases where this decision is merely influenced by the international market trends rather than companies’ actual needs as it is also witnessed in this particular case study. Though, this trend can be good or bad, both ways.
- The research participants shared the view that majority of the CRM client organizations’ vision lack clarity. Particularly in the banking sector, CRM is no more than an automation of call center operations. Such clients either turn out to a blessing for the vendors as they know less and get easily convinced or totally get the other way around and increasing the project failure risk.
- Higher managements mostly work in an autocratic way, where middle management and non-managerial level employees are less empowered. They mostly had the view that it’s not important to tell their subordinates about an upcoming system. A call centre agent neither has the time, nor interest in knowing what system his company would be launching. He would do whatever he was being told. In this particular research study it was found that neither the staff was officially
informed about the system nor sufficient training sessions were conducted. That resulted in job
insecurity, lack of confidence and resistance to use the new system. Although, this factor was
later eliminated by enforcing system use by the managers.

- The renowned theme of CRM being integration of people, process and technology which is
  emphasized by the research participants, mostly found as a lip praying service and not as a
  working practice. Technology is mostly considered as the ultimate solution to achieve potential
  benefits and end users of technology are mostly ignored. But, the extent of this trend tends to vary
  from company to company as well. Firms’ employees having a history of frequent changeovers of
  systems and technology are found to be more adaptable and vice versa.

- Two very interesting comments were made by the CRM implementation consultants with regards
to CRM practices in Pakistan. One view was “CRM will blend and merge itself with the local
colours and flavours just like the international fast food chain McDonalds did, not only bringing
new recipes from foreign countries but also adding and maintaining the taste of local culture”
and the other view was “The concept of democracy which is so popular in the Western world can
not be applied with its full strengths in Pakistan due to less education, knowledge and awareness
to select the right leader, same goes for CRM. The solution is indeed international but it will take
immense learning experience to get it matured in the Pakistani corporate environment”

- It is also believed that problems with CRM implementation in Pakistan are same as they are
internationally. Rather, they can become worse due to less technical culture and rigid working
environment especially in the public sector.

6. FUTURE RESEARCH AND CONCLUSION

In this particular research study, an effort has been made to understand and analyze the implementation
of one of the most popular and novel IS innovations i.e. CRM in a particular social context by using a
Pakistani bank’s case study. The research has its limitations in certain areas which include i.e.
unavailability of official data about the current size and capabilities of Pakistan’s CRM industry, CRM
industry is in still early stages of adoption in Pakistan, research study has limited to only one detailed
case study due to limited participation from industry and also the time and resource constraints has
restricted the effective generalization of findings. It is therefore suggested that, future research can take
larger research samples to identify and verify the established relationships and also the researchers
need to identify the organizational level changes in Pakistani working culture which has a promising
future in IT growth.

The research study has revealed that certain social and environmental factors affect any IS project’s
success and its possible outcomes. Also, the project success or failure has various dimensions
therefore, analysing a project from multiple dimension has provided a better and clear view of the
overall picture. People factor has play a very critical role in project’s success therefore; they are needed
to be managed wisely. Technology alone can not perform miracles if people do not use it properly. In
this regard, relationship shared by different stakeholders is also critical to any CRM project success.

There are certain culture differences within Pakistani CRM market which can adversely affect CRM
success therefore, instead of merely following the international trends, an effort should be made by the
companies to increase awareness at all levels within a company so that better results can be achieved.
The future of Pakistani CRM market is getting better and better and there is a great potential for
getting the promised benefits from the CRM systems. Only suggestion I would like to make is that
instead of reinventing the wheel, stakeholders should learn from the mistakes of the western
companies. If they have imported the solution from these developed countries, then those failure lessons should also be imported and applied in order to avoid damage and financial losses.

References
PREDOMINANTLY ELECTRONIC OR PERSONAL SERVICE DELIVERY? A CASE IN THE WEALTH MANAGEMENT CONTEXT

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0200.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Consumer behavior / choice / demand / empowerment / reviews / consumerism, E-Service, E-commerce (B2B / B2C / B2G / G2C), Face-to-Face Communications</td>
</tr>
</tbody>
</table>
PREDOMINANTLY ELECTRONIC OR PERSONAL SERVICE DELIVERY? A CASE IN THE WEALTH MANAGEMENT CONTEXT

Sunikka, Anne, Helsinki School of Economics, Department of Business Technology, Runeberginkatu 22-24, 00100 Helsinki, Finland, firstname.surname@hse.fi

Abstract

Financial services have been a recurrent subject of a multichannel inquiry but investigation into the wealth management area is scarce. This paper intends to fill the gap and presents the results of a questionnaire directed at customers of a financial conglomerate. The objective of this research is to examine which variables influence consumers’ channel preferences in the wealth management context, and to find out possible differences between the customers who prefer predominantly electronic service or personal service delivery. Logistic regression and t-tests are used in the analysis. The perceived channel attributes of personalization, convenience and safety, relationship strength, and the internet and wealth management knowledge influence the channel preferences. Typical wealth management customers prefer multichannel service delivery; only 4% of customers prefer pure electronic service, and 14% of customers prefer pure personal service. There are several aspects that differentiate those customers who prefer predominantly electronic or personal service. The preference for the electronic channel indicated investments in shares, independent decision making style in wealth management tasks and reliance on electronic information channels. In addition, the customers who perceive relationship strength with the service provider as weaker prefer predominantly e-services, which should give impetus for action among the management in the financial service companies.

Keywords: Multichannel, Wealth Management, Electronic Service, Personal Service, Relationship Strength
1 INTRODUCTION

Multichannel customer management is the design, deployment, coordination, and evaluation of channels through which firms and customers interact, with the goal of enhancing customer value through effective customer acquisition, retention and development (Neslin et al. 2006). Multichannel retailing is assumed to offer synergies, as it can result in an increased customer base, added revenue, and higher market share (Berman and Thelen 2004). A well-integrated multichannel strategy includes product, service and quality consistency across channels, highly-integrated promotions and integrated information systems that share customer and offering information. Much of the multichannel research has taken place in the financial industry for several reasons. Financial services have been in the forefront of the new technology deployment, since the financial industry is essentially an information-based industry (Dewan and Seidmann 2001). In addition, financial service providers control both their online and offline channels and can decide fairly freely how to develop the channels, and where the focus of the development should be. In the financial sector, multichannel service delivery is thus commonplace. The average adoption rate of online banking is high in the Nordic countries making the need for multichannel strategy essential. According to the statistics from spring 2008, 83% of Finns use the internet, and 72% of the population are online banking users (Statistics Finland, 2008).

This study is based on the results of both qualitative and quantitative data that were collected under a larger research project. The main aim of the project was to develop a comprehensive wealth management service concept for individual customers. The qualitative methods included financial expert interviews and consumer focus group discussions. The main focus of this paper is, however, on reporting the results of a questionnaire sent to customers of a Nordic financial conglomerate (N = 291). The sample customers have some property and savings, representing thus a customer segment that financial service providers will find increasingly interesting in the future. They are, however, not entitled to private banking services. In our case financial conglomerate only customers who have 100 000 € of assets for investment purposes can use private banking services. The sample customers thus manage their assets either independently or after infrequent consultations with a financial advisor.

This paper broadens the research focus from transactional services (mainly payment of bills and checking account balances) to wealth management services. For the purposes of this study, we understand wealth management broadly as those activities focusing on financial issues in the households; real and financial assets and liabilities, including insurances for protection of possessions and persons. We emphasize a comprehensive view to wealth management services, and recognize that wealth management requires knowledge and experience that are not necessary in day-to-day running of financial affairs. For example, in the preliminary testing phase of the wealth management service concept, a financial security check” – the opportunity for consumers to map their own financial situation – and compare it with others belonging to the same demographic group, was appreciated.

The objective of this research is to examine which variables influence consumers’ channel preferences in the wealth management context. Consequently, a model of variables influencing the consumers’ channel preferences is developed. We examined the channel attributes; convenience, security and personalization, and whether customers’ perceptions of the relationship strength with the service provider are associated with channel preferences. In addition, internet and wealth management knowledge and experience were hypothesized to have an impact on the channel preferences. Our paper thus concentrates on finding an answer to the question: What variables influence consumers’ channel preferences in the wealth management context? In addition, we carry out an exploratory analysis on how predominantly electronic service (PES) and predominantly personal service (PPS) customer groups diverge from each other.

This paper is structured as follows. We first discuss the literature investigating the multichannel usage, and then develop the hypotheses for the model. In section three we describe the data and methodology used in the empirical study. After that, the results based on the binary logistic regression and t-test analysis are presented. Finally, the results and theoretical and managerial implications are discussed.
2 THEORETICAL BACKGROUND AND HYPOTHESES

Sousa and Voss (2006) define multichannel service as a service composed of components that are delivered through two or more channels. According to Neslin et al. (2006), practitioners have five challenges to address in the effective management of the multichannel environment: i) data integration, ii) understanding consumer behaviour, iii) channel evaluation, iv) allocation of resources across channels, and v) coordination of channel strategies. Academic research has mostly addressed the question of consumer behaviour, and has concentrated on three main channels: catalogs, bricks-and-mortar stores, and the internet.

Previous research has given evidence for reasons for channel choice and concluded that multichannel consumers, in general, buy more (Kumar and Venkatesan 2005). According to Neslin et al. (2006), the main determinants of customer channel choice can be divided into five groups: *marketing efforts* (Ansari and Mela 2003), *channel attributes* (Devaraj et al. 2006), *channel integration* (Montoya-Weiss et al. 2003), *social influence and situational factors* (Burke 2002, Dabholkar and Bagozzi 2002) and *individual differences* (Durkin 2004). In addition, the *task characteristics* of goal-directed or experiential tasks (Hoffman and Novak 1996) and the *type of products purchased* (Chiang et al. 2006) influence the channel choice. Less research has been conducted in the area of data integration even though it is important for financial companies (see however, Cappiello et al., 2003).

It is likely that the characteristics of wealth management services influence the channel choice and the need for channel interaction. Long-term wealth instruments and services consist mainly of credence attributes (Darby and Karni 1973), and are marketed and sold with promises of future revenue streams and credibility of the service provider (Harrison 2000). The channel – service framework (Apte and Vepsäläinen 1993) concludes that complex and infrequently used services would typically require personal interaction whereas simple and frequent transactions can be carried out as a self-service. In addition, the media richness theory (Daft and Lengel 1986) emphasizes the richness of personal contact, and its superiority in dealing with complicated issues in comparison to, for example, phone calls or e-services.

Most of the multichannel studies in the financial services context have examined the association between channel choice and loyalty, and the findings have been contradictory. On one hand, multichannel environment can be seen eroding loyalty because it encourages extensive search and enables easy switching to another service provider. In addition, electronic channels entail little human contact, which itself can erode loyalty. Wright (2002) claims that new channel technologies have loosened the relationship between the banker and the customer. On the other hand, multichannel usage might also enhance loyalty (Shankar et al. 2003). According to Wallace et al. (2004), multichannel usage is associated with higher perceptions of the provider’s service offering, which in turn leads to higher customer satisfaction and loyalty. Coelho et al. (2003) investigated 62 UK financial service companies and found that multichannel companies enjoyed higher sales levels but lower profits. It seemed that multichannel companies suffered especially in terms of customer service and customer retention. Thus, providing good, coordinated service is a challenge for multichannel companies. In general, however, several research results indicate that channel choice has no association with loyalty in the banking context (e.g. Colgate and Smith 2005; Herington and Weaven 2007).

Only few studies have examined services related to maintaining and accumulating wealth. Falk et al. (2008) found that in Germany, the status quo bias is a powerful hindrance for consumers to turn to electronic services in the investment context, and especially older, male consumers and inexperienced internet users preferred continuing to use purely personal service. Ding et al. (2007) examined what combination of features financial institutions should offer to satisfy the needs of the high involvement customers who preferred using the e-service, multichannel or personal service for investment needs. Within these different customer segments the customer needs regarding online features differed, indicating a need to personalize the multichannel offering for each customer segment. The electronic service segment included 52 %, multichannel 37 %, and personal service 11 % of the sample’s customers in the study (Ding et al. 2007).
Channel preferences (dependent variable in the model)

Channel choice can range from pure electronic channel (self-service) to pure personal service. The dominant view of the experts we interviewed was that in wealth management tasks, personal encounter with a financial advisor is the best way to influence customers, especially when the customers are inexperienced. More experienced customers might be able and willing to turn to electronic services. Thus, the multichannel view emerged very strongly in the interviews. However, a linkage between the relationship strength and channel preferences could not be made. In the consumers’ focus group discussions, the internet was mentioned as an information source but e-services of banks were not discussed. Topics like security or privacy did not emerge directly in the discussions. Instead, consumers talked spontaneously about the need for trust in the wealth management relationship and their partial distrust in financial advisors (Sunikka and Peura-Kapanen 2008a, 2008b).

The channel preferences are measured with one item “I prefer taking care of my wealth” with anchors at 1, signifying electronic channel delivery and 7, signifying interaction with a financial advisor. The midpoint 4 denotes customers that prefer electronic and personal service delivery to the same extent. For the purposes of the binary logistic regression analysis, customers were divided into two groups; those who prefer predominantly electronic service, PES-group (choices form 1 to 3, n = 72) and those who prefer predominantly personal service, PPS-group (choices from 5 to 7, n = 160). We decided to omit those customers who chose the mid-point 4 from further analysis (n = 59).

Channel attributes

Previous research has identified several reasons why consumers adopt new technology. Especially in the financial service sector, research has concentrated on examining the antecedents of electronic banking adoption (e.g. Durkin, 2004). The technology acceptance model, TAM, (Davis 1989) has identified usefulness and ease of use as the main variables for explaining why consumers adopt new technology. In addition, the consumer readiness (Meuter et al. 2005) and the technology readiness (Parasuraman 2000) constructs explain why some consumers feel more comfortable with new technology than others.

Convenience, security and personalization represent channel attributes in this study. For example, Szymanski and Hise (2000) found convenience to be an important factor in e-satisfaction. Convenience is understood in terms of consumers’ time and effort perceptions related to using a service. Fun of interaction was not included as an attribute since electronic banking (also for wealth management tasks) is considered as a utilitarian electronic service (van den Heijden 2004). In an online context, security, and the perception of trust have been identified as important determinants of customer willingness to use electronic services (Casaló et al. 2007, Yousafzai et al., 2003). According to a recent Eurobarometer (2008) study, consumers in the Nordic countries trust banks and financial institutions to use their personal data in an appropriate way. Personalization has been taken into new levels in the electronic environment, and research interest has continuously increased (for a review, see Fan and Poole, 2006). The chosen attributes were combined with two tasks: information search about wealth-related issues, and buying or selling of assets. Hence, we combined three channel attributes with two wealth management tasks.

We hypothesize that all relationships between the channel attributes and the channel preference are positive. In other words, if a respondent thinks that information search is convenient with the help of a personal financial advisor, the customer is expected to choose options from 5 to 7 and thus indicate her/his preference for the personal service delivery channel. If, on the other hand, a customer regards electronic service as a convenient channel in information search, s/he is expected to choose options from 1 to 3 and thus show preference for the predominantly electronic service delivery. Descriptive statistics of the items and constructs are presented in Appendix 1.

**H1:** Customer’s perception of the channel convenience will be associated with the customer’s channel preferences.
**H2:** Customer’s perception of the channel security will be associated with the customer’s channel preferences.

**H3:** Customer’s perception of the channel personalization will be associated with the customer’s channel preferences.

### Relationship strength

Incorporating the relationship strength as one of the variables in the model was motivated by the question whether financial service relationships are perceived as strong or weak by the consumers. Financial service relationships are often used as examples of the relationship management approach, which is defined as activities directed towards establishing, developing and maintaining successful relational exchanges (Morgan and Hunt 1994). However, there are doubts about the strength of ties that consumers feel towards their financial service providers in the current era where “the competitor is only a click away”.

In contrast to the research on service quality, satisfaction and loyalty (e.g. Beerli et al., 2004), there are only few studies on relationship strength. Donaldson and O’Toole (2000) and Hausman (2001) have examined relationship structure and strength and its impact on performance in a non-finance B2B context. Wong and Sohal (2006) have developed a model of relationship strength in the retail sector, and claim that their results could be generalized to banking and insurance industries. Other researchers have concentrated on the salesperson’s role (Bove and Johnson 2001), or on comparing varying levels of relationship strength in different service industries (Ward and Dagger 2007). Ward and Dagger (2007) found out that of the five service contexts they studied, customers in general perceive that the relationship with a bank represents medium level strength; the strongest relationship is usually with the hairdresser, and the weakest with a cinema.

In this study, relationship strength is defined as the extent, degree, or magnitude of relationship which is governed by the amount of trust and the level of commitment the customer feels towards the service provider (Bove and Johnson 2001). The relationship strength is particularly applicable in situations where the service involves a high component of interpersonal delivery, and when the service is varying and high in experience or credence qualities, making quality difficult to predict or evaluate and therefore increasing the customer perceived risk (Sheth and Parvatiyar 1995).

Ball et al. (2004) grouped the antecedents of loyalty as follows: characteristics of the environment (perceived switching costs or technological changes), characteristics of the dyadic relationship (shared norms or relationship duration), characteristics of the consumer (relationship tendency or involvement in the category), and consumer perceptions of the relationship with the company (service satisfaction, trust and service quality). Consumers’ perception of the relationship strength with the service provider is thus perceived as an antecedent of loyalty. In this paper, the perceived relationship is understood purely from the consumer’s point of view. Originally, we had three items to measure the relationship strength; however, the transactional item had to be excluded from further analysis since it did not fit with the other two items of the construct.

**H4:** Customer’s perception of the relationship strength with the service provider will be associated with the customer’s channel preferences.

### The internet variables: knowledge and experience

In line with previous studies (e.g. Montoya-Weiss et al. 2003), increased familiarity (knowledge and frequency) of the internet usage is likely to result in increased use of electronic services. According to Alba and Hutchinson (1987), knowledge can be divided into subjective and objective knowledge. Subjective knowledge is the own perceived level of knowledge, in this study the self estimated ability to use the internet. Objective knowledge was not measured in this study.

Experience with the internet, both the length of time the consumer has used the internet and the frequency of the internet usage, as well as the versatility of the tasks, are expected to influence channel preferences. In this study, frequency of usage (in hours per week) represents the internet experience.
**H5:** The internet knowledge will be associated with the customer’s channel preferences.

**H6:** The internet experience will be associated with the customer’s channel preferences.

### Wealth management variables

Knowledge on different instruments is assumed to increase the confidence of customers in their own ability to conduct wealth-related tasks independently, without turning to customer representatives for assistance. The frequency of wealth management transactions is also assumed to increase the likelihood of electronic channel usage. Associated with this is the nature of the financial assets owned since, for example, ownership of liquid assets (e.g. shares) can require frequent transactions and might thus result in preference for multichannel and electronic channel service delivery.

**H7:** Wealth knowledge will be associated with the customer’s channel preferences.

**H8:** Wealth experience will be associated with the customer’s channel preferences.

---

**3 METHODOLOGY**

Our qualitative data examined wealth management behaviour both from the point of view of consumers and service providers. At first, focus group discussions were carried out with 33 consumers in six sessions and 11 individual interviews with financial experts. Based on the results of this qualitative phase and an earlier literature, a questionnaire was constructed. The questionnaire included several themes ranging from motivations for financial behaviour to customer views of total wealth management services. Most of the items used in the questionnaire were adapted from previous studies. In addition, practitioners and academics participating in our research project commented extensively on the questions resulting in several changes to the final questionnaire form. 1500 questionnaires were sent out to customers of a financial conglomerate. Two mailings were used. The customers were chosen from the database of the financial services company, and are representative of relatively wealthy customers. The response rate was 20.6 % (309 returned questionnaires), which can be considered satisfactory considering the length of the questionnaire, the sensitivity of the topic, and the fact that this was a mailed survey. After having removed incomplete responses, 291 usable responses were available for our analysis.

---

**4 RESULTS**

The demographics of the 291 respondents were the following: 57 percent were female, the average age was 49 (range from 25 to 75), and 49 percent had a polytechnic or a university degree. Nearly 70 percent earned less than 40 000 € per year and the group owning between 100 001 and 250 000 in assets was the biggest group (35 %). 83 percent of the respondents lived in smaller towns and rural municipalities.

---

**Figure 1. Model for channel preference in the wealth management context**

---

Proceedings ECIS 2009
Confirmatory factor analysis (CFA) was conducted using principal component analysis and orthogonal varimax rotation. Appendix 1 depicts the constructs that were used in the logistic regression model. The security attribute on buying and selling loaded on the convenience attribute, leaving only one item, the security of information search, to represent the security construct. The reliabilities of the constructs (Cronbach’s alpha values) ranged from 0.76 for convenience to 0.84 for relationship strength. Nearly all tasks were considered more convenient, more personalized and more secure when conducted as PPS. Only the information search item was considered slightly more convenient as PES. Correlations among variables ranged from -0.35 (personalization and frequency of wealth management decision making) to 0.64 (channel attributes personalization and convenience).

We conducted a stepwise binary logistic regression with SAS 3.0 to test the hypotheses. The explanatory variables were derived from the factor analysis and the significant variables are presented in Table 1 below. The dependent variable is channel preference, 1 for PES and 0 for PPS. We used the question “I prefer taking care of my wealth” with anchors at 1 for purely electronic channel and 7 for purely personal channel to distinguish those who prefer PES (from 1 to 3, 25% of the respondents) from those who prefer PPS (from 5 to 7, 55% of the respondents). We decided to omit those customers who chose the mid-point 4 to denote their channel preference (20% of the respondents). However, multichannel service delivery was clearly the dominant service delivery model since the pure electronic channel was preferred only by 4 percent and pure personal service by 14 percent of the respondents.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate (B)</th>
<th>Standard Error</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; Chi Sq</th>
<th>Exp (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>1.5401</td>
<td>0.3292</td>
<td>21.8828</td>
<td>&lt;.0001</td>
<td>4.665</td>
</tr>
<tr>
<td>Channel Convenience</td>
<td>1</td>
<td>1.227</td>
<td>0.2938</td>
<td>17.4402</td>
<td>&lt;.0001</td>
<td>3.411</td>
</tr>
<tr>
<td>Internet knowledge</td>
<td>1</td>
<td>-0.7583</td>
<td>0.3194</td>
<td>5.6354</td>
<td>0.0176</td>
<td>0.468</td>
</tr>
<tr>
<td>Wealth mgmt knowledge</td>
<td>1</td>
<td>-0.5927</td>
<td>0.2925</td>
<td>4.1053</td>
<td>0.0427</td>
<td>0.553</td>
</tr>
<tr>
<td>Channel Personalization</td>
<td>1</td>
<td>3.1297</td>
<td>0.5074</td>
<td>38.0397</td>
<td>&lt;.0001</td>
<td>22.867</td>
</tr>
<tr>
<td>Relationship strength</td>
<td>1</td>
<td>0.8288</td>
<td>0.2838</td>
<td>8.5305</td>
<td>0.0035</td>
<td>2.291</td>
</tr>
<tr>
<td>Channel Security</td>
<td>1</td>
<td>0.9689</td>
<td>0.2932</td>
<td>10.9212</td>
<td>0.001</td>
<td>2.635</td>
</tr>
</tbody>
</table>

Table 1: Influential variables for channel preferences

The -2Log likelihood of the final model was 83.651, R-Square 0.556 and Max-rescaled R-Square 0.874. In the likelihood ratio test for the global null hypothesis (beta = 0), chi-square was 160.901 with 6 DF (<.0001). No additional effects met the 0.05 significance level after Step 6. Somers’ D denotes the strength and direction of the relation between pairs of variables. It is the difference between the percent concordant and the percent discordant divided by 100, in our model (96.6 - 3.4)/100 = 0.933. The C measure (equivalent to a well-known measure of ROC) is 0.966, which corresponds to the model discriminating the responses well.

The channel attributes of personalization, convenience and safety are all positively related with the channel preference (H1 – H3); those who rate these attributes high for personal service also prefer PPS delivery. Relationship strength is also positively related with the channel preference; those who perceive a stronger relationship with the service provider prefer PPS delivery in the wealth management context (H4). The relationships with the internet knowledge (H5) and wealth management knowledge (H7) with channel preference are reverse; that is, the increase in these variables indicates a move from personal service preference to multichannel and electronic service preference. The relationships between the internet experience and wealth management experience with channel preference were not statistically significant; we thus reject H6 and H8.

In order to develop a deeper understanding of the differences and similarities between the customer groups that preferred PES or PPS delivery we carried out t-tests with variables that were not included in the logistic regression model (see Appendix 2 for details). We used the same binary variable as in the logistic regression to group the data. Even though we did not think that demographic variables
would have an impact on the channel preferences, there was a statistically significant difference between the groups. PPS was preferred by older customers; the average age of PPS customers was 50 years in comparison with 46 years in PES group (p = 0.0308). PPS group consisted of 63% of females whereas PES group included only 36% of female. In addition, members of the PPS group owned less shares. We asked about the sources of information that customers considered as the most important in the financial decision making context. PPS group turned to financial advisors for information whereas PES group used a variety of information sources to support them in the financial decision making: especially electronic channels but also printed articles and own previous experience. When decision making style in financial matters was specifically inquired, PES group acted independently, relying on their own expertise whereas PPS group needed more advice and reassurance from others. PPS group had a more favourable view of the marketing activities of the financial institutions; seeing marketing as aiming at the good of the customer rather than pushing products to customers. Customers in PPS group intended to increase the use of services of the current service provider indicating behavioural loyalty even though there was no statistically significant difference with the intentions to use services of other companies between the PPS and PES groups. The PPS group also felt that they had invested more effort in finding a suitable service provider creating thus higher switching barriers. In general, PES group had a more critical attitude towards financial service providers than PPS group.

5 DISCUSSION AND CONCLUSIONS

This paper focused on understanding customers’ multichannel usage in the wealth management context, and especially the linkage between the relationship strength and the channel preferences. Multichannel service delivery is clearly preferred to pure electronic service delivery (4%) or pure personal service delivery (14%) in wealth related tasks. Only the information search task was considered slightly more convenient as predominantly electronic service (PES) than personal service, whereas convenience, personalization and safety constructs were all seen as predominately personal service (PPS) attributes. Channel preferences in wealth management are also influenced by the relationship strength; those who perceive a stronger relationship prefer using PPS. Increasing internet and wealth management knowledge, on the other hand, makes customer prefer switching to the multichannel and electronic service delivery.

The t-tests between the PPS and PES groups revealed quite distinguished profiles of the two customer groups. Those who preferred PES were demographically more likely to be younger males and they had more investments in shares than the PPS group. The preference for increased electronic channel usage might be partly explained by their willingness to use e-services for share trade transactions. The PES group used more versatile information sources than the PPS group who mainly relied on the information provided by their primary financial service provider. In addition, customers in the PES group preferred making independent decisions instead of turning to financial advisors for advice. Perceptions about the advice that the financial service companies offered differed too; the PES group thought that the advice of the financial service companies mostly promoted the products or services of the particular company, and not the good of the customer. Furthermore, the behavioral loyalty and perceived switching costs diverged confirming the stronger perceived relationship between the PPS customers than the PES customers with their financial service provider.

Traditional wealth services require expertise and are thus labor intensive. As the number of consumers willing and able to invest is expected to increase in the future, the challenge of suitable service channel combinations will intensify. When electronic and personal services are combined the wealth management service providers are not only able to offer services to a larger number of customers but can also provide more comprehensive and more integrated services to cover consumers’ wealth management needs as a whole. The findings of our study show that those who prefer PES are more knowledgeable and willing to make independent decisions in wealth management issues than the PPS customers. However, there are only 25% of PES customers in this sample compared with 55% of PPS customers. Thus, even though customers are active users of online banking in everyday monetary affairs, the majority of customers seem to need personal interaction with the financial advisor in a
more complex context at the moment. Also, since the multichannel strategy emerged as the most preferred channel choice both electronic and personal channels have to be developed in an integrated manner.

As it seems that the PES customers perceive a weaker relationship with their service provider, the financial service providers should emphasize the added value that their electronic channel provides for the customers. Nowadays, information technology offers several possibilities for automated and personalized service delivery. However, e-services are not necessarily perceived as personalized even though customers find their own account and transaction information in the online applications. The information provided by the financial service provider is not personalized, and not necessarily relevant for the customers’ situation, or their financial objectives in life. A more personalized experience could be offered with a tool that aggregates the customers’ financial transactions in order to show their financial status and the allocation of wealth automatically, without manual calculations. A more hedonic application is a widget that is provided by an e-service of one bank: the customer can choose a certain objective for savings, for example, a vacation trip, and the widget records all the savings by depicting a piggy bank that is becoming rounder when the savings amount accumulates. In addition, chat and VoIP could be harnessed for customer service purposes since one of the recurrent sources of complaints is that it is difficult to get in contact with financial advisors. However, the linkage to personal service should always be kept in mind, and the PES customers should be served efficiently when they need personal service. For example, Colgate and Smith (2005) studied multichannel financial services, and concluded that a good relationship with a financial advisor can build trust among the e-service customers.

Our research has the following limitations. Only customers of one service provider were surveyed, limiting the external validity of the study. Because the questionnaire was long (11 pages), and there were several themes, the number of items for the channel and relationship questions was restricted. According to Drolet and Morrison (2001), one-item constructs are not necessary harmful in service research. For example, Shankar et al. (2003) only used one-item constructs in their multichannel study. The main reason for this was to avoid the excessive length of the questionnaire, as in our case.

It should be noted that we do not know if customers who perceive weaker relationship prefer PES, or if the usage of PES leads to weaker relationship due to the decreasing amount of personal contact. More studies should concentrate on finding out how the financial service offering should be combined in various channels, and what kind of services are regarded as attractive in the electronic channel. Further studies should also examine other industries and companies to confirm the relationship between the electronic service usage and relatively weaker relationship with the service provider.

References

Proceedings ECIS 2009


**Appendix 1: Variables, mean averages, standard deviations, and Cronbach alphas for the constructs**

<table>
<thead>
<tr>
<th>Construct</th>
<th>Item</th>
<th>Factor loadings</th>
<th>N</th>
<th>Avg</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Channel convenience</strong></td>
<td>Information search about wealth is the most convenient (1 = as a pure electronic service, 7 = as a pure personal service)</td>
<td>0.5711</td>
<td>291</td>
<td>3.73</td>
<td>1.67</td>
</tr>
<tr>
<td>CA = 0.7611</td>
<td>Purchasing and selling wealth is the most convenient...</td>
<td>0.8057</td>
<td>291</td>
<td>4.15</td>
<td>1.74</td>
</tr>
<tr>
<td>Avg: 4.29, Std: 1.34</td>
<td>Purchasing and selling wealth is the most secure...</td>
<td>0.7656</td>
<td>291</td>
<td>4.99</td>
<td>1.45</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
### Appendix 2: T-test results

<table>
<thead>
<tr>
<th>Source of Information</th>
<th>Predominantly personal service (PPS)</th>
<th>Predominantly e-service (PES)</th>
<th>Significance *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share ownership (1 = none, 7 = all assets)</td>
<td>Avg. 2.06 SD 1.667 N 140</td>
<td>Avg. 2.82 SD 1.928 N 65</td>
<td>0.0048*</td>
</tr>
<tr>
<td>Personal advice from own financial advisor</td>
<td>Avg. 5.20 SD 5.202 N 158</td>
<td>Avg. 4.37 SD 4.767 N 71</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Material distributed at branch offices</td>
<td>Avg. 4.11 SD 1.371 N 158</td>
<td>Avg. 3.67 SD 1.322 N 72</td>
<td>0.0231*</td>
</tr>
<tr>
<td>E-services of own financial institution</td>
<td>Avg. 4.03 SD 1.451 N 152</td>
<td>Avg. 4.54 SD 1.401 N 70</td>
<td>0.0135*</td>
</tr>
<tr>
<td>E-service of independent service providers</td>
<td>Avg. 3.73 SD 1.461 N 154</td>
<td>Avg. 4.33 SD 1.411 N 70</td>
<td>0.0043*</td>
</tr>
<tr>
<td>Articles in print magazines and newspapers</td>
<td>Avg. 4.23 SD 1.295 N 157</td>
<td>Avg. 4.67 SD 1.322 N 72</td>
<td>0.0193*</td>
</tr>
<tr>
<td>Articles in the internet</td>
<td>Avg. 3.26 SD 1.395 N 155</td>
<td>Avg. 4.86 SD 1.387 N 71</td>
<td>0.0029*</td>
</tr>
<tr>
<td>Own knowledge and experience</td>
<td>Avg. 4.70 SD 1.320 N 158</td>
<td>Avg. 5.10 SD 1.109 N 72</td>
<td>0.0058*</td>
</tr>
</tbody>
</table>

**Decision making style in financial affairs**

<table>
<thead>
<tr>
<th>Decision making style</th>
<th>Predominantly personal service (PPS)</th>
<th>Predominantly e-service (PES)</th>
<th>Significance *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn to financial advisor for assistance (1 = never, 7 = always)</td>
<td>Avg. 4.93 SD 1.481 N 157</td>
<td>Avg. 3.78 SD 1.366 N 72</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Rely more in (1 = own deliberation, 7 = others’ advice) in financial decision making</td>
<td>Avg. 4.47 SD 1.534 N 158</td>
<td>Avg. 3.31 SD 1.469 N 72</td>
<td>&lt;.0001***</td>
</tr>
</tbody>
</table>

**Advice and marketing**

<table>
<thead>
<tr>
<th>Advice and marketing</th>
<th>Predominantly personal service (PPS)</th>
<th>Predominantly e-service (PES)</th>
<th>Significance *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective of marketing: Push and sales (1) vs. good of the customer (7)</td>
<td>Avg. 4.13 SD 1.433 N 158</td>
<td>Avg. 2.76 SD 1.409 N 72</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Marketing directed to individual products (1) vs. to total wealth concept (7)</td>
<td>Avg. 4.73 SD 1.096 N 157</td>
<td>Avg. 3.83 SD 1.424 N 71</td>
<td>&lt;.0001***</td>
</tr>
</tbody>
</table>

**Relationship strength**

<table>
<thead>
<tr>
<th>Relationship strength</th>
<th>Predominantly personal service (PPS)</th>
<th>Predominantly e-service (PES)</th>
<th>Significance *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention to increase service of the current service provider (1 = disagree, 7 = agree)</td>
<td>Avg. 4.42 SD 1.350 N 158</td>
<td>Avg. 3.37 SD 1.434 N 67</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Investments in finding a suitable financial services company (1 = disagree, 7 = agree)</td>
<td>Avg. 2.81 SD 1.515 N 158</td>
<td>Avg. 2.24 SD 1.169 N 67</td>
<td>0.0026*</td>
</tr>
<tr>
<td>Perception of wealth service prices (1 = very low, 7 = very high)</td>
<td>Avg. 4.81 SD 1.096 N 156</td>
<td>Avg. 5.24 SD 1.177 N 71</td>
<td>0.0077*</td>
</tr>
</tbody>
</table>
Reconceptualising the Information System as a Service

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0239.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Benchmarking, Formative constructs, Customer satisfaction / service, IT/IS management</td>
</tr>
</tbody>
</table>
RECONCEPTUALISING THE INFORMATION SYSTEM AS A SERVICE
(RESEARCH IN PROGRESS)

Gable, Guy, Queensland University of Technology, Brisbane, Australia  g.gable@qut.edu.au
Rai, Arun, Georgia State University, Atlanta, USA  arun.rai@ceprin.org

Abstract

The study will cross-fertilise Information Systems (IS) and Services Marketing ideas through reconceptualising the information system as a service (ISaaS). The study addresses known limitations of arguably the two most significant dependent variables in these disciplines - Information System Success or IS-Impact, and Service Quality. Planned efforts to synthesise analogous conceptions across these disciplines, are expected to force a deeper theoretical understanding of the broad notions of success, quality, value and satisfaction and their interrelations. The aims of this research are to: (1) yield a conceptually superior and more extensively validated IS success measurement model, and (2) develop and operationalise a more rigorously validated Service Quality measurement model, while extending the ‘service’ notion to ‘operational computer-based information systems in organisations’. In the development of the new models the study will address contemporary validation issues.

Keywords: ISaaS, IS as a Service, IS-Impact, IS success, Information system success, Service quality, IS service quality, Formative construct validation.
1 INTRODUCTION

The study aims to make parallel contributions to the two fields of Information Systems and Services Marketing. The study proceeds from a central interest in the importance of evaluating information systems (IS) in organisations, and adopts the IS-Impact model of (Gable et al. 2008) as the primary commencing theory-base. Akin to analytic theory1 (Gregor, 2006), IS-Impact is conceptualised as a formative, multidimensional index, wherein the dimensions have a causal relationship with the overarching measure – IS-Impact.

IS have evolved to be intertwined with organisational processes, users increasingly interacting with the IS to co-produce the service and the IS accommodating heterogeneity in process and information needs (Rai, Lang, & Welker, 2002). Yet, research to date measuring IS success (including IS-Impact) has, typically, implicitly regarded the IS as a ‘product’, thereby possibly underestimating its service value to stakeholders. In parallel with continuing validation and generalisation of the ‘IS-Impact’ model (Gable et al., 2008), there is value in reconceptualising the IS ‘as a service’ (ISaaS) and ultimately triangulating measures of ‘ISaaS’ with IS-Impact. Important research questions are: ‘Does conceptualising the IS as a service yield a more complete and accurate estimate of its impact?’ and ‘Is the Service Quality of the IS already accounted for wholly or partially by existing IS-Impact dimensions and measures?’

From an Information Systems perspective, main intended outcomes are to: (1) explore implications for evaluating IS from conceptualising the IS as a service; (2) bring conceptual clarity to the Information Systems discipline regarding the alternative notion of Service Quality as a ‘dimension’ of Information Systems Success (SVQD) as suggested by DeLone and McLean (2003); and (3) further validate the IS-Impact model of Gable et al. (2008). From a Services Marketing perspective, main intended outcomes of the study are to: (1) develop, and operationalise a more rigorously, empirically validated Service Quality measurement model; and (2) extend the ‘service’ notion to ‘operational information systems in organisations’ (IS).

1.1 Significance for practitioners

The Information and Communications Technology (ICT) market in Australia alone is worth $38B, 3rd largest in Asia-Pacific (10th in the world). This spend is by both private and public sector, and by both large and small organisations (SME spend in Australia in 2003 was $12B). Complex and expensive IT innovations (e.g. Enterprise Systems) are transforming organisations and industries, but not always for the better. Yet, IT investments are seldom systematically evaluated post-implementation, and where assessed, the process and measures are typically idiosyncratic and lacking credibility. Reliable, valid and comparable indicators of the impact of IS are required for organisations to know how their IT investment is performing, to maximise benefits, and to better plan future IT investments.

1.2 Significance for research

The study addresses known limitations of what may be the two most important dependent variables in

---

1 The first of Gregor’s (2006) five types of theory in IS, analytic theories, “analyse ‘what is’ as opposed to explaining causality or attempting predictive generalizations … they describe or classify specific dimensions or characteristics of individuals, groups, situations or events by summarizing the commonalities found in discrete observations” (2006, p.612).

2 The concept of IS-as-a-service (ISaaS), coined here for the first time, should not to be confused with the notion of ‘Software as a Service’ (SaaS), which refers to a means of software deployment whereby payment is typically on a usage basis. Neither should ISaaS be confused with the ‘Service Quality of the IS function’ (sometimes ambiguously referred to as ‘IS Service Quality’), most notably introduced in 2003 by DeLone & McLean (2003) to their IS success model.
Information Systems and Services Marketing research, namely Information System Success or IS-Impact, and Service Quality. The study is novel in aiming to cross-fertilise learnings of these two disciplines; efforts to harmonise conceptions of IS-Impact and Service Quality being expected to force a deeper understanding of the broad notions of success, quality and satisfaction and their interrelations.

2 LITERATURE REVIEW

2.1 IS success measurement

While the assessment of IS Success is consistently reported by organizational executives throughout the world as a key issue (e.g., Irani and Love, 2000, Thatcher and Oliver, 2001), there is little consensus among practitioners or researchers on how best to measure the impact of IS in organizations. (Sabherwal et al. 2006:1849) observe, “Despite considerable empirical research, results on the relationships among constructs related to information systems (IS) success, as well as the determinants of IS Success, are often inconsistent.”

The (DeLone and McLean, 1992) IS Success model is most widely cited, and includes six constructs: System Quality, Information Quality, Satisfaction, Use, Individual Impact, and Organisational Impact. Nonetheless, a range of concerns have been suggested with past attempts to validate that model, including: poor measurement (e.g., incomplete or inappropriate measures) (DeLone and McLean, 1992, DeLone and McLean, 2002, DeLone and McLean, 2003, Gable, 1996, Melone, 1990), lack of theoretical grounding and, hence; lack of agreement on appropriate measures (Bonner, 1995, Myers et al., 1998); myopic focus on financial performance indicators (Ballantine et al., 1996, Kaplan and Norton, 1996); weaknesses in survey instruments employed (e.g., constructs lacking in validity); or inappropriate data collection approach (e.g., asking the wrong people, unrepresentative sample) (Seddon et al., 1999). Moreover, the lack of consensus on such a central dependent variable, compromises the comparability of study results and hinders the cumulative research tradition.

Gable et al. (2008) introduce the IS-Impact model (see also Gable et al. 2003, Sedera & Gable 2004), which, based in DeLone and McLean’s work, overcomes many concerns with past IS Success models (see Figure 1). In attention to proliferation of overlapping measures, (Gable et al. 2008) comprehensively evaluated existing items, resolving redundancy and identifying new measures for contemporary IS. Their model reconciles persistent confusion regarding the role of the DeLone and McLean constructs as measures versus explanandum, conceptually demonstrating their value as both. Their analysis represents the first test of the sufficiency and necessity (or not) of the six DeLone and McLean constructs; they ultimately evidence the sufficiency and necessity of the four IS-Impact dimensions. They argue the redundancy of Use, and consistent with contemporary views in Information Systems and other disciplines, they present a strong rationale for conceiving Satisfaction as a consequence of success (and antecedent) rather than a dimension (Figure 1).

User satisfaction has been possibly the most extensively employed single measure for IS evaluation [DeLone and McLean, 1992, Doll and Torkzadeh, 1988, Etezadi-Amoli and Farhoomand, 1991, Gatian, 1994, Igbaria and Tan, 1997, Lucas, 1975]. Several widely cited studies developed standard instruments that measure satisfaction [Bailey and Pearson, 1983, Baroudi and Orlikowski, 1988, Doll and Torkzadeh, 1988]. Early satisfaction constructs in IS success evaluation (e.g., user information satisfaction—Bailey and Pearson 1983) have been found to mix measures of multiple success constructs (e.g. quality and impact) rather than measuring a distinct satisfaction construct [Gable 1996]. Rai et al (2002), state that user satisfaction has been measured indirectly through Information-Quality, System-Quality and other variables in prior studies. Additionally, [Sedera and Tan, 2005] demonstrated – through content analysis of 192 satisfaction-related items from 16 Satisfaction instruments – that 98% (189) of the measures readily map into existing measures pertaining to: System-Quality, Information-Quality, Individual-Impact and Organizational-Impact; with only 2% of the items (3 items) appearing to measure Satisfaction explicitly.

The conception of Satisfaction as immediate consequence of IS-Impact too has support in the Marketing discipline. Services marketing researchers e.g. [Brady et al 2005; Anderson and Sullivan 1993; Spreng and MacKoy 1996] employ a nomological
Gable et al. (2008) define the IS-impact of an Information System (IS) as “a measure at a point in time, of the stream of net benefits from the IS, to date and anticipated, as perceived by all key-user-groups”. The four-dimensional IS-Impact measurement model consists of two halves; the “impact” half includes Organizational-Impact and Individual-Impact dimensions; the quality half includes System-Quality and Information-Quality dimensions. The IS-Impact model, by design, is intended to be robust, simple and generalizable, to yield results that are comparable across time, stakeholders, different systems and system contexts. The model and measurement approach employ perceptual measures and offer an instrument that is relevant to all key stakeholder groups, thereby enabling the combination or comparison of stakeholder perceptions. Such a validated and widely accepted IS-Impact measurement model has both academic and practical value.

![IS-Impact Model](image)

*Figure 1 – The IS-Impact Model as mapped into IS-Net* (from Gable et al. 2008:395)

### 2.2 Service quality measurement

Practitioners and academics have sought since the 1980s to better understand the quality of services (Brogowicz, Delene and Lyth 1990). The measurement of Service Quality has been of central interest to the Services Marketing discipline for several decades. Online services today are pervasive and service quality is recognised as a key contributor to national economies and an increasingly important competitive differentiator.

SERVQUAL (Parasuraman, Berry & Zeithaml 1988) a multi-item scale that decomposes the notion of service quality into five constructs: tangibles, reliability, responsiveness, assurance and empathy, has been the most widely employed and cited measurement model. (Parasuraman, Berry and Zeithaml 1988) define ‘service quality’ broadly as “a global overarching judgment or attitude relating to the overall excellence or superiority of a service;” nonetheless, the term has many interpretations. Though there has been considerable progress on how to measure service quality perceptions, little consensus has been achieved on what should be measured. There is agreement that service quality is multi-dimensional, but little agreement as to the nature or content of these dimensions (Brady and Cronin 2001). In example, within the banking sector, there exist multiple service quality models, each

---

net that positions Satisfaction as immediate consequence of Service Quality; Satisfaction being antecedent of Behavioural Intention.

5 The IS-Impact model is shown here as mapped into the (Benbasat and Zmud (2003) nomological net IS-Net. (Gable et al. 2008) reconcile DeLone and McLean, IS-Net and IS-Impact.
consisting of varying numbers and nature/content of dimensions. More recent work by Brady and Cronin (2001) has provided a new and integrated conceptualisation of service quality. Brady and Cronin argue convincingly that customers form service quality perceptions on the basis of their evaluations of three primary dimensions: outcome quality, interaction quality, and environmental quality; these three primary dimensions are composed of multiple sub-dimensions.

In the 1990s, in recognition of the expanding service role of the IS function, Information Systems researchers such as Pitt et al. (1995) proposed that SERVQUAL be adapted to provide information about user satisfaction with the IS function. Despite some criticism (e.g. van Dyke, Prybutok, & Kappelman, 1999) the IS-adapted SERVQUAL has been praised for its practical relevance (Jiang, Klein, & Carr, 2002) and continues to be used to evaluate technical support service interactions (e.g. Carr (2002)). More recently IS researchers have also adapted SERVQUAL for use in the evaluation of electronic service environments such as e-commerce web sites (e.g. van Iwaarden, van der Wiele, Ball, & Millen, 2003) and Internet banking sites (e.g. Jayawardhena, 2004).

Conceiving a product as a set of service characteristics is not entirely new. It was first suggested by Lancaster (1966) four decades ago, and several researchers have since applied Lancaster's ideas. Saviotti and Metcalf (1984) used the idea of a product as a set of service characteristics to develop a framework for the evaluation of technological phenomena such as cars, and Gallouj and Weinstein (1997) used it to interpret innovation processes.

2.3 Contemporary Construct Validation

Construct validation issues and concerns have generally been under-addressed in many fields of research endeavour, with the Information Systems and Marketing disciplines responding relatively recently. Burton-Jones and Straub (2006) focus on the importance of operationalisation in light of the specific research context, theory and hypotheses, highlighting concerns with over-reliance and inappropriate reliance on ‘omnibus’ constructs.

Recent work by Jarvis et al. (2003) and Petter et al. (2007) suggests extensive misspecification and validation of constructs as ‘reflective’ that are on closer scrutiny in fact ‘formative’. Reflective constructs have observed measures that are affected by an underlying latent, unobservable construct (MacCallum & Browne 1993), while formative constructs are a composite of multiple measures. A change in the reflective construct affects the underlying measures, while changes in the formative measures cause changes in the underlying formative construct. Because measurement error impacts the structural model, misspecification of constructs as formative or reflective increases the potential for both type I and type II errors. Though not explicitly acknowledged nor addressed in the Services Marketing literature, the SERVQUAL and B&C models are ‘formative’ (as is IS-Impact). Thus, while the B&C Model (Brady & Cronin 2001) represents the most rigorously validated existing Service Quality measurement model, validation of the B&C model has to date not accounted for its formative nature.

Petter et al. (2007) have cast doubt on the validity of many mainstream constructs employed in IS research over the past 3 decades. They critique the almost universal conceptualisation and validation of these constructs as reflective. They are politic in not citing specific infractions but, rather, they list a range of studies and example constructs that have been ‘properly’ specified as reflective or formative. It is noteworthy that no examples of the proper specification of either the Individual-Impact or Organisational-Impact dimensions are cited (recognising that their list is not intended to be comprehensive). Cited examples of the proper specification of other of the DeLone and McLean dimensions are few, particularly in light of their extensive employment in IS research (e.g. only 1 example each of System-Quality and Information-Quality, both from the same study (Wixom & Todd 2005)). The IS-Impact model (Gable et al. 2008), however, is formative and is validated as such.

Each of IS-Impact, SERVQUAL and B&C and whatever ‘new’ model that derives from the Identification-Phase, employs survey instruments and multi-item scales that aim to measure real world
phenomena or latent constructs. Allport and Kerler (2003, p.356) suggest that ‘measurement is perhaps the most difficult aspect of behavioural research’. Validation of such latent constructs typically employs the so-called classical test theory (CTT), however, other approaches are gaining traction. In example, Item Response Theory (IRT) Rasch (1993) employs statistical models to achieve the ‘objective’ measurement of latent traits. And though relatively few papers in leading IS journals report the use of the Rasch IRT model [exceptions include Dekleva and Drehmer (1997) and Alvarez et al. (2007)], several recent marketing papers have suggested advantages of this approach (e.g. Salzberger and Sinkovics 2006).

3 THE CONCEPTUAL STUDY MODEL

The initial study model, developed in the first phase of the study, is depicted in Figure 2. This conceptual model incorporates three sub-models, and evaluates the relative power of IS-Impact, IS-Impact+ (IS-Impact plus SvQD) and ISaaS to predict Satisfaction.

The 1st sub-model: IS-Impact (as reproduced top Figure 2), with a strong basis in work by DeLone and McLean (1992, 2003), includes 4 dimensions in two halves, representing ‘the stream of net benefits from an Information System to date and anticipated as perceived by all key-user-groups’ (Gable et al. 2008). The ‘impact’ half measures benefits to date, or Individual-Impact and Organizational-Impact. The ‘quality’ half, uses System-Quality and Information-Quality as proxy measures of probable future impacts.

The 2nd sub-model: IS-Impact+, is a variant on IS-Impact which includes SvQD as a 5th dimension. There is potential for confusion here between our broader notion of ISaaS and the DeLone and McLean (2003) notion of Service Quality as a ‘dimension’ of Information System Success (SvQD). DeLone and McLean (1992), the most widely cited IS success model, extended their model in 2003 by including Service Quality (SvQD) as additional to their original dimensions of system success. Citing Pitt et al. (1995) they emphasise the dual role of contemporary organisations as both information provider (producing an information product) and service provider (providing support for end user developers) and argued for the inclusion of ‘service quality of the IS function’ as a dimension when evaluating IS success. Various researchers have employed ‘service quality of the IS function’ as a dimension in evaluating system success (Pitt et al. 1995; Jiang et al. 2002; Laurn & Lin 2003; DeLone & McLean 2004). (Watson, Pitt and Kavan 1998) argue the need to identify actions at the strategic, tactical and operational levels that will yield maximum improvement in IS service quality.

The 3rd sub-model: ISaaS, derives from more recent work by Brady and Cronin (2001). Consistent with ideas espoused by (Dabholkar et al. 1996, 2000), it represents service quality perceptions as multilevel and multidimensional, the rationale being that customers tend to cognitively decompose service quality dimensions into various sub-dimensions (Carman 1990), and a hierarchical structure accounts for the complexity of human perceptions (Dabholkar et al. 1996). Brady and Cronin (2001) bring substantial clarity and harmony to the measurement of service quality, their model in some sense subsuming SERVQUAL. Their third order model (ISaaS in Figure 2) relates service quality perceptions to the three dimensions – Interaction quality, Environment quality, and Outcome quality; each of which has three sub-dimensions that define the basis of service quality perceptions. They further suggest that for each of these sub-dimensions to contribute to improved service quality perceptions, the quality must be perceived to be reliable, responsive and empathetic.

The study aims to exploit perceived potential from cross-fertilisation of ideas between the IS and Marketing fields, the hierarchical structures of the IS-Impact and B&C models being analogous in several ways. In example, (Gronroos 1982), as cited in (Brady & Cronin 2001, p.35), suggests two main service quality dimensions where ‘Functional quality represents how the service is delivered; that is, it defines customers’ perceptions of the interactions that take place during service delivery. Technical quality reflects the outcome of the service act, or what the customer receives in the service
encounter.’ With the ‘operational’ IS (the focus of IS-Impact – and the unit-of-analysis in this study), where the system itself is conceived as a stream of services or a systematised (automated) service, the system (and its quality) are the ‘functional’ and its impacts are the ‘technical’ (or outputs). Note that the key distinction made in the IS-Impact model, between its Quality and Impact halves, is also similar to Alter’s distinction (in Seddon et al. 1999) between ‘internal performance’ and ‘external performance’ which respectively refer to ‘how well the system operates internally’ versus ‘how well the system achieves it purpose’ (1999, p.48).

The IS-Impact Model


IS-Impact+

Details of the ISaaS sub-model in Figure 2, tentatively (thus the cloud) reflect the B&C model. The B&C model may ultimately be supplanted by a partially or substantively different model deriving from the exploratory Identification-Phase of the proposed study (see Figure 3). In the simplest scenario, the B&C model is adapted to the purpose of ISaaS. This may involve some variation of its dimensions or sub-dimensions (potentially the only variation being with the adaptation of its items). In example, an open issue is whether all nine B&C dimensions are relevant when conceptualising the IS as a service. Dependent on the operational definition of the IS, the B&C ‘Environment’ dimension – a) may not vary with different IS (‘I use different systems, but all from the same PC and desk’); b) may be optional (‘with my WiFi laptop I work from wherever I am’); or c) may more appropriately refer to the web environment (the look-and-feel of the virtual environment). In a more extreme scenario, inductive attention to quality-citations gathered in the exploratory Identification-Phase may suggest a radically different ISaaS model structure from that of the B&C model.

Finally, note that in the study model, the Satisfaction construct is: a) the immediate consequence of each of IS-Impact, IS-Impact+ and ISaaS; b) the only reflective construct in the study model (all others expected to be formative), this being a requirement for full identification of the formative IS-Impact, IS-Impact+ and ISaaS models (see Jarvis et al. 2003; Petter et al. 2007); and c) to be operationalised in full light of the model and study intent, as per Burton-Jones and Straub (2006).
4 APPROACH AND METHODOLOGY

4.1 The Study Design

To operationalise and validate the ISaaS model, the study employs a longitudinal, multi-method research design, extending the research cycle proposed by MacKenzie and House (1979) and McGrath (1979) for developing and validating a measurement model. The research design (Figure 3) entails two main phases and three surveys: (1) an exploratory-phase, to develop the hypothesised measurement model, and (2) a confirmatory-phase, to test the hypothesised measurement model against new data gathered. The exploratory phase adheres with the two-step approach of Burton-Jones and Straub (2006) for operationalising constructs and identifying measures, the related aim being to adequately account for the context of contemporary IS, and to ensure model completeness and an appropriate and complete choice of measures and dimensions.

The exploratory phase consists of two surveys, an identification-survey followed by a specification-survey. The identification-survey, akin to the ‘function’ phase of the Burton-Jones et al. (2006) approach, is intended to identify the salient dimensions and measures of the study ISaaS conceptual model (bottom sub-model in Figure 2); these dimensions and measures will later become the basis of an a-priori model to be operationalised in the specification-survey.

The study model will be tested empirically, primarily employing 1-7 point Likert survey data gathered longitudinally in relation to Enterprise Systems – Financials (the financial module of large enterprise Systems packaged software suites. E.g. SAP, ORACLE, Peoplesoft, …). The study will be restricted to Financials, they being ‘relatively’ simple and homogenous (across organisations); and, as intra-organisational systems, only internal stakeholders need be canvassed. In order to gain a holistic view, all key-user-groups (Gable et al. 2008) will be surveyed, namely, Strategic, Management, Operational and Technical. Target sources of data include organisations with which the team has a track-record; e.g. Queensland Government agencies (mainly SAP Financials) and Universities (various).

In response to concerns expressed by Salzberger and Sinkovics (2004) and others, with classical test theory (CTT), an extension of the proposed study will entail the operationalisation and validation of all study constructs, employing both CTT and Item Response Theory (IRT) approaches, and corresponding comparison and triangulation of results.

The Identification-survey aims to generate a comprehensive inventory of ‘qualities’ of services experienced. Respondents are prompted to describe (both closed and open questions), and anchor their responses in, the target Financial system. Part B of the instrument seeks brief, qualitative, textual statements on perceived ‘qualities’ of that system as a service. These statements are decomposed into their component quality-citations (content analysis), then mapped into candidate frameworks (target: minimum inter-coder reliability of 80%). To the extent that any one of the frameworks (e.g. SERVQUAL, IS-Impact, B&C …) is fully instantiated and fully accommodates the quality-citations, it may be considered complete (manifests content validity) and appropriate. Should none of the candidate frameworks adequately accommodate the quality-citations, a bottom-up and more grounded approach will be attempted, seeking to identify salient dimensions and sub-dimensions from the data. The main outcome of this stage is the ISaaS a-priori model.

The Specification-survey (the 2nd survey) aims to further specify and test the a-priori ISaaS model employing data gathered (primarily 7-point Likert scales) with an instrument that operationalises the dimensions, sub-dimensions and measures deriving from the identification-survey (possibly a variant of the B&C model; possibly something quite different). In order to establish internal validity for a formative index, we follow guidelines by Bagozzi (1994), Diamantopoulos and Winklofer (2001) and Spector (1992). We also follow Jarvis et al.’s (2003) procedures for achieving identification of formative indicators, aiming for construct identification through both measurement (employing reflective criterion measures) and structural relations (the path with Satisfaction).
The Confirmation-survey aims to further validate the study model and instrument deriving from the exploratory-phase, and to further illustrate the mutual exclusivity and additivity of the measures and dimensions in the Model using confirmatory data analysis techniques and new data. In this phase of the study, we have the further objective of parallel testing of CTT and IRT data measurement approaches. Thus, four main bodies of data are required: (1) ISaaS CTT data, (2) ISaaS IRT data, (3) IS-Impact+ (subsumes IS-Impact) CTT data, and (4) IS-Impact+ IRT data. The effort and complexity of operationalisation is not underestimated. And though the team has comparatively advantageous access to large organisations for evidence, careful consideration in detailed design of data collection will need to address potential respondent fatigue and CMV (e.g. through random sub-sampling, longitudinal sampling). To complete the research cycle proposed by Mackenzie et al. (1979), construct validation tests similar to the Specification-Survey are conducted on the Confirmation-Survey data.

5 PROGRESS TO DATE AND NEXT STEPS

A literature survey has been completed, examining relevant issues relating to IS Success and to Service Quality measurement in the Services Marketing discipline. An evaluation has been made of current challenges in relation to construct measurement and validation in research. This has incorporated a literature review supported by a conceptual analysis. A basis for synthesising concepts from IS Impact measurement and from contemporary Service Quality measurement has been established. An initial conceptual Service Quality model has been developed as the basis for creating the a-priori model to be tested and refined into a validated ISaaS model.

The study has been proposed to the Australian Research Council, and acknowledges the generous support of that effort from a panel of two experts who offer alternative perspectives on the study design and developments. (1) Professor Teck Hua Ho, William Halford Jr. Family Professor of Marketing, Berkeley, brings a strong quantitative marketing perspective and a track record of research into consumer satisfaction (e.g. Ho 2006). (2) Associate Professor Judy Drennan, Faculty of Business, Queensland University of Technology, has specialist expertise in Services Marketing and Service Quality measurement (e.g. Drennan et al. 2003). The authors are optimistic the study will continue with ARC support. The study is regardless in 2009 proceeding with the Identification-survey.

Proceedings ECIS 2009
References


Measurement of e-service quality: an empirical study on online travel service

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0449.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-Service, Empirical study, Quality, Service Quality (SERVQUAL)</td>
</tr>
</tbody>
</table>
MEASUREMENT OF E-SERVICE QUALITY: AN EMPIRICAL STUDY ON ONLINE TRAVEL SERVICE

Li, Hongxiu, Turku School of Economics, Turku Center for Computer Science, Joukahaisenkatu 3-5 B, 6th floor, 20520 Turku, Finland, hongxiu.li@tse.fi

Liu, Yong, IAMSR, Åbo Akademi University, Turku Center for Computer Science, 20520 Turku, Finland, yong.liu@abo.fi

Suomi, Reima, Turku School of Economics, Turku Center for Computer Science, Rehtorinpellonkatu 3, 20500 Turku, Finland, reima.suomi@tse.fi

Abstract: Information communication technologies (ICTs) have significantly revolutionized travel industry in the last decade. With an increasing number of travel companies participating in the Internet market, low price has become a minimum qualification to compete in the Internet market. As a result, e-service quality is becoming even more critical for companies to retain and attract customers in the digital age. This study focuses on e-service quality dimensions in the Internet market with an empirical study on online travel service. The purpose of this study is to develop a scale to evaluate e-service quality from the perspectives of both online companies and customers, which provides fresh insight into the dimensions of e-service quality. The results in this study indicate that trust from the perspective of customer and ease of use from the perspective of online company are the most critical and important facets in customers’ perception of online travel service quality, while reliability, system availability and responsiveness have influence on customer’s perception of online travel service quality as well, but the influence is not so strong as that of trust and ease of use. Online travel service companies should pay attention to the facets of reliability, system availability and responsiveness while focusing on the facets of ease of use and trust in order to improve their online travel service quality to customers.

Key words: E-service, Quality, SERQUAL, Online Travel Service.
1 INTRODUCTION

With the rapid development of information and communication technology, Internet and World Wide Web (WWW) have become important tools in business, which has a major impact on business world. Internet has significantly revolutionized travel industry in the last decade. In the early stage of Internet market, online travel service companies had competed with traditional travel service companies by providing travel service at lower prices. With an increasing number of travel companies participating in the Internet market, the initial price advantages in online travel service has nullified, and low price has become a minimum qualification to compete in the Internet market (Sohn and Tadisina 2008). Therefore, travel companies operating in the Internet market need to find something else to attract customers to their online travel service. Oliveria et al. (2002) state that electronic service (e-service) might be the key to long-term advantages in the digital times, and e-service quality is becoming even more critical for companies to retain and attract customers in the digital age (Oliveria et al. 2002).

Though e-service is very important in the Internet market, most online travel companies seem not to focus on their online service to customers. They do not recognize that the competition among online travel companies relies mainly on their online service, and they do not have the motivation to adopt some strategies to develop or improve their online travel service to customers. Even though some companies have realized the importance of online travel service to customers, they seem not to understand customer’s perception of online travel service and how customers assess their online travel service quality.

Currently, despite many studies concerning traditional service quality, relatively few studies have been conducted in the Internet market, and even less on online travel service quality. This study focuses on e-service quality dimensions in the Internet market with an empirical study on online travel service. The purpose of this study is to develop a scale to evaluate e-service quality, which provides fresh insight into the dimensions of e-service quality. Internet-based travel companies mainly provide intangible service to customers. Thus, in this study the meaning of e-service indicates any intangible product-related services over the Internet provided by Internet-based travel companies.

The remainder of this paper is organized as follows. Following the introduction, the second section provides a relative literature review on both service quality and e-service quality construct. The third section presents the research hypotheses and research model. The fourth section is on empirical data analysis, and the fifth section discusses the results and the findings of this study. Finally, the study concludes with a discussion of the limitation of the study, and the future research in this field is presented as well.

2 LITERATURE REVIEW

2.1 Service Quality

Traditional service quality refers to the quality of all non-Internet based customer interactions and experiences with companies (Parasuraman et al. 1988). Service quality is determined by the difference between expected service and perceived service from companies (Zeithaml 1998). Parasuraman et al. conduct empirical studies in different service industries to develop and refine the service quality instrument (SERQUAL) to assess companies’ service quality (Parasuraman et al. 1988, 1991, 2005). They aim at providing a generic instrument for measuring service quality across a broad range of service categories. The widely used SERQUAL instrument is composed of five dimensions (Parasuraman et al. 1988), which is based on the original ten dimensions of service quality put forward by Parasuraman et al. (Parasuraman et al. 1985). The five dimensions of SERQUAL are:

- **Tangibles:** The appearance of physical facilities, equipment, personnel and communication materials;
- **Reliability:** The ability to perform the promised service dependably and accurately;
Responsiveness: The willingness to help customer and provide prompt services;
Assurance: The knowledge and courtesy of employees and their ability to convey trust and confidence;
Empathy: Care and individualized attention provided to customers.

The five dimensions of SERVQUAL was initially used to measure service quality in various service companies, including banks, credit card companies, telephone companies and travel companies, and was found to be valid in the traditional service market. Some academic researchers applied the SERVQUAL scale to measure service quality in the context of e-service. However, SERVQUAL has been considered problematic and may not be quite appropriate for e-service quality evaluation. The reason is that e-service is different from traditional service with three aspects standing out: the absence of sales staff, the absence of traditional tangible element, and self-service of customers. In this light it is clear that the SERVQUAL is not suitable for measuring e-service quality, and it is meaningful to develop an instrument for measuring e-service quality.

2.2 Development of E-service Quality Measures

With the increase of e-service adoption in business field, the importance of measuring and monitoring e-service quality in the virtual world has been recognized, and e-service quality posits to be a hot topic in research field. Some academic researches have already been conducted to develop e-service quality measurement. From the existing literature, it is evident that these studies have been conducted mainly in the domains of online retailing service, online shopping website quality, and e-service quality. There has been limited attention on online service sector.

Much of the studies in e-service quality take a combination of traditional service quality dimensions and web interface quality dimensions as the starting point. Dabholkar (1996) conducts a research work on the dimensions of e-service quality focusing on website design, and he argues that 7 dimensions of e-service quality can be illustrated as the basic parameters in the judgement of e-service quality, including website design, reliability, delivery, ease of use, enjoyment and control (Dabholkar 1996). Yoo and Donthu (2001) develop a 4-dimension scale called SITEQUAL to measure online service quality of website, and the four dimensions are ease of use, aesthetic design, processing speed, and interactive responsiveness (Yoo and Donthu 2001). Cox and Dale (2001) set up 6 dimensions of online retailing service quality with the comparison of the traditional dimensions of service quality, and the six dimensions are website appearance, communication, accessibility, credibility, understanding and availability (Cox and Dale 2001). Wolfinbarger and Gilly (2002) develop an e-service quality scale which was initially titled COMQ and later was progressed to eTailQ with the following four dimensions: website design, reliability, security and customer service (Wolfinbarger and Gilly 2002, 2003). Lociacono et al. (2002) develop an e-service quality scale called WEBQUAL, which is composed of 12 dimensions (Lociacono et al. 2002).

There is growing recognition of different variability in the outcome of e-service quality studies in terms of the quality dimensions (Waite 2006; Kim et al. 2006). Recently research on e-service quality shows more different dimensions in e-service quality (Madu and Madu 2002, Surjada et al. 2003; Santos 2003; Yang et al. 2003, 2004; Field et al. 2004; Kim and Stoel 2004; Yang and Fang 2004; Long and McMellon 2004; Gounaris et al. 2005; Lee and Lin 2005; Kim et al. 2006; Fassnacht and Koese 2006; Cristobal et al. 2007). Madu and Madu (2002) develop a 15 dimensions scale of e-service quality based on better understanding of customer and providing services to meet the needs and expectations of customers (Madu and Madu 2002). Santos (2003) argues that both active and incubative dimensions are important in e-service quality, and both of the dimensions should be taken into account in e-service quality assessment. An 11 sub-dimensions scale is put forward based on the two dimensions of e-service quality (Santos 2003). Field et al. (2004) develop a process model for assessing and improving service quality by identifying e-service system entities and transactions between those entities and mapping key quality dimensions onto them (Field et al. 2004). Gounaris et al. (2005) argue that the dimensions of perceived e-service quality are influenced by different
antecedents (Gounaris et al. 2005). Yang (2002) identifies the differentiation among dimensions between online-purchaser and non-purchaser (Yang 2002). Yang and Fang (2004) further examine the differentiation of dimensions to online service satisfaction and dissatisfaction. They suggest that there are four salient quality dimensions leading to both satisfaction and dissatisfaction: responsiveness, reliability, ease of use and competence (Yang and Fang 2004). Zeithaml et al. (2000, 2002) and Parasuraman et al. (2005) carry out a study on Internet service quality based on their earlier research on service quality in the traditional distribution channels, and develop an E-S-QUAL scale based on the 7 dimensions proposed by Zeithaml (Zeithaml 2000, 2002; Parasuraman et al. 2005). The E-S-QUAL scale comprises 11 dimensions in e-service quality, and later Parasuraman et al. (2005) develop the E-S-QUAL into a seven dimensions scale (Parasuraman et al. 2005). Kim et al. (2006) extend the dimensions developed by Parasuraman et al. (2005) into a 9 dimensions scale in e-service quality in order to use them for content analysis and evaluation of websites in the apparel retailing sector (Kim et al. 2006). Sohn and Tadisina (2008) put forward a 6-dimension model for e-service quality assessment based on their empirical study in Internet-based financial institutions (Sohn and Tadisina 2008).

There are some other significant discussions related to e-service quality as well, for example, in the contexts of technology readiness, service experience, customer satisfaction and web site loyalty. Yen (2005) articulates that the importance of attributes of online customer’s satisfaction is dependent on technology readiness (Yen 2005). Research on the antecedents to e-service adoption also suggests that e-service experience has impact on customers’ perception and evaluation of e-service quality (Yang and Jun 2002; Rowley 2006). Cristbal et al. (2007) suggest that the perceived quality of a web site or the degree of customer’s satisfaction to a web site is especially relevant to customer’s loyalty to a web site, and propose a four dimensions scale of e-service quality based on customer’s satisfaction and web site loyalty (Cristbal et al. 2007).

3 RESEARCH MODEL AND HYPOTHESIS

Though studies on service quality and e-service quality have been conducted, and different scales have already been developed for measuring e-service quality, the existing research on e-service quality has been described as fragmented (Wolfinbarger and Gilly 2003). A comprehensive framework is needed to identify the dimensions of e-service quality. After combining and synthesising the existing construct of both service quality and e-service quality, a perceived e-service quality construct is proposed, which consists of the dimensions from both online companies’ and customers’ perspectives. The proposed e-service quality model comprises 9 dimensions: ease of use, website design, reliability, system availability, privacy, responsiveness and empathy from the perspective of online companies, and experience and trust from the perspective of customers (See Figure 1).

Ease of use

Ease of use is defined how easy it is for customers to use website. Website should be designed for customer’s ease of use, including searching, navigating and use. Ease of use is an important determinant in the incubative dimension of e-service quality. Ease of use has been highly rated in customer’s e-service quality measurement, and it has been noted by some researchers (Dobhokar 1996, Zeithaml et al. 2002, Yang 2001, Fassnacht and Koese 2006).

Website design

In the virtual environment of e-service, for customers website is the main access to online organizations and to a successful purchase process. The deficiency of website design can result in a negative impression of the website quality to the customers, and customers may exit the purchase process. Website is the starting point for customers to gain confidence. Website design can influences customers’ perceived image of company, and attract customers to conduct purchasing online easily with good navigation and useful information on the website. Website should provide appropriate information and multiple functions for customers.
Reliability
Reliability refers to the consistency of performance and dependability of companies (Parasuraman et al. 1985, 1988). According to some empirical studies, reliability is the most important dimension of e-service quality. In the virtual environment, it is vital to make customers to trust that the company is going to perform what it promises to do. Reliability can make customers recognize the consistency and credibility of the company as well.

System availability
System availability refers to the correct technical function of the website. In e-service, the system availability makes customers always accessible to the online service offered by online companies, which can help customers to have a good image of online companies. If customers can not use the online system when they need online service, they will switch to some other online companies.

Privacy
Privacy refers to the degree to which the website is safe and customer information is protected. This dimension holds an important position in e-service. Customers perceive significant risks in the virtual environment of e-service stemming from the possibility of improper use of their financial data and personal data.

Responsiveness
Responsiveness refers to effective handling of problems and returns via the Internet. In e-service, company’s prompt service to customers via the Internet can make customers feel more comfortable during purchasing and continue purchasing without interruption.

Empathy
Even though there is no direct human interaction in the virtual e-service process, some human contacts are involved in e-service, for example e-mail communication. Providing customer individual attention
shows empathy to customers. Response to customers should always be cognizant of customer’s needs and show understanding of customer’s needs. In the virtual environment of e-service, empathy is important in customer’s perception of the e-service quality without face-to-face encounter. Accordingly, it is hypothesized that:

\[ H1: \text{Ease of use positively relates to customer’s perception of e-service quality.} \]
\[ H2: \text{Website design positively relates to customer’s perception of e-service quality.} \]
\[ H3: \text{Reliability positively relates to customer’s perception of e-service quality.} \]
\[ H4: \text{System availability positively relates to customer’s perception of e-service quality.} \]
\[ H5: \text{Privacy positively relates to customer’s perception of e-service quality.} \]
\[ H6: \text{Responsiveness positively relates to customer’s perception of e-service quality.} \]
\[ H7: \text{Empathy positively relates to customer’s perception of e-service quality.} \]

**Experience**

Experience is related to customers’ previous e-service usage behaviour. Online experience is customers’ total impression about the online company resulting from customers’ exposure to a combination of virtual marketing tools. Customers’ online experience embraces elements like searching, browsing, finding, selecting, comparing and evaluating information as well as interacting and transacting with the online company (Constantinides 2004). Customers’ online experience can influence their future purchasing intentions, their attitude toward e-service and their satisfaction.

**Trust**

In the context of the Internet, trust toward online companies is often regarded as a key factor of e-commerce growth, online success and competitiveness (Gounaris et al. 2005). Trust in e-service is related to the buying and payment process, the reliability of the website, privacy and securities issues, order fulfilment, service delivery, after sales service and the reputation of the company. Customers’ trust to online companies is critical for online companies’ success. Accordingly, it is hypothesized that:

\[ H8: \text{Experience positively relates to customer’s perception of e-service quality.} \]
\[ H9: \text{Trust positively relates to customer’s perception of e-service quality.} \]

The definitions of the nine dimensions are illustrated in Table 1.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of use</td>
<td>Easy for customers to use the website.</td>
</tr>
<tr>
<td>Website design</td>
<td>The website interface should be well designed and visually appealing.</td>
</tr>
<tr>
<td>Reliability</td>
<td>The consistency of performance and dependability of the website.</td>
</tr>
<tr>
<td>System availability</td>
<td>The correct technical function of the website.</td>
</tr>
<tr>
<td>Privacy</td>
<td>The safety of the website and the protection of customer information.</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>Effective handling of problems and returns via the internet.</td>
</tr>
<tr>
<td>Empathy</td>
<td>Care and individualized attention provided to customers via electronic channels.</td>
</tr>
<tr>
<td>Experience</td>
<td>The impression about online companies resulting from customers’ previous e-service usage behaviour.</td>
</tr>
<tr>
<td>Trust</td>
<td>Confidence among customers by providing prompt and information rich service.</td>
</tr>
</tbody>
</table>

*Table 1. Definitions of e-service quality constructs*
4 RESEARCH METHODOLOGY

4.1 Data Sample

The empirical sample of this study was the customers of some online travel companies in China. The customers were asked to indicate the dimensions which influence their evaluation of online travel service quality. It is based on their previous experience of online travel service booking. The questionnaire was developed mainly based on the scales from previous researches. A five-point Likert-scale ranging from strongly disagree (1) to strongly agree (5) was used to measure each item. The items for measuring the nine dimensions of e-service quality are built based on the instruments developed in previous researches as discussed in the literature review. Some modification and reword in the scale has been conducted to meet the requirements of this study.

The questionnaire was distributed to the customers by mail. Totally 1500 questionnaires were mailed to potential respondents, and 503 of the 1500 individuals replied. Among the 503 responses, 50 of them indicated that they have little or no experience of online travel service booking from online travel companies. The remaining 453 responses are the basis of this study. The 30% response rate can be considered acceptable since generally the response rate for questionnaires in information systems domains is between 8 and 15%. The age profile of the respondents represents most age groups with the majority (58%) being in the 18 to 35 age range and 62.7% of the respondents are male. The sample is considered to represent the Internet users in China. The demographic information of the respondents is shown in Table 2.

<table>
<thead>
<tr>
<th>Demographic profile</th>
<th>Frequency</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>284</td>
<td>62.7</td>
</tr>
<tr>
<td>Female</td>
<td>169</td>
<td>37.3</td>
</tr>
<tr>
<td>Total</td>
<td>453</td>
<td>100.0</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-25</td>
<td>121</td>
<td>26.7</td>
</tr>
<tr>
<td>26-35</td>
<td>142</td>
<td>31.3</td>
</tr>
<tr>
<td>36-45</td>
<td>138</td>
<td>30.5</td>
</tr>
<tr>
<td>46-55</td>
<td>32</td>
<td>7.1</td>
</tr>
<tr>
<td>56-65</td>
<td>20</td>
<td>4.4</td>
</tr>
<tr>
<td>Total</td>
<td>453</td>
<td>100.0</td>
</tr>
<tr>
<td>Frequency of using Internet (hours per week)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 5 hours</td>
<td>105</td>
<td>23.2</td>
</tr>
<tr>
<td>5 to 10 hours</td>
<td>51</td>
<td>11.3</td>
</tr>
<tr>
<td>More than 10 hours</td>
<td>297</td>
<td>65.5</td>
</tr>
<tr>
<td>Total</td>
<td>453</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 2. Demographic Information of Participants

4.2 Measurement Validity

The data analysis was performed using Partial Least Squares (PLS), PLS-Graph Version 3.0, to obtain estimates for the measurement and structural parameters in our structural equation model (Chin et al. 2003). PLS has enjoyed the increasing popularity in recent years because of its ability to model latent construct under the conditions of non-normality and that the theoretical framework is not fully crystallized (Chin 1998).

Convergent validity indicates the extent to which the measures of a construct that are theoretical related are also related in reality. Convergent validity can be evaluated by inspecting the factor loadings of the measures on their respective constructs (Chin 1998; Hulland 1999; Tenenhaus et al. 2005), and the reliability of the measures can be assessed using composite reliability (CR) and average variance extracted (AVE). In this study most of the factor loading are satisfactory with the cut-off value above 0.7, except that the factor loading of three items are acceptable with the cut-off value
between 0.5 and 0.7 (Hair et al. 2006). In this study the values of composite reliability (CR) and average extracted variance (AVE) satisfy the threshold value of 0.7 and 0.5 respectively (See Table 3), which demonstrates good internal consistency and suggests good convergent validity and reliability of the measures in this study (Fornell and Larcker, 1981).

<table>
<thead>
<tr>
<th>Constructs and items</th>
<th>Loading</th>
<th>St. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of use (EOU)</td>
<td>α=0.789</td>
<td>CR= 0.869</td>
<td>AVE=0.693</td>
</tr>
<tr>
<td>EOU1 It is easy to look for travel information.</td>
<td>0.906</td>
<td>0.011</td>
<td>82.245</td>
</tr>
<tr>
<td>EOU2 It is easy to move around the website.</td>
<td>0.900</td>
<td>0.009</td>
<td>94.034</td>
</tr>
<tr>
<td>EOU3 It is easy to do what I want to do, for example searching information, making an order.</td>
<td>0.669</td>
<td>0.068</td>
<td>9.784</td>
</tr>
<tr>
<td>Website design (WD)</td>
<td>α=0.611</td>
<td>CR=0.826</td>
<td>AVE=0.707</td>
</tr>
<tr>
<td>WD1 The user interface has a well-organized appearance.</td>
<td>0.928</td>
<td>0.375</td>
<td>2.473</td>
</tr>
<tr>
<td>WD2 The pages load quickly.</td>
<td>0.743</td>
<td>0.339</td>
<td>2.189</td>
</tr>
<tr>
<td>Reliability (REL)</td>
<td>α= 0.774</td>
<td>CR=0.860</td>
<td>AVE=0.672</td>
</tr>
<tr>
<td>REL1 The delivery service is accurate.</td>
<td>0.817</td>
<td>0.048</td>
<td>17.029</td>
</tr>
<tr>
<td>REL2 The order service is complete.</td>
<td>0.804</td>
<td>0.055</td>
<td>14.608</td>
</tr>
<tr>
<td>REL3 Its online booking records are always accurate.</td>
<td>0.837</td>
<td>0.039</td>
<td>21.370</td>
</tr>
<tr>
<td>System availability (SA)</td>
<td>α= 0.742</td>
<td>CR=0.839</td>
<td>AVE=0.637</td>
</tr>
<tr>
<td>SA1 System is always available for business.</td>
<td>0.913</td>
<td>0.023</td>
<td>38.659</td>
</tr>
<tr>
<td>SA2 System does not crash.</td>
<td>0.740</td>
<td>0.040</td>
<td>18.222</td>
</tr>
<tr>
<td>SA3 System runs smoothly in the transaction process.</td>
<td>0.728</td>
<td>0.049</td>
<td>14.850</td>
</tr>
<tr>
<td>Privacy (PRI)</td>
<td>α= 0.718</td>
<td>CR=0.826</td>
<td>AVE=0.615</td>
</tr>
<tr>
<td>PRI1 It protects the information about customers’ online shopping behaviour.</td>
<td>0.770</td>
<td>0.052</td>
<td>14.744</td>
</tr>
<tr>
<td>PRI2 It does not share customer’s information with others.</td>
<td>0.887</td>
<td>0.038</td>
<td>23.193</td>
</tr>
<tr>
<td>PRI3 It protects customers’ credit card information.</td>
<td>0.683</td>
<td>0.077</td>
<td>8.932</td>
</tr>
<tr>
<td>Responsiveness (RES).</td>
<td>α=0.700</td>
<td>CR=0.777</td>
<td>AVE=0.550</td>
</tr>
<tr>
<td>RES1 Adequate responsive time.</td>
<td>0.535</td>
<td>0.128</td>
<td>4.177</td>
</tr>
<tr>
<td>RES2 Prompt service.</td>
<td>0.937</td>
<td>0.033</td>
<td>28.444</td>
</tr>
<tr>
<td>RES3 Timely response.</td>
<td>0.700</td>
<td>0.098</td>
<td>7.083</td>
</tr>
<tr>
<td>Empathy (EMP)</td>
<td>α=0.887</td>
<td>CR=0.947</td>
<td>AVE=0.898</td>
</tr>
<tr>
<td>EMP1 Address complaints friendly.</td>
<td>0.956</td>
<td>0.004</td>
<td>227.916</td>
</tr>
<tr>
<td>EMP2 Consistently courteous.</td>
<td>0.939</td>
<td>0.007</td>
<td>131.981</td>
</tr>
<tr>
<td>Experience (EXP)</td>
<td>α=0.711</td>
<td>CR=0.831</td>
<td>AVE=0.621</td>
</tr>
<tr>
<td>EXP1 I have arrived at most of my expectation in my last experience.</td>
<td>0.825</td>
<td>0.031</td>
<td>26.894</td>
</tr>
<tr>
<td>EXP2 My last online booking experience is positive.</td>
<td>0.771</td>
<td>0.031</td>
<td>24.680</td>
</tr>
<tr>
<td>EXP3 I have good impression of the online travel company I booked my travel service from last time.</td>
<td>0.767</td>
<td>0.029</td>
<td>26.144</td>
</tr>
<tr>
<td>Trust (TRU)</td>
<td>α=0.634</td>
<td>CR=0.843</td>
<td>AVE=0.730</td>
</tr>
<tr>
<td>TRU1 I believe that the travel service provider can keep their promise.</td>
<td>0.892</td>
<td>0.020</td>
<td>45.844</td>
</tr>
<tr>
<td>TRU2 I believe that the travel service provider is honest and competitive.</td>
<td>0.815</td>
<td>0.033</td>
<td>24.397</td>
</tr>
<tr>
<td>E-service quality (ESQ)</td>
<td>α= 0.836</td>
<td>CR=0.924</td>
<td>AVE=0.858</td>
</tr>
<tr>
<td>ESQ1 Based on my previous online booking experience, I feel the online travel service quality is good.</td>
<td>0.919</td>
<td>0.013</td>
<td>73.449</td>
</tr>
<tr>
<td>ESQ2 The online service quality is better than I expected.</td>
<td>0.935</td>
<td>0.008</td>
<td>115.965</td>
</tr>
</tbody>
</table>

Table 3. Psychometric properties of measures

Discriminant validity can be verified with the square root of the average variance extracted for each construct higher than any correlation between this construct and any other construct (Fornell and Larcker 1981). As shown in Table 4, each construct shares a greater variance with its own measures than with any other construct. This reveals that each construct is more closely related to its own measures than to those of other constructs, and discriminant validity is supported in this study (Fornell and Larcker 1981).
### Table 4. Correlations between constructs

<table>
<thead>
<tr>
<th></th>
<th>EOU</th>
<th>WD</th>
<th>REL</th>
<th>SA</th>
<th>PRI</th>
<th>RES</th>
<th>EMP</th>
<th>EXP</th>
<th>TRU</th>
<th>ESQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>EOU</td>
<td>0.832</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WD</td>
<td>-0.094</td>
<td>0.841</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REL</td>
<td>0.201</td>
<td>0.093</td>
<td>0.820</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SA</td>
<td>0.156</td>
<td>0.029</td>
<td>0.557</td>
<td>0.798</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRI</td>
<td>0.318</td>
<td>0.092</td>
<td>0.467</td>
<td>0.398</td>
<td>0.784</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RES</td>
<td>0.229</td>
<td>0.002</td>
<td>0.243</td>
<td>0.367</td>
<td>0.115</td>
<td>0.740</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMP</td>
<td>0.397</td>
<td>0.119</td>
<td>0.495</td>
<td>0.553</td>
<td>0.491</td>
<td>0.400</td>
<td>0.950</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXP</td>
<td>0.438</td>
<td>-0.169</td>
<td>0.060</td>
<td>0.221</td>
<td>0.195</td>
<td>0.277</td>
<td>0.299</td>
<td>0.788</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRU</td>
<td>0.303</td>
<td>0.002</td>
<td>0.114</td>
<td>0.317</td>
<td>0.209</td>
<td>0.129</td>
<td>0.415</td>
<td>0.323</td>
<td>0.854</td>
<td></td>
</tr>
<tr>
<td>ESQ</td>
<td>0.518</td>
<td>-0.020</td>
<td>0.283</td>
<td>0.364</td>
<td>0.274</td>
<td>0.267</td>
<td>0.414</td>
<td>0.369</td>
<td>0.514</td>
<td>0.927</td>
</tr>
</tbody>
</table>

Note: The bold items on the diagonal represent the square roots of the AVE, and off-diagonal elements are the correlation estimates.

### 4.3 Hypothesis Testing

In this study we use a bootstrapping procedure to test the effects and the statistical significance of the parameters in the structural model. The findings in this study provide significant support for some of the hypotheses proposed in this study. Ease of use, reliability, system availability, responsiveness and trust are supported to be positively related to customer’s perception of e-service quality, while website design, privacy, empathy and experience are not supported to be related to it. Ease of use is the most important factor in e-service quality evaluation ($\beta=0.345$, p<0.001). Trust is the second important variable ($\beta=0.342$, p<0.001). In addition, reliability, system availability and responsiveness significantly affect customer’s perception of e-service quality while website design, privacy, empathy and experience have no significant impacts on customer’s perception of e-service quality. The proposed research model explains 45.7% of e-service quality. The results are shown in Figure 2.

![Figure 2. Structural analysis of the e-service quality model](image-url)

*Note: ***: p-value<0.001, **: P-value<0.01, and *: p-value<0.05*
5 DISCUSSION

Customer perceived e-service quality is one of the critical determinants of the success of online business (Yang et al. 2004). Accordingly, there is a rise of research on the construct of e-service quality. Currently, most of the research is conceptual in nature or based on a few case studies. Though there are survey-based empirical studies examining the construct of e-service, most of them focus on the dimensions from the perspective of online companies, and few studies have been conducted to investigate the dimensions from the perspective of customers, for example customers’ attribute.

In order to fill the research gap, this study empirically examines the construct of e-service quality from the perspectives of both the online companies and customers in the context of online travel service. This study develops a comprehensive instrument to evaluate e-service quality. The results in this study indicate that there are five key dimensions of e-service quality, including ease of use, reliability, system availability and responsiveness from the perspective of online companies, and trust from the perspective of customers. The results of reliability and validity test of the scale demonstrate good psychometric properties of the scale.

The five dimensions have different influence on customers’ perception of e-service quality. Ease of use and trust are ranked as the most critical and important facets of e-service quality, and have the strongest influence on customer’s perception of e-service quality. In this regard, it is noteworthy that ease of use attribute relates to how easy it is for customers to use the website and trust is involved in the confidence among customers by providing prompt and information rich service. Thus, the results of the study suggest that there is a need for online travel companies to place extra emphasis on their website attribute and their service process pertaining to these two dimensions of ease of use and trust.

For online travel companies, providing e-service with good quality involves much more than creating an excellent website for customers. Online travel companies should design its websites to be as easy as possible for customers to use. Online travel companies still need to pay attention on the building of trust in customers, which can help to improve their e-service quality, which prompts the managers to develop right strategies to build trust in customers. Of course, offering true and accurate information on the website might be the first strategy to help online travel companies to build up good image in customers’ mind, and the accurate service delivery might help to improve the confidence of customers. The two strategies can be the most basic strategies to build trust in customers, and some other strategies are needed as well.

Though reliability, system availability and responsiveness have significant influence on customers’ perception of e-service quality, the influence is not so strong compared to the influence of ease of use and trust. The reason might lie in the fact that the service on the three dimensions in online travel industry in China has been developed into a mature period.

Surprisingly, the study results indicate that privacy does not have significant influence on customer’s perception of e-service quality. However, for online companies it is still necessary to assure customers through website design cues and communications to signal customers the privacy of their websites.

In summary, the results in this study support the following conclusions. First, in general comparing the influence to customer’s perception of e-service quality, the dimensions from perspective of the online companies is stronger than that from the perspective of customers. Secondly, trust and ease of use are the most critical and important facets in customers perception of e-service quality. Third, though reliability, system availability and responsiveness are not as important as ease of use and trust, they are still important facets of e-service quality, and online companies should pay attention to them as well while focusing on the facets of ease of use and trust in order to improve online travel service quality. Forth, though privacy has not found to be a significant facet in customer’s perception of online travel service quality, it is still necessary to for online travel companies to assure customers the privacy of their websites.
6 LIMITATIONS AND FUTURE RESEARCH

This study has offered some valuable insight into studies on e-service quality, which involves a number of limitations that need to be acknowledged. First, the empirical study was conducted only in China. It is recommended to replicate the study in different nations to get international sample. Second, the study tests how the experience and trust of customers influence customers’ perception of e-service quality, other aspects of customers, such as attitude to e-service, education, social presence, have not been considered. Thus, there is scope for further work on the other measurement of e-service quality from the perspective of customer.

References


Chin, W.W., Marcolin, B.L. and Newsted, P.R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic- mail emotion/adoption study. Information Systems Research, 14(2), 189-217.


THE EFFECTS OF REGULATORY PRESSURE ON INFORMATION SYSTEM ADOPTION SUCCESS: AN INSTITUTIONAL THEORY PERSPECTIVE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0349.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Adoption, Organisational Change, Institutional theory, Sarbanes-Oxley</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
THE EFFECTS OF REGULATORY PRESSURE ON INFORMATION SYSTEM ADOPTION SUCCESS: AN INSTITUTIONAL THEORY PERSPECTIVE

Krell, Katharina, The University of Queensland, UQ Business School, Brisbane, 4072 Queensland, Australia, k.krell@business.uq.edu.au
Matook, Sabine, The University of Queensland, UQ Business School, Brisbane, 4072 Queensland, Australia, s.matook@business.uq.edu.au
Rohde, Fiona, The University of Queensland, UQ Business School, Brisbane, 4072 Queensland, Australia, f.rohde@business.uq.edu.au

Abstract

In today’s economy, firms are affected by various government regulations that have implications for their information systems (IS). Regulatory pressure has become a major driver behind IS adoption. Nevertheless, it is currently unknown how regulatory pressure affects adoption success. Using an institutional theory lens, this research proposes that regulatory pressure affects some well established success factors of IS adoption. In particular, we hypothesise that regulatory pressure has a positive effect on top management support, project champion and formal project management; but a negative effect on project team competence. Hypotheses will be tested using a survey approach.

Keywords: Regulatory Pressure, Compliance, Adoption, Success.
1 INTRODUCTION

Experience from practice shows that the success rate of information system (IS) adoption is generally low, with estimations ranging from 16% (Lee, 2003) to 40% (Liang et al., 2007). Motivated by the need to increase the success rate, much research has been conducted into IS adoption (Wixom & Watson, 2001; Premkumar et al., 1997; Lam, 2005). Prior studies suggest that adoption success is influenced by a range of success factors such as top management support (Akkermans & van Helden, 2002; Wixom & Watson, 2001) and formal project management (Aladwani, 2002). Nevertheless, because success factors appear mainly as independent variables in prior studies, there is a lack of knowledge about parameters that influence these success factors. For example, it is not clear why top management support varies across projects (Akkermans & van Helden, 2002). Understanding what affects the success factors can provide a much-needed new perspective on adoption success.

This paper proposes regulatory pressure as one of the parameters that influence success factors of IS adoption. Regulatory pressure occurs when governmental agencies directly or indirectly force firms to change their IS. In today’s economy, firms are affected by various government regulations that have implications for their IS (Krell & Matook, 2007; Braganza & Franken, 2007). For example, many firms in Europe and the US have recently been affected by new auditing regulations based on the Sarbanes-Oxley Act (SOX) (Marnet, 2007). Numerous firms found that their current IS were insufficient to achieve SOX compliance, for instance because of low security standards or a lack of financial monitoring systems. Thus, firms needed to alter their IS to ensure compliance (Sipior & Ward, 2007). Indeed, compliance affects varies IS projects in firms – examples reach from changes to spreadsheet software to the adoption of ERP systems (Gartner, 2008a). Many firms have started to appoint information technology (IT) compliance managers (Gartner, 2008b), and it is estimated that firms spend up to 15% of their IS budgets on regulatory compliance (Gartner, 2006).

IS adoption projects that are affected by regulatory pressure are different from other adoption projects because they are influenced by requirements defined in government regulations. For example, government regulations influence the schedule of adoption projects. Most government regulations define compulsory due dates by which compliance must be achieved (Haworth & Pietron, 2006). As a result, firms are forced to bring forward adoption projects that were actually planned for a later time, or even initiate unplanned adoption projects (Garcia, 2004). Consequently, it is reasonable to assume that regulatory pressure should affect adoption success factors. For example, if a government regulation influences the adoption schedule, and imposes penalties if the adoption is not completed on time, a firm might apply formal project management to ensure that the schedule is adhered to. Understanding how regulatory pressure affects success factors is important, particularly in a time like now when many industry sectors face increasing levels of governmental regulation. Nevertheless, there are currently no studies investigating how regulatory influences affect success factors.

Our paper applies an institutional theory perspective to address this gap. Using DiMaggio’s and Powell’s (1983) work on institutional pressures, we propose that regulatory pressure affects adoption success factors established in prior research. In particular, we hypothesise that regulatory pressure has a positive effect on top management support, project champion and formal project management; but a negative effect on project team competence. The mostly positive effects of regulatory pressure might surprise practitioners, but can be explained through our lens of institutional theory.

This research is expected to contribute to the IS literature in two ways. First, it is an initial attempt to investigate the effects of regulatory pressure on adoption success and provides a new perspective on success. This new perspective is important because in the information age, more and more regulations require firms to make changes to IS. Further, the recent economic and political development in many countries suggests that in the near future, some industries like banking or stock trading will face higher levels of government regulation. Against this backdrop, understanding the effects of regulatory pressure on IS adoption success is useful to increase the currently low success rate of IS adoption. Second, this paper contributes to IS research because it identifies regulatory pressure as a parameter influencing success factors of IS adoption. Future research can use our study as a starting point to identify other parameters, analyse how and why success factors of IS adoption occur in firms, and

Proceedings ECIS 2009
examine interactions between parameters. This work will help firms to deliberately model the success factors and increase IS adoption success. The remainder of this paper is organised as follows. Next, we outline how we use institutional theory as a theoretical lens to study regulatory pressure. Then, the hypotheses are developed, and the methodology is described. Finally, we discuss the current stage of the research.

2 AN INSTITUTIONAL THEORY PERSPECTIVE ON REGULATORY PRESSURE

In contrast to other theories on macro organisational behaviour like the resource-based view (Teece et al., 1997; Barney, 2002) and transaction cost economics (Williamson, 1981), institutional theory assumes that changes in structures and behaviours of firms are less driven by the desire to increase efficiency or create competitive advantage, but rather by a need of legitimacy (Meyer & Rowan, 1977). According to institutional theory, firms constantly aim to maintain and increase legitimacy through complying with pressures that arise from their institutional environment (Mizruchi & Fein, 1999). Three types of pressure exist: mimetic pressure, normative pressure, and coercive pressure (Dimaggio & Powell, 1983). Mimetic pressure arises from uncertainty. In situations when firms possess insufficient knowledge to evaluate alternative behaviours, the mere fact that another institution pursues a particular behaviour increases the legitimacy of this behaviour and hence, the firm mimics this behaviour (Haveman, 1993). Normative pressure, by contrast, results from norms defined by institutions such as professional or industrial associations. Once a firm has internalised a norm and decision makers identify with the norm, behaviours that comply with the norm legitimate the firm (Palmer et al., 1993). Coercive pressure, finally, stems from institutions in a firm’s environment that are powerful enough to reward or sanction a firm’s behaviour, for example large customers, suppliers of scarce resources, or governmental agencies (Guler et al., 2002). Complying with requests from these institutions enables the firm to benefit from rewards and avoid negative sanctions (Dimaggio & Powell, 1983). The focus on legitimacy makes institutional theory an ideal theoretical lens to study regulatory pressure because the main motive behind regulatory compliance is legal legitimacy (Liang et al., 2007).

Regulatory pressure is a special form of coercive pressure (Hu et al., 2007). It arises exclusively from the requirements forced upon a firm by governmental agencies (Braganza & Franken, 2007). There are two reasons why regulatory pressure tends to affect firms stronger than other institutional pressures. First, the regulatory environment in most countries is constantly changing due to ongoing changes in the national and international political development (Damianides, 2005). In contrast to other pressures, one particular change in government regulations affects a tremendous number of firms in a market or country at the same time (Haworth & Pietron, 2006). As a result, regulatory pressure affects firms more frequently than other pressures. Second, it lies in the nature of governmental agencies that they are powerful and impose strong negative sanctions on non-complying firms. For example, the retail multinational TJX failed to comply with US data security regulations and was convicted by a US court to pay compensations of several million US dollars after customer’s credit card details were stolen from the TJX customer relationship management system (DataBreaches, 2008). Firms are highly motivated to comply with regulatory pressures because they aim to avoid the negative sanctions which are associated with non-compliance (Abrahami, 2005).

In the information age, regulatory compliance often requires changes to IS. In most cases, firms do not initiate adoption projects solely for the purpose of compliance. Rather, firms review previously planned adoption projects and make changes to these projects to ensure that compliance is achieved (Hu et al., 2007). Think, for example, of a firm that needs to adopt a financial monitoring system to comply with a regulation. If this firm had planned to adopt a new accounting system in the next time anyway, managers might decide to modify the accounting systems to ensure that the system can be used for financial monitoring as required for compliance. Only if this is not possible the firm will consider initiating a completely new adoption project (Ghandforoush et al., 1999). Regulatory pressure is mostly not the sole driver behind adoption projects, but one of many drivers; and the strength of regulatory pressure as an adoption driver varies across adoption projects.
3 EFFECTS OF REGULATORY PRESSURE ON ADOPTION SUCCESS

For the purpose of this research, IS adoption is defined as the process during which a firm becomes capable of using an IS (Iacovou et al., 1995). Adoption includes, for example, the installation of technical system components, and the revision of firm processes. Adoption is completed when all technical components have been implemented, processes have been revised, and the firm possesses the necessary knowledge to use the IS. The actual usage of the IS is not part of the adoption, rather, it happens directly after the adoption (Iacovou et al., 1995). Iacovou’s definition was selected for this research because it enables the identification of a point of time when the adoption process is completed, and adoption success can be measured.

Commonly, adoption is considered to be successful if the new IS is implemented within budget, if all critical deadlines are met, and if all pre-defined technical requirements are implemented (Wixom & Watson, 2001). Success according to these criteria has been labelled differently in the literature and is referred to as “project implementation success” in this paper. Although prior research agrees that project implementation success is vital for adoption success as a whole, it is possible that an adoption is unsuccessful even if budget, schedule, and technical specifications are completely adhered to (Debrabander & Edstroem, 1977). This can happen, for example, if the pre-defined technical requirements are insufficient to support the firm’s business processes. In this case, managers might consider the adoption unsuccessful because they feel that the usability of the adopted system is low.

To account for this possibility, we build on Hong and Kim’s work on organisational fit (Hong & Kim, 2002) to expand on our definition of adoption success. Organisational fit is defined as the congruence between the IS and the business context in which it is adopted. The business context includes the business processes that an IS facilitates or supports. Many practitioners argue that IS adoption can only be successful if the IS is adequate for a firm’s processes (Gattiker & Goodhue, 2004) and hence, organisational fit is an important component of adoption success. For the purpose of this paper, it is therefore defined that adoption success is comprised of two components: project implementation success and organisational fit success, i.e. the level of organisational fit achieved during the adoption project.

Adoption projects that are initiated because of regulatory pressure differ from other adoption projects. The timing of the project and the technical characteristics of the adopted system are determined externally because they are strongly affected by a government regulation (Braganza & Franken, 2007; Haworth & Pietron, 2006). The stronger regulatory pressure becomes as a driver, the stronger will the adoption project be determined by these externally defined characteristics. By nature, government regulations are composed without special attention to the situation of a particular firm. Hence, timing and technical characteristics are affected by decisions made without consideration of a firm’s strategy, technical infrastructure, or usual IS adoption procedures. For example, a firm might normally go live with new ISs at a time of the year when key IT staff can be expected to be present. However, because of a particular law, the firm could be forced to go live on January 1st, which might conflict with the vacation roster. If the firm has other reasons than compliance to go live with this system on January 1, and compliance is a weak driver behind a particular adoption project, it can be expected that measures will be taken long in advance to ensure that key staff is available. Hence, the government regulation does not require much additional effort. However, if regulatory pressure is a strong driver behind the adoption, and there are no other reasons why the firm would go live on January 1, the additional effort due to the government regulation is tremendous, and ensuring that all key staff are available is a challenge. In this paper, we argue that challenges like this affect some of the adoption success factors identified in prior research. Next, we hypothesise which success factors we propose will be affected.
3.1 The Effect of Regulatory Pressure on Top Management Support

Many studies have demonstrated that adoption success is affected by top management support for an adoption project (Caldeira & Ward, 2003; Wixom & Watson, 2001; Akkermans & van Helden, 2002). Institutional theory suggests that top management support is affected by regulatory pressure. According to institutional theory, institutional pressures affect a firm if powerful agents within the firm recognize the pressure, and perceive that compliance is important and will increase legitimacy (Dimaggio & Powell, 1983). Hence, regulatory pressure will only affect macro organisational behaviour if top executives are aware of a change in government regulations, and of rewards and sanctions associated with compliance and non-compliance (Haworth & Pietron, 2006). Because non-compliance is associated with strong negative sanctions, compliance is considered a top management issue (Garcia, 2004). Even though related decisions can be delegated to a lower management level, top management supervises compliance-related actions and has a strong interest in the successful completion of these activities (Braganza & Franken, 2007). Thus, if regulatory pressure is a strong driver behind an IS adoption project, and the completion of the project is important for compliance, top management will be strongly interested in the adoption project. Hence, top management will strongly support the project. The stronger the regulatory pressure behind an IS adoption, the more will top managers be interested in the project and support the project. Hence, it is hypothesised:

**H1:** The strength of regulatory pressure has a positive effect on top management support.

3.2 The Effect of Regulatory Pressure on Project Champion Commitment

A project champion is a person who actively promotes his vision of an adoption project in a firm with the goal to ensure project success. The level of champion commitment varies across adoption projects. While in some cases project champions are formally appointed but do not engage in any particular activities to promote the project, other project champions frequently and vigorously communicate the benefits of the project (Orlikowski et al., 1995; Lai, 1997). Prior studies have shown that adoption is more likely if a project champion strongly commits to an adoption project (Orlikowski et al., 1995; Akkermans & van Helden, 2002; Lai, 1997; Wixom & Watson, 2001; Premkumar & Ramamurthy, 1995).

The project champion is either a senior manager himself, or he is appointed and supervised by a senior manager (Teo et al., 2003). As discussed before, institutional theory suggests that if regulatory pressure is a strong driver behind an adoption project, senior managers will be highly aware of sanctions associated with compliance, and will commit to avoiding sanctions. Therefore, it can be expected that the project champion will also be committed to avoiding sanctions, either because of his position in the firm, or through the supervision of the manager who appoints him. Consequently, the champion will be highly motivated to engage in actions to promote the project and hence ensure adoption success. Therefore, it is hypothesised

**H2:** The strength of regulatory pressure has a positive affect on project champion commitment.

3.3 The Effects of Regulatory Pressure on Formal Project Management

Formal project management is defined as the degree to which formal methods are used to plan, organise, and monitor an adoption project (PMI, 2004). Varies studies showed a positive relation between formal project management and adoption success (Akkermans & van Helden, 2002; Umble et al., 2003; Aladwani, 2002). In the context of regulatory compliance, formal project management is important because as discussed before, adoption projects are affected by externally determined requirements defined through government regulations, for example requirements that refer to the adoption schedule or to technical specifications. The stronger regulatory pressure becomes as an adoption driver, the more is the adoption project defined by externally determined requirements. Typically, such requirements occur within short periods of time and cannot be planned in advance (Garcia, 2004). Nevertheless, according to institutional theory, firms will attempt to comply with all governmental requirements even if the time frame is unusually short to ensure legal legitimacy.
Thus, firms are challenged to complete adoption projects within a short period of time even if these projects conflict with their usual practices. We argue that firms will opt for highly formal project management to meet such challenges. For example, in case the adoption schedule forced upon a firm through a regulation conflicts with the firm’s usual adoption practices, the firm might use a formal project staff plan to ensure that all required staff are available for the adoption. The more dominating regulatory pressure becomes as an adoption driver, the more challenges will arise, and the more will firms attempt to ensure adoption success through formal project management. Hence, it is hypothesised:

**H3:** *The strength of regulatory pressure has a positive effect on formal project management.*

### 3.4 The Effects of Regulatory Pressure on Project Team Competence

IT competence is defined as “the set of IT-related knowledge and experience” of a person (Basselier et al., 2003). Consequently, project team competence is the set of knowledge and experience of the project team related to a particular adoption project. The required set of knowledge and experience is different for each adoption project (Aladwani, 2002). In times of rapid technology changes, team members need to learn and redevelop their skills for each project (Stratman & Roth, 2002). Hence, firms often regroup project teams, form new teams for particular adoption projects, and retrain team members formally or on-the-job.

In cases when regulatory pressure is a strong driver behind IS adoption and technical requirements are largely determined by government regulations, it is particularly important to retrain team members to ensure that the team possesses the necessary competences to implement the technical requirements. Hence, it is necessary to define formal training measures, expose team members to similar adoption projects where they can get informal training “on the job”, or even hire new team members who possess the necessary qualifications (Walz et al., 1993; Newell et al., 2004). However, most measures to increase team competence require a certain period of time, at least several months, before they yield results (Biros et al., 2002). This is problematic because of the low predictability of changes in the regulatory environment (Garcia, 2004) and the externally defined due-date for compliance (Haworth & Pietron, 2006). According to institutional theory, achieving compliance at the required time will be more important for the firm than completing all competence-increasing measures because of the immense importance of legitimacy (Hu et al., 2007). Hence, it can be expected that in situations when regulatory pressure is a strong driver of IS adoption, the firm will have insufficient time to develop project team competence, and hence it is hypothesised

**H4:** *The strength of regulatory pressure has a negative effect on project team competence.*

### 3.5 The Effects of the Success Factors in the Context of Regulatory Compliance

As discussed before, top management support, project champion, formal project management, and team competence are well established success factors in the literature. An overview over prior research on these success factors is provided in Table 1.

<table>
<thead>
<tr>
<th>Success Factor</th>
<th>Supporting Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top management support</td>
<td>(Caldeira &amp; Ward, 2003; Wixom &amp; Watson, 2001; Akkermans &amp; van Helden, 2002)</td>
</tr>
<tr>
<td>Project champion</td>
<td>(Orlikowski et al., 1995; Akkermans &amp; van Helden, 2002; Lai, 1997; Wixom &amp; Watson, 2001; Premkumar &amp; Ramamurthy, 1995)</td>
</tr>
<tr>
<td>Formal project management</td>
<td>(Akkermans &amp; van Helden, 2002; Umble et al., 2003; Aladwani, 2002)</td>
</tr>
<tr>
<td>Project team competence</td>
<td>(Akkermans &amp; van Helden, 2002; Aladwani, 2002; Caldeira &amp; Ward, 2003; Cox et al., 1981; Dewar &amp; Dutton, 1986; Pennings &amp; Harianto, 1992; Slevin et al., 1991; Sharma &amp; Yetton, 2007)</td>
</tr>
</tbody>
</table>

*Table 1. Success factors of IS adoption.*
Due to the lack of prior studies on IS adoption success in the context of regulatory pressure we need to re-confirm that these factors are positively associated with adoption success in our special case. There are no hints in institutional theory why any of the success factors would be negatively associated with success, or not associated with success at all, when regulatory pressure is an adoption driver. Hence, in accordance with prior literature, it is hypothesised:

**H5a:** Top management support has a positive effect on adoption success.

**H5b:** Project champion commitment has a positive effect on adoption success.

**H5c:** Formal project management has a positive effect on adoption success.

**H5d:** Project team competence has a positive effect on adoption success.

### 3.6 Control Variables

While we believe we developed sound hypotheses and use a valid approach to test them, we still acknowledge that adoption success is affected by a range other variables that are not accounted for in our research model. We therefore introduced control variables in our research design which are summarised in Table 2. The research model with all hypotheses and controls is presented in Figure 1.

<table>
<thead>
<tr>
<th>Control Variable</th>
<th>Definition</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firm size</td>
<td>Size of a firm in terms of number of employees and sales in the last financial year</td>
<td>Firms of different sizes might be affected by different government regulations. Further, firms might struggle to successfully adopt IS due to a lack of resources.</td>
</tr>
<tr>
<td>Industry</td>
<td>Industry in which the firm primarily operates</td>
<td>Firms of different industries might be affected by different government regulations. Further, firms might use industry-specific systems that are easier or more difficult to adopt than systems used in other industries, hence success might vary across industries.</td>
</tr>
<tr>
<td>System complexity</td>
<td>Complexity of the adopted system in terms of technology and required integration with other systems</td>
<td>Complex systems might be more difficult to adopt successfully than less complex systems.</td>
</tr>
<tr>
<td>Project size</td>
<td>Size of the adoption project in terms of number of team members, duration of the project, and project budget</td>
<td>Large projects might exceed budgets and schedules more often than smaller projects because they are more difficult to plan.</td>
</tr>
<tr>
<td>Time elapsed since adoption</td>
<td>Time elapsed since users started to use the adopted system for their regular work tasks</td>
<td>The respondent’s recall of the adoption project might fade or change in the time after the adoption.</td>
</tr>
</tbody>
</table>

Table 2. Control variables.

![Figure 1. Research Model.](image-url)
4  METHODOLOGY AND CURRENT STAGE OF THE RESEARCH

This research applies a survey approach because it provides statistical power and is an appropriate basis for generalisation (Hair, 2006).

4.1  Research Instrument Development

The survey instrument was developed in a four-step procedure as follows. First, we conducted a literature review to find previous survey instruments for our constructs (Singleton & Straits, 2005). For each construct, several instruments existed in prior research.

Second, we built on prior survey instruments to design the first version of the survey instrument for this research. Because none of the selected instruments had been used in a regulatory pressure context before, we slightly adjusted the wording of some items to make them adequate for the purpose of this research. With the exemption of adoption success, all constructs were identified as reflective. Adoption success was modelled as a formative second-order construct because prior research showed that the two dimensions, project implementation success and organisational fit success, do not necessarily correlate (Hong & Kim, 2002).

In a third step, we evaluated the survey instrument using an academics panel and a practitioner panel. The academics panel comprised five experienced IS academics who were presented with a list of items for each construct. Using a 7-point scale, the academics were asked to evaluate how well each list of items represented the related construct. Further, the academics were also asked to provide comments on the items and constructs. The practitioner panel comprised four IT managers who represented our target participants. Again, participants were shown a list of items for each construct in our questionnaire. On a scale from 1 to 7, the participants indicated how easily answerable the items were. Further, the managers were also asked to provide comments. The average score of the constructs awarded by the academic panel was 5.8, and the average score awarded by the practitioner panel was 5.7. Given the mostly positive feedback from the academic panel, these scores were deemed acceptable. Some changes to the item wordings were made based on comments provided by the panels.

The fourth and final step was the pilot study. The research instrument was electronically distributed to IT managers in Australian firms. 30 usable responses were received. Construct reliability was evaluated using Cronbach’s Alpha. Generally, values above 0.7 are acceptable (Hwang, 2008). All calculated Alphas were above this threshold (see Table 3). Thus, no further changes were made to the research instrument. The final survey instrument is summarised in Table 3. All items are measured on a 7-point Likert scale.

<table>
<thead>
<tr>
<th>Construct</th>
<th># of Items</th>
<th>Adapted from</th>
<th>Cronbach’s Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strength of regulatory pressure</td>
<td>3</td>
<td>(Teo et al., 2003; Liang et al., 2007)</td>
<td>0.78</td>
</tr>
<tr>
<td>Top management support</td>
<td>5</td>
<td>(Teo &amp; Pian, 2003)</td>
<td>0.95</td>
</tr>
<tr>
<td>Project champion</td>
<td>4</td>
<td>(Teo et al., 2003; Wixom &amp; Watson, 2001)</td>
<td>0.97</td>
</tr>
<tr>
<td>Formal project management</td>
<td>6</td>
<td>(Henry et al., 2007; PMI, 2004)</td>
<td>0.92</td>
</tr>
<tr>
<td>Project team competence</td>
<td>7</td>
<td>(Stratman &amp; Roth, 2002; Aladwani, 2002)</td>
<td>0.90</td>
</tr>
<tr>
<td>Adoption Success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimension 1: Project implementation success</td>
<td>6</td>
<td>(Wixom &amp; Watson, 2001)</td>
<td>0.91</td>
</tr>
<tr>
<td>Dimension 2: Organisational fit success</td>
<td>7</td>
<td>(Hong &amp; Kim, 2002)</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Table 3.  Survey instrument (control variables omitted).
4.2 Data Collection and Analysis

Data will be collected through a survey amongst IT managers, IT directors and CIOs. This target group was chosen because prior research has shown that IT managers and IT executives are able to evaluate the success factors investigated in this research and the level of adoption success (Cragg et al., 2002; Harland et al., 2007; Wixom & Watson, 2001). Further, they are also aware of compliance-related IS decisions (Garcia, 2004) and thus, they are able to evaluate the strengths of regulatory pressure as a driver behind IS adoption.

Originally, we aimed to apply an online survey approach. However, an analysis of comments from companies that had been invited to participate in the pilot study and had replied that they felt unable to do so showed that in response to recent social engineering attacks, when IT managers were deceived in fraudulent surveys to reveal confidential IT information that would later be used for hacking attacks, many companies have implemented policies that restrained IT managers from participating in online surveys. Therefore, we will apply a mixed electronic and paper survey approach in the main study to increase the response rate. We aim at 200 usable responses.

Data will be analysed using structural equation modelling (SEM)/Partial Least Squares (PLS). PLS is a latent SEM technique that utilises a principal-component based approach to estimation. It can handle formative constructs (Chin & Newsted, 1999) and is robust to small sample sizes (Majchrzak et al., 2005).

5 CURRENT STAGE OF THE RESEARCH

Data was collected in March 2009. A total of 210 responses from eligible respondents were received. At the time of writing the camera-ready version of this paper, the SEM analysis is in progress.

References


A STUDY OF COMPLIANCE MANAGEMENT IN INFORMATION SYSTEMS RESEARCH

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0444.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
</tbody>
</table>
A STUDY OF COMPLIANCE MANAGEMENT IN INFORMATION SYSTEMS RESEARCH

Syed Abdullah, Syed Norris Hikmi, The University of Queensland, 4072 Brisbane, Australia, norris@utm.my

Indulska, Marta, The University of Queensland, 4072 Brisbane, Australia, m.indulska@business.uq.edu.au

Sadiq, Shazia, The University of Queensland, 4072 Brisbane, Australia, shazia@itee.uq.edu.au

Abstract

Regulatory compliance has become a critical concern for many industries around the globe and investment to achieve compliance has increased drastically inline with that concern. While Information Systems (IS) are considered a part of the support architecture, anecdotal evidence suggests that organisations struggle with finding the right tools and guidance on approaches for compliance management. For this reason, we undertake a review of the current research on compliance management topics in the Information Systems domain, with the ultimate goal to carry out a gap-analysis between research-based solutions and the current needs of compliance management professionals. In this paper, we consider thirteen Information Systems journals and perform an exhaustive analysis of the type of compliance management research published at these venues in the last five years. The analysis found forty-five relevant articles, which were then further classified depending on the type of their contribution. The results of the analysis suggest that IS research in managing compliance has received increasing attention in the recent years. The study also suggests that research has predominantly focussed on exploratory studies, rather than proposition of solutions that can assist organizations in their compliance management regimens.

Keywords: IS Journals, Regulation / Deregulation, Literature Review, Business Process Management
1 INTRODUCTION

Today, regulatory compliance has attracted much investment by organisations across the globe (Braganza & Franken 2007). The boost in compliance related investment is primarily a consequence of regulatory mandates that emerged as a result of events that led to some of the largest scandals in corporate history, such as Enron, WorldCom (USA), HIH (Australia) and Societe Generale (France). Compliance essentially means ensuring that business processes, operations and practice are in accordance with a prescribed and/or agreed set of norms. Introduction of regulations such as Sarbanes-Oxley Act of 2002 (SOX), Health Insurance Portability and Accountability Act of 1996 (HIPAA), Gramm-Leach-Bliley Act of 1999 (GLB) has made regulatory compliance a focal point of many organisations. Organisational units or department such as finance, administrative and information systems are affected by these changes. The investment is necessary for organisations to remain in business (Perskow 2003, Anon et al. 2007), since non-compliance to some government and legal requirements can have dire consequences.

Each new introduced regulation carries with it a potentially significant cost of implementation for those organizations that fall under its control. SOX spending has reached US$6B in 2007 in USA alone, which represents only a fraction of the total compliance effort (Reilly 2007). Meanwhile, analysts predict that overall USA spending on governance, risk and compliance (GRC) will exceed US$32B in 2008 (McGreevy 2008). Hence, increasing compliance expectations are a significant financial drain on organizations and are often considered as a burden rather than a business opportunity. The burden is magnified by the apparent lack of guidance from the research community on the best ways to approach compliance management. In this study, we define compliance management as mechanisms to keep enterprise’s businesses safe from possible violation of regulatory compliance. In his work, Kharbili (2008) stated that compliance management also refers to standards, frameworks, and software used to ensure the company’s observance of legal texts.

There is ongoing discussion on the roles of Information Systems (IS) as enabling technology that facilitates achieving and demonstrating compliance. Although some of the discussion focuses more on the use of IT as a supporting technology to achieve compliance, research also exists on management related issues, for example, how compliance affects CIO responsibilities in the organisation (Berghel 2005) and (Braganza & Franken 2007). While compliance management does not appear to have become a main stream of IS research at this point in time, the apparent lack of guidance for compliance management professionals motivates us to develop a comprehensive understanding of the focus, extent and shortcomings of compliance management research in the Information Systems community. The final goal is the derivation of an industry-relevant research agenda for compliance management. In order to develop such an agenda, two aspects are required, viz. an understanding of the current state of research, and an understanding of current problems exhibited in practice. In this paper we address the first aspect.

We proceed as follows. The next section presents a discussion of the methodology employed to assure a rigorous and relevant analysis of compliance management research. Section 3 presents an in-depth analysis of the relevant research. In section 4 we present a discussion of the results and conclude the paper in Section 5 with a discussion of limitations and future work in this area.

We expected that the majority of the research would be published since 2006, however we considered all papers published at these outlets in 2001-2008 so as to identify any changes in trends and also a change in foci of compliance management research (e.g. a shift of focus to SOX). In total, the data set consisted of 5633 articles. Each paper was prepared and included in a full text search for the purposes of identifying contributions relevant to the compliance management domain. Full text searches were conducted on the data set, using a keyword of “compliance” and “compliant”. Following a stage of eliminating paper duplicates in the search results (i.e. ensuring that a paper that matched both search criteria is only counted once). This analysis provided us with a set of 510 articles. As a further step to assess paper relevance to the domain of compliance management, we inspected the occurrence of the search terms in the paper text, and included only those that had three or more hits. This step reduced and focused the data set on 178 papers.

With the reduced set of papers, a review of each paper and its contribution was carried out. This review included reading the abstract, introduction, and scanning through the main contributions of the remainder of the paper as well as its conclusions. This stage was the most critical part of the study, since it provided an understanding of the contribution of the paper and what aspects of regulatory compliance were addressed (and in what way). This stage of the analysis was helped through a framework that was developed to analyse the contributions of compliance management research. The framework is presented in Figure 1. Within this framework, we differentiate between case study and exploratory papers, versus papers that provide a solution to a compliance management related problem. These identify the two main classes of research approaches. The solution papers were further classified in terms of their focus on preventative, detective or corrective measures. The papers were also classified by geographic region of application of the paper (North America, South America, Africa, Asia-Pacific, Europe), domain of application or industry sector (Financial, Healthcare, E-Commerce, etc) and type of compliance considered (to (1) Regulations or Legislation, (2) Standards and Code of Practice, (3) Business Contracts, Service Agreements etc. and (4) Corporate Policy). The latter classification is intended to provide research context.

This last stage of the analysis and the careful reading of each paper resulted in a further reduction of the data set. Despite the papers having more than three references to “compliance” or being “compliant”, many were determined not to present a main contribution to the domain of compliance management. Instead, they mentioned compliance in various parts of the discussion and future work, however did not focus on the topic, or the notion of compliance was significantly different, e.g.
compliance to a network protocol, or XML format, etc. Accordingly, the analysis reduced the set of papers from 178 to 45.

3 ANALYSIS

Table 1 shows the breakdown of papers relevant to compliance management and their source of publication. Out of 5633 articles, only 45 articles match the regulatory compliance context as explained above. This represents 1.3% of BPMJ articles, 2.4% of CAIS articles, 0.8% each for CACM and JI&M, 0.6% EJIS, 0.7% of JISR, 1.4% of JAIS, and 0.4% for MISQ. Five other journals (JIS, ISF, ISJ, IS and JMIS) did not contain any articles that match the compliance context. We expect that the significantly higher number of matches in the CAIS and CACM journals, may be influenced by the nature of the journal itself, which caters to exploratory type articles. At the early stage of the compliance management research in Information Systems, one would expect that there would be an increased focus on exploratory papers first, so as to understand the problems at hand, before a surge in papers presenting Information Systems solutions. Although the relatively low number of publications may be seen as not encouraging, the IS research community should not take it as an indicator of insignificance. Indeed, the roles of IS or IT as enablers of regulatory compliance have increased year by year (Smith & McKeen 2006).

<table>
<thead>
<tr>
<th>JOURNALS</th>
<th>TOTAL</th>
<th>Matched with Regulatory Compliance</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAIS</td>
<td>594</td>
<td>14</td>
<td>2.4</td>
</tr>
<tr>
<td>JAIS</td>
<td>147</td>
<td>2</td>
<td>1.4</td>
</tr>
<tr>
<td>BPMJ</td>
<td>320</td>
<td>4</td>
<td>1.3</td>
</tr>
<tr>
<td>JI&amp;M</td>
<td>491</td>
<td>4</td>
<td>0.8</td>
</tr>
<tr>
<td>CACM</td>
<td>2100</td>
<td>17</td>
<td>0.8</td>
</tr>
<tr>
<td>JISR</td>
<td>150</td>
<td>1</td>
<td>0.7</td>
</tr>
<tr>
<td>EJIS</td>
<td>350</td>
<td>2</td>
<td>0.6</td>
</tr>
<tr>
<td>MISQ</td>
<td>271</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>JIS-Sarasota</td>
<td>117</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ISF</td>
<td>195</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ISJ-Blackwell</td>
<td>185</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ISJ-Elsevier</td>
<td>385</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>JMIS</td>
<td>328</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TOTAL</td>
<td>5633</td>
<td>45</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 1. Sources and Frequency of Publication

The next step in the analysis carried out the classification with respect to the type of publication, viz. case study/exploratory and solution. As expected in an emerging research domain, the majority of the publications were found to be in the case study or exploratory paper category - 35 (76.1%) of the articles are case study/exploratory articles and nine (19.6%) are solution articles. However, there are two (4.3%) articles that matched both types of articles. The results suggest that research on regulatory
compliance solution has being initiated but remains still in the early exploratory stages, not yet progressing to a stage where many Information Systems solutions are proposed or discussed. This finding is inline with the need to identify the problems at hand first, before proposing solutions. Figure 2 presents the breakdown of the papers by type of contribution.

**Figure 2. Distribution of Articles by Type**

Furthermore, we were interested to determine the emergence of compliance management research in Information Systems publication outlets. The breakdown of compliance management per year of publication is shown in Figure 3. The figure clearly shows that a spike in publications was recorded in 2006. We posit that this finding is in line with the increased focus on SOX Act of 2002 and also an early focus on HIPAA. A characteristic of publishing in the Information Systems discipline is a lag of generally one to two years from time of writing to time of actual publication. Accordingly, this situation would imply that increasing focus began around 2002, when HIPAA was being seriously considered and the SOX Act was introduced. Prior to this event, little literature on compliance management exists, despite some other regulations having already been proposed (e.g. HIPAA). We expect the increasing trend to continue in 2008, despite the perceived drop in compliance management publications. One of the limitations of the study pertaining to the specific year of 2008 is that while some of the newly published contributions have already been indexed, many have not. Due to this restriction, Figure 3 only includes some of the early 2008 research.

**Figure 3. Distribution of Article per Type per Year**
Further analysis was done to identify the type of domain application that was the focus of the paper. Figure 4 shows the results of that analysis, which found that the application of regulatory compliance discussed in the articles was dominated by three domains, viz. auditing, finance and healthcare, which made up 76% of the articles. On the other hand, domains such as e-commerce and the environment were only discussed in 13% of the articles proportion. The remaining articles, which represent 11% of the focused data set, did not have a specific domain of application with regard to regulatory compliance.

![Figure 4. Articles by Domain of Applications](image)

In looking at the distribution of articles per region, we have been able to identify the main geographical areas of application as reported in the papers. The analysis was performed by inspecting not the country of publishing authors but rather the locality of case studies or exploration on regulatory compliance discussed in the articles. The analysis found that 22 out of 37 case study/exploratory articles represent the North America region. This proportion of almost 60% of the articles that focus on the North America region might be due to most of the regulatory compliance requirements being introduced earlier in the North America region than others. The finding indicates that there is a need to explore compliance management in different regions, in particular focusing on whether differences exist between practices in the various regions. The focus on other regions is not, however, nonexistent, with two and three articles representing Asia-Pacific and Europe regions respectively. The other 10 articles were categorised as non-specified because they applied to either all regions or did not explicitly state the region of the case study.

![Figure 5. Articles Population by Region](image)

The careful classification of the papers also resulted in the elucidation of details in terms of the regulations that are the focus of the Information Systems research community. The analysis found that 42 out of 46 articles had specified the regulations that the research was relevant to. While eight articles discussed the details on regulatory compliance type only in general without focusing to specific
regulation name (e.g. contract compliance in general), the 34 remaining articles specifically named a
single regulation in the discussion. Detail for the distribution of articles by their regulatory focus is
presented in Table 2.

<table>
<thead>
<tr>
<th>Regulation/Policy/Standard/Contract</th>
<th>No. of Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>8</td>
</tr>
<tr>
<td>HIPAA</td>
<td>8</td>
</tr>
<tr>
<td>SOX</td>
<td>6</td>
</tr>
<tr>
<td>ISO</td>
<td>6</td>
</tr>
<tr>
<td>SLA</td>
<td>2</td>
</tr>
<tr>
<td>CFIP</td>
<td>2</td>
</tr>
<tr>
<td>Others</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 2. Distribution of Regulatory Compliance Type by Name

Most prominently, HIPAA (Health Insurance Portability and Accountability Act of 1996) had the
highest perceived focus, being the focus of eight publications. SOX (Sarbanes-Oxley Act of 2002) and
ISO (ISO17799, ISO14000, ISO9000, and ISO9001:2000) followed with six articles each. Lesser
focus was placed on SLA (Service Level Agreement) and CFIP (Code of Fair Information Practice).
Nine papers mentioned a variety of regulations. These include Controlling the Assault of Non-
Solicited Pornography and Marketing Act of 2003 (CAN-SPAM), New Zealand Privacy Act,
Electronic Record Management Policy, European 1995 Personal Data Protection Directive,
Accreditation Board for Engineering and Technology / Computing Accreditation Commission
Standard (ABET/CAC), American Election Committee Voting System Standards (AEC), Eco-
Management Audit Scheme (EMAS), Capability Maturity Model Integration - Software
Engineering(CMMI-SW), and Fair Credit Reporting Act 1970.

We were further motivated to investigate whether clear specific foci exist with the main geographic
regions of interest. In the North America region, the focus is led by Auditing applications (with seven
articles). This focus was followed by Healthcare and Finance applications with six and seven articles
respectively. While Auditing remains as focus of the application in Asia-Pacific and Europe with one
article for each region, Healthcare applications did not attract any discussion from Asia-Pacific and
Europe regions.

Furthermore, we reviewed the same articles by discovering the type of regulatory compliance being
discussed in those articles with consideration of the region of application. The review shows that most
of North America region articles (17 articles) focus on regulatory compliance. This focus is followed
by a focus on compliance with standards as second popular type in North America (with 3 articles).
The foci for Asia-Pacific region include regulatory compliance and contract compliance, with one
article for each type. The Europe region also appears to focus on standard and regulation compliance,
with two and one articles respectively. Our analysis did not identify any contributions that specifically
address compliance with policies.

Further to the detailed investigation carried out on the case study/exploratory articles, the articles that
were classified as solution providing articles were also reviewed in terms of the type of solution
presented. This process involved identification of articles that referred to solution approaches. These
articles were classified as either preventive, detective, or corrective in nature. The study reveals that 7
out of 11 articles offer a preventive solution, while the remaining (four) articles offer a detective
solution. No corrective solutions we identified in the data set. The situation might be explained be a
preference in compliance management for preventive solutions, given the high penalties for non-
compliance with some regulations (not only in terms of costs but also in terms of reputational damage,
etc). The actual solutions presented in the papers vary from contribution to contribution. Some of the solutions, for example, (Agrawal et al. 2007) and (Weitzner et al. 2008) addressed the detective type of solution. In particular, (Agrawal et al. 2007) introduced the Hippocratic Database Compliance Auditing component, which facilitates the privacy officer to conduct a series of audits, in a matter of minutes, to reliably isolate potential sources of information leaks in electronic health records. (Banker et al. 2007), have discussed a simple model in which contract monitoring cost in addition to search and coordination cost is introduced to capture the complexity in buyer-supplier relationships. Another work, (Volonino 2003) presented an overview of e-evidence and computer forensics and their implication to IS. Thus there is significant diversity in the various solution papers.

4 DISCUSSION

The findings from the analysis have provided us with a number of interesting facts with regard to the current focus of IS journal articles associated with regulatory compliance. We summarize these as follows:

- Research Trends and Focus
  The trends of regulatory compliance IS research articles show significant growth starting in year 2003. The majority of these contributions feature exploratory research as compared to research that provides specific solutions. Our findings also reveal that the articles are dominated by work that discusses regulatory compliance associated with North America cases. This research is also linked to a number of regulations that have been introduced in the region i.e. HIPAA, GLB, COPPA, SOX, and CAN-SPAM Act.

- Domain of Application
  Auditing, Healthcare and Finance are the domains that attracted most concern in the articles. This implies the critical applications of compliance. However, since these domains are predominant in North America region the necessity of other domain of applications might emerge differently in other region.

- Type of Regulatory Compliance
  Perhaps because articles were dominated by North America region, regulation is the anticipated type of compliance and is shown to have attracted most discussion. In contrast, other types such as policy, contract and even standard did not get as much attention. This further show that most of the solutions discussed also focus on achieving compliance with regulations.

- Focus of the Solutions
  All solution associated with regulatory compliance discussed in the articles focus either on preventive or detective type of solutions. It may be hard in general for IS community to contribute to this aspect as corrective measures are an outcome of business/legal advice and strategy. However, it needs to note that it is not always possible to mitigate the risk even when a compliance product is involved and hence corrective measures become inevitable (Mercuri 2004).

- Regional
  As mentioned earlier, North America cases feature prominently in the discussion of regulatory compliance in IS research. Future work is clearly needed to study the impact of regulatory compliance in other regions.
This paper presents a first comprehensive snapshot of the development and focus of compliance management related research in the Information Systems discipline. The analysis is motivated by the recent changes that drastically change the IS/IT functions within organisations so as to achieve compliance. The articles were gathered from 13 IS Journals, which consist of articles from 2001 until early 2008. The sources for the analysis, which constitute of 5633 articles, were filtered through three stages of filtration that finally narrow it down to 45 articles. The filtered articles were then subjected to classification and analysis based on a proposed framework. The study reveals that the majority of the related IS publications are exploratory in nature. We posit this to be an indication of the relative immaturity of the research in this domain, where problems are still being identified rather than solutions being provided. While we concede that many solutions might perhaps be published in more technical venues (Sadiq et al. 2006), the role of Information Systems in supporting and demonstrating compliance should be exhibited with solutions papers in the IS domain also.

The work has a number of limitations. While all care was taken to download the full set of journal papers that represents the selected journals and years, this selection was limited to soft copy papers only. Papers available in print version (and not already digitised) were excluded from this analysis. Given the focus on recent years, we do not expect this limitation to have a significant impact on the data set. Additional limitations stem from the currently incomplete set of journals for the year 2008. Only a small number of papers published this year have been indexed, these papers will be included in the analysis as soon as they become available so as to present an up to date snapshot of compliance management research.

Future work in this area involves two major steps. The first is an extension of the IS journals to also IS conference venues, which have a shorter time to publication and would provide more detail on the development of maturity in the compliance management research area. Second, a series of interviews and focus groups has been initiated in order to collect the compliance management problems, as experienced by compliance management professionals, and as observed by compliance management consultants and auditors. The ultimate goal is to perform a gap-analysis between industry needs and the focus of the research community, resulting in an industry-relevant research agenda for the coming years.

References
ANALYSING THE EFFECT OF SECURITY ON INFORMATION QUALITY DIMENSIONS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0688.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Data quality, Information management, Quality, Cost-benefit analysis</td>
</tr>
</tbody>
</table>
ANALYSING THE EFFECT OF SECURITY ON INFORMATION QUALITY DIMENSIONS

Helfert, Markus, Dublin City University, Glasnevin, Dublin 7, Ireland, markus.helfert@computing.dcu.ie
Foley, Owen, Galway Mayo Institute of Technology, Dublin Road, Galway, Ireland, Owen.Foley@gmit.ie
Ge, Mouzhi, Dublin City University, Glasnevin, Dublin 7, Ireland, mge@computing.dcu.ie
Cappiello, Cinzia, Politecnico di Milano, Dipartimento di Elettronica e Informazione, Via Ponzio 34/5, 20133 Milano, Italy, cappiell@elet.polimi.it

Abstract

Information quality (IQ) has become a critical concern in today’s organisations. Although recent studies of information systems indicate an increasing importance of IQ, foremost research is still limited. Indeed, little is known about the impact of various design decisions on IQ. Recent research shows that security measures are increasingly important for any information system; however security measures are often introduced without considering the effect on IQ. At the same time, literature provides us with indications that trade-offs between various IQ dimensions exist. In this article we aim to investigate how security measures impact on different IQ dimensions. We carried out an experiment, which indicates that security measures have a significant effect on timeliness, whereas other dimensions are not particularly influenced. This observation led us to the proposal of cost-benefit considerations, an important aspect for IQ management. The study is valuable for both research and practitioners. Further research studies can build on our observations and extend the research. Practitioners are provided with arguments for considering IQ trade-offs in relation to security measures.

Keywords: Information Quality, Trade-offs, Security, Cost-Benefit model
1 INTRODUCTION

Over the last two decades researchers have addressed data quality (DQ) and information quality (IQ) from various view points. Researchers have developed many frameworks, criteria lists and approaches for assessing and measuring IQ. Also, literature provides us with numerous case studies, investigating IQ in practice. However, despite the increasing interest in this topic, little is known about the effects and relations between different criteria of IQ. Knight and Burn (2005) point out that despite the sizeable body of literature available relatively few researchers have tackled quantifying some of the conceptual definitions such as security and accessibility. We aim to address this limitation of foremost research and intend to provide insight into associations of different IQ criteria.

Due to the increasing importance of security and accessibility we focus on these aspects and their implications on other IQ dimensions. In our previous research (Fehrenbacher and Helfert, 2008) we show that the importance of security and accessibility as IQ criteria has increased. This is accompanied with an increase in security requirements and complexity of information systems. Due to the increasing complexity and variety of access methods, question about its impact arises. What are implications of security measures on other IQ criteria? Does architecture have a significant (moderating) effect on the relationship between IQ criteria? What is the difference in the impact of accessibility from a workstation compared to a mobile device?

In order to address current limitations, this research focuses on the security and accessibility dimension of IQ. Review of related research shows that most IQ frameworks consider accessibility and security; however researchers classify or consider these IQ dimensions diversely among various IQ frameworks. Furthermore, our research indicates an impact of security and accessibility on other IQ dimensions. An experiment is conducted to evaluate the effect on IQ dimensions of varying levels of security to an Information System (IS). It allows for a thorough analysis of accessibility as a dimension of IQ. We propose a research model and illustrate results of an experiment, which support our research hypothesizes.

The paper is structured as follows. Section 2 reviews related work and provides indications for IQ assessments and trade-offs. Section 3 centers on selected IQ dimensions and proposes a research model and the underlying assumptions. Section 4 presents an experimental research and illustrates the key results. Section 5 discusses the implications of our research and proposes some considerations concerning cost-benefit considerations. We conclude our paper in section 6, in which we discuss some limitations of our research and summarize further research directions.

2 RELATED WORK

Many studies have confirmed that IQ is a multi-dimensional concept (Ballou and Pazer 1985, Redman 1996, Wand and Wang 1996, Wang and Strong 1996, Huang et al. 1999). Over the last two decades, different sets of IQ dimensions have been identified from both the database and management perspectives. Often IQ and DQ alike has been defined as "fitness for use", in that way that data or information of high quality “meets or exceeds users’ requirements.” (Wang and Strong, 1996). Most researchers consider IQ by a set of dimensions that are able to describe different characteristics of data or information. Following many other research, we do not distinguish explicitly between DQ and IQ since our findings are general and suitable for both concepts. Therefore, both terms are used in this article interchangeable.

The literature provides numerous definitions and classifications of IQ dimensions analyzing the problem in different contexts and from different perspectives. Common examples of IQ dimensions are accuracy, completeness, consistency, timeliness, interpretability, and availability. Many researchers have proposed several measures and approaches for each IQ dimension. Some suggestions
include aggregation functions (e.g. weighted sum, ratio, max, and min) in order to provide a unique IQ index. Considering different measurement values of the same dimensions or different measures associated with heterogeneous dimensions is challenging and the subject of current research in IQ.

A variety of IQ assessment methodologies have been proposed over the last decade. We select five popular methodologies (Redman 1996, Huang et al. 1999, Lee et al. 2002, Pipino et al. 2002, and Stvilia et al. 2007) and evaluate these by following criteria: definition of IQ dimensions, classification of IQ dimensions, model, tool, and case study. Definition of IQ dimensions describes which IQ dimensions and perspectives are defined. Classifications of IQ dimensions are used to compare the classification of dimensions in each methodology. The theoretical basis of the methodology is described in the category model. Tool expresses how the methodology is implemented. Case study concentrates on the empirical feasibility of these methodologies.

If the methodology is only applied to a specific domain, it is considered as a specific methodology. If the methodology can be applied to multiple domains, it is regarded as a generic methodology. If the case study is provided in the literature, we classify the methodology as a practical study otherwise it is theoretical. We summarize our evaluation of the five methodologies and their characteristics in table 1.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>12 IQ dimensions are defined from the database community</td>
<td>16 IQ dimensions are defined from management community</td>
<td>15 IQ dimensions are defined from both communities</td>
<td>16 IQ dimensions are defined from both communities</td>
<td>22 IQ dimensions are defined from both communities</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td>A step by step procedure adapted from statistical process control</td>
<td>Adopt Deficiency model of Wand and Wang (1996)</td>
<td>Adopt PSP/IQ model of Kahn et al. (2002)</td>
<td>The model combines subjective and objective assessment</td>
<td>The model consists of activity types, IQ Problems, and IQ taxonomy</td>
</tr>
<tr>
<td><strong>Tool</strong></td>
<td>DCI system</td>
<td>IQ assessment survey</td>
<td>IQ assessment survey</td>
<td>IQ assessment software</td>
<td>IQ assessment survey</td>
</tr>
<tr>
<td><strong>Case Study</strong></td>
<td>Telstra Co. Ltd.</td>
<td>Appliance Company</td>
<td>1, Global Consumer Goods, Inc., 2, Data Product Manufacturing, Inc.</td>
<td>1, Simple Dublin Core 2, English Wikipedia</td>
<td></td>
</tr>
<tr>
<td><strong>Conclusion</strong></td>
<td>Specific, practical</td>
<td>Specific, practical</td>
<td>Generic, Theoretical</td>
<td>Generic, practical</td>
<td>Generic, practical</td>
</tr>
</tbody>
</table>

Table 1. Comparison of IQ assessment methodologies

Pipino et al. (2002) categorizes IQ assessment into objective and subjective assessment. Objective IQ assessments reveal IQ problems in databases while subjective IQ assessments reflect the needs and experiences of data consumers. In order to discuss IQ assessments from objective and subjective perspectives, we follow this general classification.

Objective IQ assessment measures the extent to which information conforms to quality specifications and references. We distinguish objective IQ assessments into two categories: intrinsic and real-world IQ assessment. Intrinsic IQ assessment follows a data perspective and uses data specifications to assess the quality of the data in the database. For example, Savchenko (2003) develops item frequency
rules and regular expression patterns to facilitate an automated intrinsic IQ assessment. Real-world assessment follows the ontological perspective and compares real-world facts to discover IQ deficiencies. For example, Wand and Wang (1996) identify data mapping deficiencies between real world states and its representation in information systems. Overall, objective IQ assessment can be considered as the procedure of comparing current data value with an ideal data value of high quality.

Subjective IQ assessment measures the extent to which information is fitness for use by information consumers. Information consumers assess IQ according to their demands and expectations. Subjective IQ assessment follows the user perspective and focuses on discrepancy between the current quality of information and the user’s expectation. In order to indicate the differences between objective and subjective IQ assessment, we provide a comparison in table 2.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Method</th>
<th>Objective assessment</th>
<th>Subjective assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool</td>
<td>Software</td>
<td></td>
<td>Survey</td>
</tr>
<tr>
<td>Measuring Object</td>
<td>Data</td>
<td></td>
<td>Information</td>
</tr>
<tr>
<td>Standard</td>
<td>Rules, Patterns</td>
<td></td>
<td>User Satisfaction</td>
</tr>
<tr>
<td>Process</td>
<td>Automated</td>
<td></td>
<td>User Involved</td>
</tr>
<tr>
<td>Result</td>
<td>Single</td>
<td></td>
<td>Multiple</td>
</tr>
</tbody>
</table>

| Table 2. Comparison of objective and subjective IQ assessment |

Objective IQ assessment uses mostly software applications to evaluate the data in a database by a set of quality rules. This can often been carried out automatically. Subjective IQ assessment uses typically a survey or interview approach to measure the contextual information by data consumers. The advantage of objective IQ assessment is that it allows one to automatically process large data sets. Subsequently to the assessment, the objective approach obtains a single or aggregated assessment result. Subjective IQ assessment normally involves data users’ opinion or evaluations on data samples. Although subjective assessment may contain different assessment results due to the different opinions and roles of information consumers (Strong et al. 1997), the advantage of subjective IQ assessment is the measurement of a comprehensive set of IQ dimensions. Furthermore, certain IQ dimensions such as believability and reputation are only suitable for subjective IQ assessment. Recognizing the advantages of both objective and subjective assessment, researchers (Pipino et al. 2002, Kahn et al. 2002) have combined objective and subjective IQ assessment.

A number of literatures have analyzed dependencies of IQ dimensions. Ballou and Pazer (1995) propose a framework to investigate tradeoffs between accuracy and timeliness in the context of decision making. Redman (1996) points out that timeliness has an impact on accuracy. Ballou and Pazer (2003) model the utility and tradeoffs between completeness and consistency. Olson (2003) implies the relationship between accuracy and completeness and states that consistency is a part of accuracy. Cappiello et al. (2004) analyze the time-related accuracy and time-related completeness in multi-channel information systems. Amicis et al. (2006) propose a data-driven approach to analyze the dependency between syntactic accuracy and timeliness as well as the dependency of completeness and timeliness.

Observing the literatures above, we divide relationships of IQ dimensions into two categories: negative correlated and positive correlated dependencies. Negative correlation refers to the improvement of one IQ dimension that may lead to a decreasing value in another dimension (often also referred as IQ tradeoffs). For example, by introducing new information to improve completeness, the new introduced information may be inconsistent with the existing information. In this manner, completeness and consistency are negatively correlated. Positive correlation means two IQ dimensions are mutually contributing to a shared set of IQ problems. For example, when timeliness and accuracy are sharing outdated data as their mutual IQ problem, the improvement of timeliness may lead to an increasing value in accuracy. In this way, timeliness and accuracy are positively correlated. According to the discussion above, we summarize correlations of IQ dimensions in table 3.
Improvement of certain information quality dimensions, may lead to a decreasing value in other dimensions: tradeoffs between completeness and other dimensions, accessibility and other dimensions, security and other dimensions, relevancy and other dimensions.

When we improve IQ dimension 1, IQ dimension 2 may be improved or remain at the same quality value. The quality values depend on the mutual IQ problem.

Table 3. Dependency of IQ dimensions

Many researchers have indicated various relations between IQ criteria, such as timeliness and availability. In table 4 we combined a list of common IQ criteria and relations described in literature (Fehrenbacher and Helfert, 2008). We indicate a potential negative (N) or positive correlation (P). As the list indicates, various trade-offs of IQ dimensions can be assumed. However, most researchers merely propose some form of relations but do not further investigate the strength or direction of the relation.

<table>
<thead>
<tr>
<th>Item 1</th>
<th>Item 2</th>
<th>N</th>
<th>P</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timeliness</td>
<td>Believability</td>
<td>◆</td>
<td></td>
<td>Eppler (2001) adapted</td>
</tr>
<tr>
<td>Timeliness</td>
<td>Consistent representation</td>
<td>◆</td>
<td></td>
<td>Scannapieco and Batini (2006) adapted</td>
</tr>
<tr>
<td>Timeliness</td>
<td>Completeness</td>
<td>◆</td>
<td></td>
<td>Scannapieco and Batini (2006)</td>
</tr>
<tr>
<td>Completeness</td>
<td>Accuracy</td>
<td>◆</td>
<td></td>
<td>Ballou and Tayi (1999), Cappiello Francalanci and Pernici (2003), Fisher et al. (2006)</td>
</tr>
<tr>
<td>Completeness</td>
<td>Consistent representation</td>
<td>◆</td>
<td></td>
<td>Ballou and Pazer (2003), Scannapieco and Batini (2006) adapted</td>
</tr>
<tr>
<td>Accessibility</td>
<td>Accuracy</td>
<td>◆</td>
<td></td>
<td>Missier et al. (2003)</td>
</tr>
</tbody>
</table>

Table 4. Selected relationships of IQ criteria

Based on a framework proposed by Wang and Strong in 1996, we evaluated in a recent empirical research the importance of IQ dimensions (Fehrenbacher and Helfert, 2008). Comparing the item rankings our research showed a similar ranking pattern, however accessibility and security among...
others received increasing importance. Due to its increased importance but limited attention in foremost literature, we decided to focus our research on security and accessibility. To illustrate the effects and implications of trade-offs among IQ dimensions, we analyse particular effects of security and accessibility on timeliness.

3 INFORMATION QUALITY DIMENSIONS AND RESEARCH MODEL

3.1 Security and Accessibility as IQ dimensions

The literature has put forward a number of frameworks and classified the dimensions associated with each of these frameworks. In addition to the variety of IQ frameworks, most provide their own definitions for security and accessibility associated with particular IQ frameworks. Generally it can be assumed, that an increased level of security impacts on the accessibility to an information system. The relationship between security and accessibility allows for examination of the attributes of accessibility. Thus, accessibility in essence is a function of security. An examination of the accessibility dimension directly relates to the accessibility dimension. Loshin (2001) describes it as the degree of ease of access to information as well as the breadth of access. Wang and Strong (1996) consider that access security is also an important concept that must be taken into account when considering the dimension. Batini and Scannapieco (2006) describe accessibility in terms of the ability of the user to access the data from his / her own culture, physical status / functions and technologies available. In summary, the definition of accessibility is framework dependent. Nonetheless, there is also an ongoing debate about the relation of accessibility to IQ and some frameworks do not even consider it as a dimension of IQ.

With a view to analyzing in more detail the frameworks that are of specific interest to the accessibility dimension, we examined selected IQ frameworks. Table 2 summarizes these frameworks outlining the dimensions associated with each framework. We selected most prominent frameworks in the field of IS and IQ research. The examination allows this research to focus in more detail on the frameworks that pertain to accessibility as noted in column three of table 2.

<table>
<thead>
<tr>
<th>Framework</th>
<th>Dimensions / Quality Category</th>
<th>Accessibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zeist and Hendricks (1996) (Extended ISO Model)</td>
<td>Functionality, Reliability, Efficiency, Usability, Maintainability, Portability</td>
<td></td>
</tr>
</tbody>
</table>
The examination of the IQ frameworks in table 5 demonstrates that accessibility does feature as a dimension to varying degrees across many frameworks. Reviewing these selected frameworks resulted in a list of drawbacks, which current frameworks do not address. We identified the following five key research challenges related to the accessibility and security dimensions of IQ:

- What are the impacts of accessibility / security on the overall IQ?
- How do accessibility / security impact on other dimensions in an IQ framework?
- Do current IQ frameworks provide valid and reliable measures?
- Is the impact of accessibility / security consistent across IQ frameworks?
- What impact do multiple access / security methods have upon IQ?

Answering these questions would allow for providing an insight or even quantifying the impact of accessibility and security on other IQ dimensions and thus on IQ.

Table 5. IQ Frameworks and Dimensions

<table>
<thead>
<tr>
<th>Framework</th>
<th>Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhu &amp; Gauch (2000) (Quality Metrics for Information retrieval on www)</td>
<td>Currency, availability, information to noise ratio, authority, popularity, cohesiveness</td>
</tr>
</tbody>
</table>

3.2 Hypothesis and Assumptions

As discussed above, research has indicated many relationships between IQ dimensions. Several studies aimed to investigate the relationship between selected dimensions, however regarding its importance
accessibility and security related trade-offs are often not considered. We centre our research on these important dimensions, and investigate their effect on other IQ dimensions. From the related literature above, we can hypothesis a general relation between the dimensions. However, little is know of which dimensions are affected or their significance.

In our article we centre on process oriented IQ dimensions such as timeliness and availability. Other intrinsic IQ dimensions, such as consistency, completeness and accuracy are less affected. Considering Security measures, we assume following indicative relationship. As security measures are increased, **timeliness and availability decrease. Other dimensions are not affected**. Security measures act as independent variable, whereas other IQ dimensions act as dependent variables. We assume that this effect can be perceived by information users, and thus result in a lower perceived IQ. There are many suggestions for discrete measures for IQ values, and thus we can assume such measure. For security we assume a categorical measure, consisting of advanced, intermediate and basic security measures.

## 4 AN EXPERIMENT TO SUPPORT THE RESEARCH MODEL

In order to support our research model, we conducted an experiment. In contrast to other research methods, experimental research is particularly suitable for the identification and study of relationships between a small number of variables. Experimental research is found to be effective in addressing the cause and effect relationship (Campbell and Stanley 1963, Jarvenpaa et al. 1985). Although our experiment represents a simplified real world scenario, it assists us to understand fundamental relationships between IQ criteria. However, one of the challenges is the isolation of and control of exogenous factors, such as decision complexity and experience.

Data can be collected in a number of ways in order to answer research questions. It can be gathered by direct observation or reported by the individual. Fisher et al. (2001) indicate that systematically collecting data to measure and analyze the variation of one or more processes forms the foundation of statistical process control. In the case of an experiment a variable is manipulated and the corresponding effect on the other variables is noted. Fisher et al. (2001) also point out that a statistical experiment is a planned activity where variables that have the potential to affect response variables are under the control of the researcher.

In order to examine the impact of accessibility dimension as an IQ dimension, we examine four IQ dimensions across three architectures and two IS domains. The aim of this experiment is to demonstrate what trade offs if any are associated with varying levels of security.

- **IQ Dimensions:** As IQ is a multidimensional concept the impact on individual dimensions is examined in the experiment. For our research, we selected four dimensions that are common across IQ frameworks free-of-error, completeness, consistency and timeliness. In order to measure IQ, a subset of the questions from the AIMQ (Lee et al. 2002) methodology are employed. The specific survey questions with respect to free-of-error, completeness, consistency and timeliness were used.

- **Architectures:** Web, Client Server, Work Station

- **Domains:** The two IS domains are a library system and a student exam result system. The major areas of functionality of both systems were employed during the experiment. Three different access methods were used namely workstation, client server and web. These are used on day to day operation of both systems. All users were also day to day operators of the systems.

The experiment sets different levels of security and measures the corresponding effects on the four dimensions. Three levels of security are manipulated in the experiment basic, intermediate and advanced. Basic security has no restrictions set while the advanced level is stringent. There were twenty seven participants for the library system and eighteen for the student exam result system. The results recorded are the average scores for the twenty seven participants of the Library IS and eighteen participants of the student exam system IS. The experiment was conducted over a two day period in March 2008. The results of our experiment are illustrated descriptively and set out in a number of tables below (tables 6 – 11). Subsequently we describe and interpret the results.
<table>
<thead>
<tr>
<th>Security Level</th>
<th>Architecture</th>
<th>Free-of-Error</th>
<th>Completeness</th>
<th>Consistency</th>
<th>Timeliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced</td>
<td>Web</td>
<td>79%</td>
<td>84%</td>
<td>73%</td>
<td>46%</td>
</tr>
<tr>
<td></td>
<td>Client Server</td>
<td>83%</td>
<td>88%</td>
<td>77%</td>
<td>48%</td>
</tr>
<tr>
<td></td>
<td>Work Station</td>
<td>81%</td>
<td>85%</td>
<td>76%</td>
<td>56%</td>
</tr>
</tbody>
</table>

Table 6. Library IS Domain – Security Level Advanced

<table>
<thead>
<tr>
<th>Security Level</th>
<th>Architecture</th>
<th>Free-of-Error</th>
<th>Completeness</th>
<th>Consistency</th>
<th>Timeliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>Web</td>
<td>74%</td>
<td>71%</td>
<td>71%</td>
<td>60%</td>
</tr>
<tr>
<td></td>
<td>Client Server</td>
<td>82%</td>
<td>78%</td>
<td>77%</td>
<td>61%</td>
</tr>
<tr>
<td></td>
<td>Work Station</td>
<td>84%</td>
<td>81%</td>
<td>79%</td>
<td>64%</td>
</tr>
</tbody>
</table>

Table 7. Library IS Domain – Security Level Intermediate

<table>
<thead>
<tr>
<th>Security Level</th>
<th>Architecture</th>
<th>Free-of-Error</th>
<th>Completeness</th>
<th>Consistency</th>
<th>Timeliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td>Web</td>
<td>78%</td>
<td>79%</td>
<td>74%</td>
<td>81%</td>
</tr>
<tr>
<td></td>
<td>Client Server</td>
<td>85%</td>
<td>81%</td>
<td>75%</td>
<td>87%</td>
</tr>
<tr>
<td></td>
<td>Work Station</td>
<td>82%</td>
<td>84%</td>
<td>77%</td>
<td>89%</td>
</tr>
</tbody>
</table>

Table 8. Library IS Domain – Security Level Basic

<table>
<thead>
<tr>
<th>Security Level</th>
<th>Architecture</th>
<th>Free-of-Error</th>
<th>Completeness</th>
<th>Consistency</th>
<th>Timeliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced</td>
<td>Web</td>
<td>74%</td>
<td>81%</td>
<td>74%</td>
<td>49%</td>
</tr>
<tr>
<td></td>
<td>Client Server</td>
<td>77%</td>
<td>83%</td>
<td>77%</td>
<td>53%</td>
</tr>
<tr>
<td></td>
<td>Work Station</td>
<td>79%</td>
<td>88%</td>
<td>80%</td>
<td>62%</td>
</tr>
</tbody>
</table>

Table 9. Student Exam IS Domain – Security Level Advanced

<table>
<thead>
<tr>
<th>Security Level</th>
<th>Architecture</th>
<th>Free-of-Error</th>
<th>Completeness</th>
<th>Consistency</th>
<th>Timeliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>Web</td>
<td>77%</td>
<td>77%</td>
<td>75%</td>
<td>64%</td>
</tr>
<tr>
<td></td>
<td>Client Server</td>
<td>81%</td>
<td>84%</td>
<td>78%</td>
<td>69%</td>
</tr>
<tr>
<td></td>
<td>Work Station</td>
<td>79%</td>
<td>85%</td>
<td>72%</td>
<td>71%</td>
</tr>
</tbody>
</table>

Table 10. Student Exam IS Domain – Security Level Intermediate

<table>
<thead>
<tr>
<th>Security Level</th>
<th>Architecture</th>
<th>Free-of-Error</th>
<th>Completeness</th>
<th>Consistency</th>
<th>Timeliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td>Web</td>
<td>80%</td>
<td>75%</td>
<td>71%</td>
<td>81%</td>
</tr>
<tr>
<td></td>
<td>Client Server</td>
<td>82%</td>
<td>78%</td>
<td>72%</td>
<td>82%</td>
</tr>
<tr>
<td></td>
<td>Work Station</td>
<td>88%</td>
<td>79%</td>
<td>76%</td>
<td>86%</td>
</tr>
</tbody>
</table>

Table 11. Student Exam IS Domain – Security Level Basic
The results are based on a subset of the AIMQ survey instrument (Lee et al. 2002). This questionnaire has been used in many studies. Appendix one lists the questions that were employed in the survey. The key findings of the experiment indicate that as security levels are manipulated the other IQ dimensions are affected. Fisher et al. (2001) point out that if data is not accessible then quality will decrease because information can not be accessed in a timely fashion. There is a trade-off between security and timeliness. Our result show timeliness is significantly affected. As the level of security became more advanced the users survey results with respect to the timeliness dimension were less and less satisfactory. This research indicates that security as a dimension of IQ can have different levels and the more access is restricted the greater the dissatisfaction with the timeliness dimension. It is not merely two states of accessible and inaccessible.

A closer examination of the timeliness dimension is warranted. For example in table 6, table 7, table 8 which reflect results with respect to the library IS there is an increase in satisfaction in the survey results with respect to timeliness as the level of security is lessened. At a high level of security the satisfaction with timeliness is 46% for web access, 48% for client server and 56% for work-station. This is an average satisfaction of 50% with the timeliness dimension. As can be seen from the results the average increases to 61.6% for intermediate and 85.6% for a basic level of security. The results for the student exam system IS domain display a similar pattern with an average of 55% satisfaction with the timeliness dimension when there is an advanced level of security where as at a basic level of security the satisfaction was at 83%.

The other dimensions surveyed; free-of-error, completeness and consistency did not radically change across IS domain. Another interesting finding was the users’ satisfaction with the IQ dimensions when web architecture was employed. It compared less favourably with client server or workstation architectures. This was the case for both domains examined.

5 IMPLICATIONS – TOWARDS A COST/BENEFIT MODEL

As our research above indicates, there is an interesting relationship between timeliness and security. In the following section we analyse the implications of these observations, and examine the impact of IQ level from a cost and benefit perspective.

The trade-off between security and timeliness is often analyzed in real-time applications. In fact, these applications have clear timeliness constraints but sometimes they can also have security issues in addition to timing constraints. The timing constraints of real-time applications are typically specified in the form of rules that require that an operation has to be completed in a specified time. Failures of such rules can cause critical situations since the provided results may be useless if not timely. Real-time systems are often adopted in particular environments such as defence systems, air traffic control or stock trading where data security is a fundamental aspect. These scenarios require access restrictions in order to differentiate the data accessibility on the basis of the users that require some information. Security and timeliness requirements conflict with each other since the implementation of methods to guarantee data security may introduce some delays in the application execution. Whether to maintain timeliness or security is dependent upon the system.

Let us consider a system A in which security is preferred to timeliness and a system B in which timeliness is preferred to security. Considering the level of security (SL), we can assume that timeliness is inversely proportional to the security level along a general exponential decay trend. On the contrary, increasing the security level the quality costs increase exponentially. In fact, large investments are needed for a secure system and thus for the adoption of complex protocols. We make the assumption that the economical benefits deriving from IQ are proportional to the value assessed for the IQ dimensions. Along these considerations, it is possible to compare costs and benefits related to the different security levels and evaluate the total profit in the two considered scenarios. Cost and benefit analysis show (see Figure 1) that it is possible to define the most suitable security level by
calculating, the maximum value of profit, resulting from the subtraction of the security costs from the IQ benefits.

![Diagram showing cost/benefit analysis]

Figure 1. Cost/benefit analysis – (a) Security is considered more important than timelines, (b) timeliness is considered more important than security

## 6 CONCLUSIONS

In literature as well as in practice, it is often assumed that in order to reach the maximum IQ and the best satisfaction of user requirements maximum IQ is required. However, this disregards the existence of trade-offs among IQ dimensions. Our research indicated that there are significant effects between information system decisions, such as security measures, and IQ. Applying three distinct security levels, advanced, intermediate and basic, we found that timeliness is significantly affected. Other dimensions are not significantly affected in our experiment. The research led us to the proposal of some cost-benefit considerations. As discussed, the perception of IQ is important in order to evaluate an optimal security level.

Although we provided an experimental research design, which provided some insights into the relationship between security and IQ as well as IQ trade-offs, our research is still rather conceptual. Due to the relatively low number of participants, the analysis is descriptive. In future research we intend to extend the number of participants and apply suitable quantitative analyses techniques. We also plan to investigate further IQ trade-offs, which subsequently assist us to understand cost-benefit considerations. However, although our research has some limitations, we believe that the results are beneficial for researchers and practitioners. Further research studies can build on our observations and extent the research with experimental or empirical research approaches. A number of the IQ frameworks examined in table two do not consider security. As a result of the initial findings of this research especially with respect to the key finding of security levels and the timeliness dimension it is suggested IQ frameworks need to take the levels of security into account. This it is argued will lead to a more comprehensive view of IQ. Practitioners are provided with arguments for considering IQ trade-offs in relation to security measures. This research suggests that the factors with respect to the architecture employed need consideration when IQ policies are being designed and implemented. Furthermore, in contrast to many empirical researches, this article applies an experimental research approach. The authors believe that this more rigorous approach can complement and enhance the IQ research area, in which case studies are dominating. The authors strongly believe that more experimental research is needed in order to complement the important but often practical oriented research in this particular domain.

## REFERENCES


### APPENDIX 1

The user is surveyed for their opinion with respect to the following questions, summarized in Table 12.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timeliness</td>
<td>This information is sufficiently current for our work.</td>
</tr>
<tr>
<td></td>
<td>This information is not sufficiently timely.</td>
</tr>
<tr>
<td></td>
<td>This information is not sufficiently current for our work.</td>
</tr>
<tr>
<td></td>
<td>This information is sufficiently timely.</td>
</tr>
<tr>
<td></td>
<td>This information is sufficiently up-to-date for our work.</td>
</tr>
<tr>
<td>Accuracy</td>
<td>This information is correct.</td>
</tr>
<tr>
<td></td>
<td>This information is incorrect.</td>
</tr>
<tr>
<td></td>
<td>This information is accurate.</td>
</tr>
<tr>
<td></td>
<td>This information is reliable.</td>
</tr>
<tr>
<td>Completeness</td>
<td>This information includes all necessary values.</td>
</tr>
<tr>
<td></td>
<td>This information is incomplete.</td>
</tr>
<tr>
<td></td>
<td>This information is complete.</td>
</tr>
<tr>
<td></td>
<td>This information is sufficiently complete for our needs.</td>
</tr>
<tr>
<td></td>
<td>This information covers the needs of our tasks.</td>
</tr>
<tr>
<td></td>
<td>This information has sufficient breadth and depth for our task.</td>
</tr>
<tr>
<td>Consistency</td>
<td>This information is consistently presented in the same format.</td>
</tr>
<tr>
<td></td>
<td>This information is not presented consistently.</td>
</tr>
<tr>
<td></td>
<td>This information is presented consistently.</td>
</tr>
<tr>
<td></td>
<td>This information is represented in a consistent format.</td>
</tr>
</tbody>
</table>

*Table 12: User Survey*
INTEGRATING VALUE-ADDING MOBILE SERVICES INTO AN EMERGENCY MANAGEMENT SYSTEM FOR TOURIST DESTINATIONS

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0305.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>Design Science, Tourism, Mobile communications, User participation</td>
</tr>
</tbody>
</table>
INTEGRATING VALUE-ADDING MOBILE SERVICES INTO AN EMERGENCY MANAGEMENT SYSTEM FOR TOURIST DESTINATIONS

Scherner, Tobias, Goethe-University Frankfurt, Grüneburgplatz 1, 60323 Frankfurt, Germany, scherner@wiwi.uni-frankfurt.de
Muntermann, Jan, Goethe-University Frankfurt, Grüneburgplatz 1, 60323 Frankfurt, Germany, muntermann@wiwi.uni-frankfurt.de
Roßnagel, Heiko, Fraunhofer Institute for Industrial Engineering (IAO), Nobelstr. 12, 70569 Stuttgart, Germany, heiko.rossnagel@iao.fraunhofer.de

Abstract

Recently, several disasters have affected the tourism industry. In order to mitigate the effects of disasters, increasing the level of preparedness is essential. However, despite the devastating effect disasters can have on tourism, few tourism organizations have properly developed disaster strategies as an integral part of their business plans. Emergency management systems that utilize mobile communication infrastructures can provide prompt information delivery to save human lives. Several supra-national initiatives and research projects are working on possibilities to facilitate mobile communication networks for emergency management systems. However, the success of such systems depends on users being familiar with the service, which is difficult to achieve, if the system is solely used for emergency management. Therefore, we propose a system design that allows the integration of mobile value-adding services. We also present exemplary services, which offer value to tourists and create business opportunities for the tourism industry. The central component of our proposed system design is a service platform, which communicates with mobile network operators and provides basic services for service providers from the tourism industry and the emergency manager via service interfaces.

Keywords: Emergency Management, Tourism, Mobile Communications, Design Science.
1 INTRODUCTION

Disasters can have negative and lasting effects on the tourism industry (Faulkner 2001). In recent years, several crises and disasters have affected the global tourism industry (Ritchie 2004). Examples comprise terrorist attacks (Paraskevas and Arendell 2007), political instability (Fuller and Fugal 2008), economic recession (Kreimer et al. 2003), bio-security threats and natural disasters (Hystad and Keller 2007). Since tourism is also an important economic factor for many countries and many destinations depend on tourism for their growth and survival, there is an increasing pressure on managers and policy makers concerned with tourism to consider the impact of crises and disasters on the industry and to develop strategies to deal with their impacts (Ritchie 2004). In order to mitigate disaster effects and to accelerate the healing process, it is essential to undertake steps to increase the level of disaster preparedness, including infrastructure investments for warning systems and training activities (Johnston et al. 2007). However, despite the devastating effect disasters can have on tourism, few tourism organizations have properly developed disaster strategies as an integral part of their business plans (Faulkner 2001). Emergency management systems (EMS) provide the capability to address this dilemma and to enable disaster forces to manage disaster events, including detection and analysis of incidents (Carver and Turoff 2007). Persons in charge should be supported to prepare evacuations, control and support disaster forces and to locate victims (Carver and Turoff 2007). Since mobile communication infrastructures offer standardized wireless communication services in almost all countries (GSMworld 2008) and allow a fast diffusion of information, they provide a promising technological basis for saving human lives in emergencies. If a disaster event occurs, they can for example allow emergency managers to distribute warnings to the effected areas by cell broadcast to ensure warnings of potential victims in time (Fritsch and Scherner 2005). One of the most promising initiatives is described in the recently published report on the Commercial Mobile Alert System (Federal Communications Commission 2008) which is supposed to be in operation in 2010.

One of the most crucial requirements for preparedness is that ordinary people are used to the system in order to react on warning signals without any delay (Gruntfest and Huber 1989). Meeting this requirement is extremely difficult if the system design is solely used for early warnings. The success of emergency management systems clearly depends on well-trained users being familiar with the service functionalities provided (Turoff et al. 2004a). For an infrequently used emergency management system, limited practical experience of users can be expected (Manoj and Hubenko Baker 2007). Technically, many services used for emergency management systems do not differ from services used in day by day use cases. For example the upload of a picture could be used to inform emergency managers but also for online community services. The challenge is that both functionalities have to be integrated in a design, which allows both perspectives and supports the user to become familiar with the functionalities itself. We will therefore argue that offering mobile value-adding and emergency management services via the same infrastructure could create new business opportunities for the tourism industry. Services can be used to improve the holiday experience for tourists and to help tourist guides to manage the customers. For example, (Zipf and Malaka 2001) argue that location-based services in tourism settings can provide the basis for a range of novel tourism-related applications and open up new business opportunities to service providers. Integrating value adding mobile services into an emergency management system thus improves users’ familiarity with the system while at the same time offering a perceived value to customers. Furthermore, revenue created by value-adding services could provide the necessary funds to finance the investments in the mobile emergency management system. Besides sharing the costs for operating the infrastructure, the idea of providing both commercial and disaster warning services based on a technically integrated platform positively affects users’ experience due to a common client base.

This paper is structured as follows. We first present the methodological approach of our work in section 2. In section 3, we derive requirements for emergency management systems from the literature, discuss how current systems rate against these requirements and propose a system design to
address them. In section 4, we introduce exemplary value-adding mobile service in the tourism sector and discuss possibilities for integrating them into an infrastructure that can be used for emergency management systems as well as possible business implications. The limitations of our work are stated in section 5 and section 6 provides a conclusion.

2 METHODOLOGICAL APPROACH

We address the introduced problem with a conceptual system design providing a technological basis for emergency management and value-adding services being offered to different parties. This system design represents an IT artefact instantiation that aims at contributing to the problem’s solution, i.e. at (1) minimizing the impact of crises or disasters and at (2) offering novel services to customers. As system architectures, system designs or prototypical software applications, IT artefact instantiations demonstrate the feasibility of an approach developed. Due to its utility-centric focus, design science research contributions present novel IT artefacts and suitable evaluation approaches that address the artefact’s appropriateness to contribute to the problems’ solution (Nunamaker et al. 1991). These two facets of rigorous design science-oriented research contribute to the foundations and the methodologies pool of Information Systems research, i.e. they contribute to its knowledge base (Hevner et al. 2004). In the following, we present a design of an emergency management system, which utilizes communication facilities of mobile communication networks such as GSM. In contrast to existing approaches such as the Global Disaster Alert and Coordination System (GDACS) (European Communities 2008), our system design provides facilities for integrating EMS functionalities and commercial value-adding services on the basis of a common platform. In order to evaluate our proposed system design, we then present exemplary value-adding services that demonstrate its feasibility. Further evaluation is provided by a qualitative analysis of potential benefits being provided to the different parties including the public sector, the industry and customers. We therefore follow the classic approach of design science-oriented research (March and Smith 1995) as we first present a developed IT artefact and we second evaluate the artefact’s benefits for potential users. Therefore, and e.g. in contrast to behavioural science oriented research, the goal of this contribution is utility, namely the benefits our system design can provide (Simon 1969).

3 EMERGENCY MANAGEMENT SYSTEMS

Having outlined our methodological approach, we will now derive system requirements from the literature and review how they are addressed by existing emergency management systems. We then use these requirements as the basis for a proposed system design.

3.1 System Requirements

From the literature we derived the following high-level requirements that need to be addressed by an emergency management system, which could also be used for secondary purposes. They comprise (1) system effectiveness (Johnston et al. 2007), (2) reliability (Zeckhauser 1996), (3) cost efficiency (Zeckhauser 1996), (4) smooth service integration (Ritchie 2004), (5) multilateral user interaction (Turoff et al. 2004b), (6) availability (Faulkner 2001) and (7) security (Valtonen et al. 2004). In the following, we further explore these identified requirements:

**Effectiveness:** Effective early warning systems have to be based on regularly used communication infrastructures and people have to know how to react on visual and acoustic warning signals (Johnston et al. 2007). This requirement holds for commercially successful infrastructures (Buhalis 1998) as well as early warning systems. This requirement suggests that commonly used infrastructures, like those operating on the GSM specification provide a suitable basis for mobile services. GSM represents a standardized communication link between disaster management forces and victims being already used by a large installed base of more than 2 billion customers worldwide (GSM-Association 2008).
Reliability: Despite effectiveness, reliability of underlying networks is an important success criterion, as up-to-date systems have to be regularly maintained to ensure reliability. Therefore, this requirement is tied to regular usage (Kron and Thumerer 2002) of the system by parties involved into emergency management. Furthermore, potentially affected people have to be involved in preparedness measures. Otherwise, alerting systems that are not regularly used for other purposes often become outdated, unmaintained and do not get adapted to changing requirements (Gruntfest and Huber 1989).

Cost efficiency: The system shall offer interfaces to existing legacy systems providing the opportunity to include available information from sources of the public sector and third parties. Furthermore, it should provide functionalities to direct messages in a cost-efficient manner (Zeckhauser 1996; Zhao et al. 2005). Major revisions on mobile communication infrastructures and on mobile devices of customers will thus violate this requirement. Secondary use of early warning systems in terms of offering value-adding services shall be supported. The European Telecommunications Standards Institute (ETSI) is developing a worldwide interoperable architecture for emergency management services, in which commercial use is explicitly mentioned (European Telecommunications Standards Institute 2007).

Smooth service integration and device interoperability: End-users should be able to use their already existing equipment without being forced to invest in additional technology. High prices for new equipment tend to result in reduced acceptance by end-users (Chen et al. 2002) and therefore the interoperability with already existing mobile devices must be considered. In the E911 project, end-user technology is regulated and mobile operators have to ensure that certain technologies are available at the end-users side (Federal Communications Commission 2001). In a report for the Federal Communications Commission (FCC) (Hatfield 2002) argues that the lack of integrated commercial services is one of the major obstacles E911 is facing. We therefore conclude that the lacking integration of emergency management and commercial location based services needs to be addressed in holistic manner.

Multilateral user interaction: For supporting front-line responders in forming dynamic teams, the underlying system design shall support multilateral point-to-point communication and provide availability of locally available human and technical resources. This requirement implies that the system design is able to assign access control right on the spot in emergency cases (Turoff et al. 2004b).

Availability: Availability of emergency management systems has been investigated in depth by many authors and is a common problem for systems dealing with disasters and other events with similar impacts on human life (Faulkner 2001). Taking into account that the convergence of data services in terms of using common communication infrastructures, such as internet protocol-based networks, affects all levels of preparedness, the problem of limited availability is not unique to our suggested design infrastructure. Looking at the category of predictable events within an expectable timeframe, such as storms, heavy rain and to certain degree earthquakes, the problem of limited availability during events is clearly dominated by the benefit that warnings communicated in textual and acoustical form can be brought out to potential victims.

Security: Besides availability, other security-related requirements have to be fulfilled by emergency management systems. End-users should be able to verify integrity and authenticity of emergency messages at any time (Valtonen et al. 2004). Furthermore, the system has to ensure that no party could deny its responsibility having sent notifications to potential victims. (Fritsch and Scherner 2005) proposed a system in which early warning functionalities, enhanced by privacy-preserving mechanisms, are described. Based on this work, a detailed analysis how authenticity, integrity and non-repudiation can be provided by using electronic signatures has been provided in (Roßnagel and Scherner 2006).
3.2 Existing Emergency Management Systems

Several supra-national initiatives and research projects are working on possibilities to facilitate mobile communication networks for emergency management systems. Most initiatives, as those described below, are mainly concentrating on certain phases of disasters (McEntire et al. 2002), such as detecting events and forecasting the impact. Other approaches concentrate on early warnings by providing mechanism for delivery of warnings and instructions to potential victims. The GDACS (European Communities 2008), an initiative initially launched by the European Commission, focuses on gathering information about earthquakes and tsunamis. The system automatically evaluates their impact and disseminates the information to prior registered persons around the world. Messages are delivered by SMS or E-mail, depending on the customers preferences. We conclude that the approach addresses the requirements of cost efficiency by using commonly available communication channels. However, effectiveness is only partially addressed since it can not be ensured that the addressee realizes the warning and could act accordingly.

The SMS-alert initiative in the Netherlands allows local residents to subscribe, and receive SMS text alerts from the local police regarding activity in their geographic area (Korteland and Bekkers 2007).

3.3 Proposed System Design

The proposed system design is based on Fritsch and Scherner (2005), in which early warning functionalities of a mobile emergency management system are described. To open up the system design to commercial service providers, we have defined roles of the public sector and the tourism industry. Both entities are operating on a subset of responsibilities and features that the whole system provides. Therefore, their roles are introduced as sub-roles. They have access to the underlying communication infrastructures and thereby, each provides specific services the stakeholder is specialized in. Figure 1 illustrates how the different parties interact within our system design and which services are provided by and to whom.

The central component in our design is the service platform, which is maintained by the platform operator. The platform communicates with mobile network operators and provides basic services for service providers from the tourism industry and the emergency manager via standardized service interfaces. The basic services include localization of mobile subscribers, message delivery via SMS and CBS, multilateral data transfer, access to information databases, support for mobile communities and billing services for mobile payment and mobile ticketing.

The platform operator can be a public or private entity. Its main task is to operate the information system infrastructure and to provide basic services via service interfaces to the involved parties.
Several different entities could take on this role. The platform could for example be operated by the tourist office of the destination. The tourism experience could be improved by new tourism services offered using this platform and by establishing a powerful emergency management system signalling preparedness to potential tourists. Naturally, commercial service providers could operate this infrastructure on behalf of the tourism office. Also, companies with a strong commitment to the destinations community, such as public transport providers, could take on the role of platform providers. For example the German research project VeRSiert (Projekt VeRSiert 2008) has proposed a similar infrastructure (Roßnagel et al. 2008a), which is operated by a public transport provider for coordinating major events in Cologne (Roßnagel et al. 2008b). The role of platform operator could also be fulfilled by commercial service providers, who charge for the basic services that are offered over the service platform.

Figure 1: Proposed system design

The emergency manager provides all emergency management related services to ordinary people and ensures efficient notifications, which is henceforth classified as service category 1. Furthermore, this role makes services available to the tourism sector and its employees (service category 2) in order to help them to prepare for emergencies and to offer guidance in emergency situations. For the implementation of these emergency services, the emergency manager utilizes the basic services provided by the service platform. The tourism industry offers commercial services to tourists (service category 3), which we elaborate in section 4. These services are built upon the basic services provided by the service platform, which allows a rapid development of services due to the already existing building blocks. The service aggregation on behalf of the tourism industry is done by the platform operator who takes on the role as an information intermediary (Bhargava and Choudhary 2004) providing n to m links between different entities of tourists and the tourism industry. Offering commercial services via the same infrastructure is one of the major challenges. The contribution of this approach is therefore that a) people get involved in mobile services as they are offered by mobile emergency services and b) commercial players have a vital interest in keeping the used infrastructure up-to-date. The presented system design allows sub-roles for providing different services via the same infrastructure. Thus, the integrated communication infrastructure (service platform) provides the public sector (emergency manager) and the tourism sector (tourism service providers) with access to their customers and clearly separated duties and responsibilities, solely for their use cases. Thereby the
emergency management system can meet the requirements effectiveness, reliability and cost efficiency by using the same infrastructure for several different use cases. It is not a stand-alone system, which has to be maintained separately and necessary adaptations to changing requirements are becoming more likely. The system now serves a regular secondary use. Furthermore, the stakeholders get used to the system in day-by-day use cases and gain experiences how to react to different messages.

4 MOBILE VALUE-ADDING SERVICES

We will now focus on exemplary mobile value-adding services and their integration with the emergency management system. After a description of potential exemplary services we will discuss service integration and economic feasibility. A broader discussion of tourist services that are suitable for integration into an emergency management system can be found in (Scherner and Muntermann 2008).

4.1 Exemplary Services

Buhalis (1998) and Buhalis and Licarta (2002) analyzed that information intermediaries are a suitable infrastructure component for offering services to an installed base of travellers by providing information from various sources in a concise manner. In Lee and Mills (2007) it is suggested that the key factors for travelling customers’ satisfaction regarding the support through mobile devices are determined by a) the degree of perception and b) by the perceived value of services provided. Satisfaction of customers is, following this argumentation, determined by using attractive (e.g. location- and profile-based) services compliant with privacy settings. In a tourism context, there are several such services imaginable. In many cases, a service may be offered free of charge to enhance the tourism experience. Examples are mobile hotel reservation services (Hotel Reservation Service 2008), services provided by local authorities for promoting points-of-interest as well as events and services (State Capital Stuttgart 2008). The same services could also be offered as commercially by charging for the added value they provide. We will now present two use cases in more detail to illustrate the potentials of mobile support in tourist destinations.

4.1.1 Mobile Destination Management

The Aladdin research project funded by the European Union aimed at developing a prototype of a mobile destination management system for tour guides, travelers and local small and medium enterprises (SME) (Aladdin Project 2008). The system provides incoming tour operators with an improved and cost-efficient mobile workspace to support their business processes, allowing small companies to compete with larger incoming tour operators. Furthermore, it allows SMEs located at the tourism destination to offer their services and content in an easy accessible way to potential customers. In addition, tourists can enhance their experience by using the offered services and by obtaining additional information about the tourist destination (Altenhofen et al. 2008). The Aladdin project had a focus on the business processes and on use cases that take place at the destination. Based on this research, different prototypes that support tour guides and travelers have been developed. An evaluation of the Aladdin prototype including field tests in three different destinations showed that users are particularly interested in information on points of interest, routing services, access to information services such as weather and traffic reports, as well as an emergency support (Aladdin Project 2007).

4.1.2 Mobile Community Services

There exists a large variety of mobile communities, which focus on satisfying needs of different types of mobile users, such as mobile gaming or mobile dating. Common to all of them is their reliance on mobile communication technologies for providing value-added services to community members. One
such community that represents an economically powerful sector within the tourism industry is the sports fishing community. Sport fishing enjoys a huge popularity in many industrialized countries and the economic impact of sport fishing itself has been well recognized in recent years (Arlinghaus 2004). The study reports that the total economic impact in Germany alone amounts to more than six billion Euros per year including direct and indirect impacts on the domestic economy\(^1\). Similar estimations have also been reported in the literature for other economies with only small variances for many highly industrialized countries (US Fish and Wildlife Service 2001; Moon and Souter 1995; Schwärzel-Klingenstein et al. 1999). In Arlinghaus (2004) the behavioural pattern of German sport fishers was analyzed and it was revealed that 60 percent of the sample population spend more than half of their total fishing time in foreign countries. Each fishing trip is characterized by different subsequent phases, which are the planning phase, the event itself, and a recollection phase (Arlinghaus et al. 2002). The first and the last phase can easily be conducted without any mobile support. However, the users’ demand for mobile services increases shortly before and during the event, because fishing is a location and event-driven activity. Sport fishers and other stakeholders of the sport fishing industry could benefit from multilateral mobile interactions: Sport fishers share secrets on fishing hotspots and favourite baits only with well selected buddies and many of them communicate their catches immediately to their community members. While looking for new and promising fishing places, sport fishers need to know where to obtain fishing permits and what the terms of usage (catch and release, allowance of living bait, etc.) of certain water courses are. Location-based services could support them in obtaining permits and help targeting advertisements for equipment, hotels and further services. The Norwegian fishing administration of Oslomarka\(^2\) already offers some of these services on their website (Oslomarkas Fiskeadministrasjon 2008), including delivery of fishing permits via SMS. Another use case for mobile support is the determination of the genus and species of the fish caught. The phenotype of fish is heavily related to environmental conditions, like water and food and therefore the same species might look entirely different in varied watercourses (Militz et al. 1984). Thus, sport fishers often see themselves confronted with a fish they cannot easily identify and might violate closing seasons or allowed minimum length by keeping the fish. Traditionally, consulting and gossiping is done between community members who need to be reachable in a reasonable timeframe and need sufficient information for valid decisions. The FishBase initiative (Froese and Pauly 2008) has set up a database which can help to figure out what kind of fish has been caught. The success of this database is due to the broad audience that contributes and enhances the data entries and provides knowledge to the on-site audience.

### 4.2 Service Integration and Benefits

The service platform provides a centralized access point to the mobile communication infrastructure, which can be utilized for emergency management and commercial mobile value-adding service as described in the previous section. It offers basic services on which tourism and emergency service can be build upon. Since the same underlying technology is used for both services, economies of scale have a significant impact on the cost structure of these services. By utilizing the basic services provided by the service platform, a quick development of value adding mobile tourism services can be achieved. Furthermore, the central platform offers service providers from the tourism industry, which is largely driven by small and medium enterprises (Werthner and Ricci 2004), a possibility to offer mobile services without the necessity to maintain the underlying infrastructure. By including tour guides and other employees of the tourism industry into preparedness activities, tourists have a contact point they are familiar with. Furthermore, these persons normally have an increased foreign language competence which is an advantage during crisis (Manoj and Hubenko Baker 2007). Combining emergency activities and mobile community services, e.g. by forming spontaneous first aid teams or

---

\(^1\) This is based on the assumption that all spending is done in the domestic economy.

\(^2\) Recreational Areas around Oslo
providing up-to-date information (Palen et al. 2007), shows on the other hand that preparedness can benefit from commercial services provided on the proposed infrastructure. Integrating these services into an emergency management system could improve users’ familiarity with the system while at the same time offering a perceived value to the customers. Figure 2 illustrates an infrastructural concept for such an integration.

![Figure 2: Service integration](image)

The public sector at the destination can also profit from the service integration. Signalling preparedness to prospective customers (i.e. tourists) has a positive impact on the destination’s image and can therefore positively influence the productivity of the domestic economy (Paraskevas and Arendell 2007). A positive image can contribute to a positive reputation in the long-term and therefore increase the popularity of the destination (Paraskevas and Arendell 2007). Via the sub-role of emergency managers, the public sector has an instrument at hand for optimizing the training and preparedness of emergency forces as well as the overall coordination of safety services. One contribution of the proposed system is that the pool of emergency forces has been enlarged to members of the tourism industry. Due to the vital interest of the tourism industry and community service providers to offer attractive services to their customers, tourists and community members could benefit from a wide range of potential services. Using the same infrastructure for emergency services and commercial services increases the familiarity of users with the system, its messages and provided functionalities. In touristic settings, the system design can be used to form dynamic travel teams for enhancing everyone’s safety and convenience in an unfamiliar environment. The same holds for mobile communities in which single members can benefit from the experiences and willingness to communicate of other community members. Furthermore, having a reliable link to a trusted group of entities may result in many use cases in an increased perceived usefulness of location-based services.

5 LIMITATIONS

First steps on how to combine emergency and commercial services via the same infrastructure have been sketched out. However, a detailed analysis and estimation of the impacts and interdependencies of a jointly used infrastructure on the development and the acceptance of mobile emergency services is still missing. The discussed advantages of this approach are limited by the degree of cooperation which decision makers of emergency services might be willing to allow. Also, usability aspects of such systems, like possible steps to ensure that the users pay attention to incoming messages, need to be further researched. A first step in that direction is to use different acoustic and visual signals for
emergency and commercial services to raise the awareness of the users. Furthermore, the availability of different infrastructural components is a major issue and precautions against breakdowns have to be incorporated into networks. Evaluating the readiness of infrastructures for mobile emergency services remains a challenging subject for further studies.

6 CONCLUSION

In this contribution, we proposed an emergency management system design based on mobile communication infrastructure, which allows the integration of mobile value-adding services to improve users’ familiarity with the emergency system. Following the design science research approach we first derived system requirements from the literature. They comprise (1) system effectiveness, (2) reliability, (3) cost efficiency, (4) smooth service integration, (5) multilateral user interaction, (6) availability and (7) security. Based on these requirements we addressed the introduced problem with a conceptual system design. In addition, we presented exemplary mobile value-adding services that could be integrated into the emergency management system and discussed the benefits to the users of such a system. We conclude that the proposed system design could help to increase users’ familiarity with the system, offer a perceived value to the tourists and create new business opportunities for the tourism industry.

References

Arlinghaus, R., Mehner, T. and Cowx, I.G. (2002). Reconciling traditional inland fisheries management and sustainability in industrialized countries, with emphasis on Europe, Fish and Fisheries, 3 (4), 261-316.


**TAXONOMY DEVELOPMENT IN INFORMATION SYSTEMS: DEVELOPING A TAXONOMY OF MOBILE APPLICATIONS**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Journal:</strong></td>
<td>17th European Conference on Information Systems</td>
</tr>
<tr>
<td><strong>Manuscript ID:</strong></td>
<td>ECIS2009-0308.R1</td>
</tr>
<tr>
<td><strong>Submission Type:</strong></td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td><strong>Keyword:</strong></td>
<td>Mobile systems, Mobile communications, M(obile)-banking / M-commerce, Wireless technology</td>
</tr>
</tbody>
</table>
TAXONOMY DEVELOPMENT IN INFORMATION SYSTEMS: DEVELOPING A TAXONOMY OF MOBILE APPLICATIONS

Nickerson, Robert C., San Francisco State University, College of Business, San Francisco, CA 94132, USA, RNick@sfsu.edu
Varshney, Upkar, Georgia State University, 35 Broad Street, Atlanta 30302-4015, Georgia, USA, uvarshney@gsu.edu
Muntermann, Jan, Goethe-University Frankfurt, Grüneburgplatz 1, 60323 Frankfurt, Germany, muntermann@wiwi.uni-frankfurt.de
Isaac, Henri, Université Paris Dauphine, DRM UMR n° 7088 CNRS, 75775 Paris Cedex 16, France, henri.isaac@dauphine.fr

Abstract
The complexity of the information systems field often lends itself to classification schemes, or taxonomies, which provide ways to understand the similarities and differences among objects under study. Developing a taxonomy, however, is a complex process that is often done in an ad hoc way. This research-in-progress paper uses the design science paradigm to develop a systematic method for taxonomy development in information systems. The method we propose uses an indicator or operational level model that combines both empirical to deductive and deductive to empirical approaches. We evaluate this method by using it to develop a taxonomy of mobile applications, which we have chosen because of their ever-increasing number and variety. The resulting taxonomy contains seven dimensions with fifteen characteristics. We demonstrate the usefulness of this taxonomy by analyzing a range of current and proposed mobile applications. From the results of this analysis we identify combinations of characteristics where applications are missing and thus are candidates for new and potentially useful applications.

Keywords: Taxonomy, Taxonomy Development, Mobile Application, M-Commerce, Mobile Business.

1 INTRODUCTION

A fundamental problem in many disciplines is classifying objects of interest into taxonomies. Biology has studied this problem extensively and developed a number of classification schemes that order the complexity in the living world and provide a foundation for biological research. In the management science and information systems fields, the importance of taxonomies is well recognized. Several taxonomies exist in various domains, including manufacturing services (e.g., Miller and Roth 1994) and decision support systems (e.g., Alter 1977). These taxonomies provide fundamental research foundations in the form of a common domain language in which problems and their solutions can be defined and explored.

Although the process of developing a taxonomy has been studied in a number of disciplines (e.g., Sokal and Sneath 1963 and Eldredge and Cracraft 1980 in biology; Bailey 1994 in the social sciences), little has been written about this process in the field of information systems. A well-conceived method for developing taxonomies in information systems would serve as a basis for developing new taxonomies that could bring order to complex areas and potentially lead to new research directions. The purpose of this research-in-progress paper is to present such a method and to demonstrate its applicability by developing an exemplary taxonomy.

We apply our taxonomy-development method to mobile applications. We have chosen this domain because of its increasing complexity with many mobile applications existing today and new
applications appearing regularly. Users, researchers, and developers need to be able to determine where a new mobile application fits with existing applications in order to determine if it is something entirely new and unique, a significant variation of an existing application, or just a retread of what we already have. A taxonomy of mobile applications would provide a basis for making this determination and could point out voids where new applications might be developed.

This paper is organized as follows. First, we discuss our research approach (section 2). Then we review the literature related to taxonomy development (section 3) before presenting our method for developing taxonomies in information systems (section 4). Next we review existing mobile application taxonomies (section 5) and demonstrate the use of our taxonomy development method to develop a new taxonomy of mobile applications (section 6), which we present in detail (section 7). Finally we show how our mobile application taxonomy can be used to analyze current and future applications (section 8) before summarizing our results and providing suggestions for future research (section 9).

2 RESEARCH APPROACH

This paper is based on the design science research paradigm that aims to address new knowledge about artificial (i.e., man made) objects that are designed to meet certain goals and provide utility to their users (Simon 1969). March and Smith (1995) present four kinds of contributions (artifacts) and two characteristic processes (research activities) that characterize design science research in IS.

The four central research outputs are constructs, models, methods, and instantiations. As a most fundamental artifact type, constructs define a conceptual vocabulary that provides the basis for representations of problem domains or for the construction of models. Models describe the relationship among developed constructs, that is, they describe how things are. Methods describe a set of defined steps to provide a solution to a given task. Finally, instantiations are implementations that operationalize other artifacts. As system architectures or system designs, they demonstrate the feasibility and applicability of the models and methods developed.

The two processes that characterize design science research are artifact building and artifact evaluation. Since utility is of vital importance, design science research aims at developing novel artifacts and suitable evaluations that assess the artifact’s appropriateness to contribute to the problem’s solution (Nunamaker et al. 1991).

We present a method that provides guidance for researchers developing a taxonomy for a specific domain. This method is an artifact that serves as a basis for future design science research, the purpose of which is to develop new taxonomies. Such new taxonomies are artifacts (models) in themselves.

We use the two processes of design science in our research. First we build an artifact (method) for developing new taxonomies. The process of taxonomy development is a complex task because the explanatory power of a taxonomy that constitutes its usefulness to users depends on a well-defined set of dimensions. The method we present guides users on how to define these dimensions.

Second we evaluate the artifact (method) we have built by using it to develop a taxonomy that describes and classifies existing or future objects in a specific domain in a sensible way. The domain we have chosen is that of mobile applications, an important research area in IS (March et al. 2000).

The result of the second step is an artifact (the taxonomy) that is subject to evaluation. Therefore, we evaluate the taxonomy by assessing its efficiency to classify objects in the problem domain of interest (i.e., existing and proposed future mobile applications). Evaluation results then provide evidence if it is necessary to redefine the underlying dimensions of the taxonomy.

3 TAXONOMY DEVELOPMENT

Suitable taxonomies play an important role in research and management because the classification of objects helps researchers and practitioners understand and analyze complex domains. The reduction of
complexity and the identification of similarities and differences among objects are major advantages provided by taxonomies (Bailey 1994). Furthermore, taxonomies enable researchers to study the relationships among objects and, therefore, to hypothesize about these relationships. As a vocabulary of a domain and as a set of defined constructs, taxonomies can add to the IS knowledge base and therefore lay the basis for future research approaches (Hevner et al. 2004; March and Smith 1995).

Developing a taxonomy is a complex process. Biology, with its well-known taxonomy of living organisms, provides some guidance. The traditional Linnaean taxonomy, commonly found in biology textbooks, classifies organisms based on a predefined hierarchy of categories from kingdom to species. Determining where a new organism falls in the taxonomy involves identifying into which classification the organism fits at each level of the hierarchy. Biological taxonomy development also includes phenetics and cladistics. Phenetics, also called numerical taxonomy, involves classifying organisms solely on the basis of their similarity. Characteristics are identified and organisms with similar characteristics are clustered using statistical methods to distinguish them from other organisms (Sokal and Sneath 1963). Cladistics, on the other hand, looks at the evolutionary relationships among organisms, not just their common features (Eldredge and Cracraft 1980). Thus two organisms may be closely related in a cladistic taxonomy because they have a common ancestor even though they do not share certain characteristics, thus putting them in different groups in a phenetic analysis.

Bailey (1994) provides a thorough review of taxonomy development in the social sciences. Bailey makes a distinction between a typology and a taxonomy, saying that the former is derived conceptually or deductively and the later is derived empirically. In the conceptual typology approach, the researcher may propose a typology of categories or types based on a theoretical ideal or model. In the process the researcher could define an ideal type, which Bailey (citing Weber 1949) explains is the “extreme” or “nirvana” of types. The ideal type is used to examine empirical cases in terms of how much they deviate from the ideal. Alternatively, the researcher could define a constructed type, which, as Bailey (citing McKinney 1966) explains, is not the ideal but based on reference to empirical cases. The constructed type is used to examine “exceptions” to the type. The ideal type can be compared to the highest value in a set of data (assuming highest is best) whereas the constructed type can be compared to the mean of the data (Bailey 1994, 23).

In the conceptual taxonomy approach the proposed classification is not based on empirical data, although such data could be brought in toward the end of the process. The empirical approach, on the other hand, starts with data and derives the classification from this data using cluster analysis or other statistical methods (Bailey 1994, 34). The goal is to find similarities among the data and to classify similar objects into the same category. Each category in the resulting taxonomy is called a taxon (plural taxa). Using the concepts from biology, this approach is phenetic.

Bailey (1984) describes the approaches just examined as different levels – conceptual and empirical – of a two-level model. Although researchers can approach classification through either level, he suggests that a common and often more useful approach is to use a three-level model that includes conceptual, empirical, and indicator or operational levels. In this method the researcher has two choices. One is to start with the deductive approach and then to examine empirical cases (deductive to empirical) to see how they fit with the conceptualization. The other choice is to start with empirical data clusters and then to deductively conceptualize the nature of each cluster (empirical to deductive).

4 AN APPROACH TO TAXONOMY DEVELOPMENT IN INFORMATION SYSTEMS

We now present our approach to taxonomy development for objects in the information systems field. Following the design science paradigm, we are building an artifact that is a method.

We choose to use the term taxonomy for our classification because it is more common and recognizable than the term typology, although we recognize that the later may be more correct in some situations. We define a taxonomy $T$ as a set of $n$ dimensions $D_i$ ($i=1, \ldots, n$) each consisting of $k_i$
mutually exclusive and collectively exhaustive characteristics $C_{ij}$ ($j=1, \ldots, k_i$) such that each object under consideration has one and only one $C_{ij}$ for each $D_i$. The goal is to develop a taxonomy with a set of dimension each consisting of a set of characteristics that sufficiently describes the objects in a specific domain of the information systems field.

Our purpose is to develop a useful taxonomy, but not necessarily a best or optimal one, as the later may be a moving target that could change over time as information systems evolve. A review of the literature found little help identifying metrics for evaluating taxonomies. Indeed, Bailey (1994, p. 2) makes this clear when he repeatedly asks which of his example classifications is “best” without giving guidance for finding the answer other than saying that “a classification is no better than the dimensions or variables on which it is based.” Thus we are left on our own to define a useful taxonomy. We proposed that a useful taxonomy has the following desirable attributes:

- It should be concise. It should contain a limited number of dimensions or a limited number of characteristics in each dimension, because an extensive classification scheme with many dimensions and many characteristics would be difficult to comprehend and difficult to apply.
- It should be sufficiently inclusive. It should contain enough dimensions and characteristics to be of interest. For example, a taxonomy with only one dimension and two characteristics within that dimension would not be very interesting. This attribute can conflict with the conciseness attribute.
- It should be comprehensive. It should provide for classification of all current objects within the domain under consideration.
- It should be extendible. It should allow for additional dimensions and new characteristics within a dimension when new types of objects appear.

Our approach to developing a taxonomy is phenetic, looking at the characteristics of the objects being examined. We do not look at the evolutionary development of the objects, and thus our approach is not cladistic. Our approach uses Bailey's three-level indicator model in which we combine both the empirical to deductive approach and the deductive to empirical approach.

Before starting to develop a taxonomy, the researcher must decide on the most comprehensive or meta-characteristic that will serve as a basis for the classification. The choice of the meta-characteristic should be based on the purpose of the taxonomy. For example, assume that the researcher is trying to classify computer platforms (hardware and operating system) into a taxonomy. If the researcher's purpose is to distinguish platforms based on processing power, then the meta-characteristic is the hardware and software characteristics, such as CPU power, memory, and operating system efficiency, that impact measures of power such as speed and capacity. On the other hand if the researcher's purpose is to distinguish among computer platforms based on how users use them, then the meta-characteristic is the capability of the platform to interact with users, such as the number of simultaneous users and user interface. The choice of the meta-characteristic must be done carefully as it impacts critically the resulting taxonomy. Further investigation of the selection process is needed.

We see meta-characteristics appearing research that develops taxonomies for various purposes, although they are not identified as such. For example, Nickerson (1997) develops a taxonomy of collaborative applications based on the meta-characteristic of communication among group members. Williams, Chatterjee, and Rossi (2008) choose two meta-characteristics – design and objectives – in developing their taxonomy of digital services. Leem, Suh, and Kim (2004) develop a classification scheme for mobile business models based starting with the meta-characteristic of “business players.”

We find Bailey's (1984) three-level indicator model to be most appealing for the development of taxonomies in the information systems field. It does not take a single approach but relies on both deduction and empiricism. We do not propose that our approach is the best or only taxonomy development method in the information systems field, only that it follows from the taxonomy development literature and is practical.
Figure 1 shows the approach that we propose based on Bailey’s model. The researcher begins by examining a subset of objects that he/she wishes to classify. These objects are likely to be the ones with which the researcher is most familiar or that are most easily accessible, possibly through a review of the literature. Next the researcher identifies general characteristics of these objects. In what ways are objects similar? What distinguishes objects from each other? Identification of these characteristics leads to the first effort at a taxonomy. The characteristics are grouped into dimensions that form the initial taxonomy. Each dimension contains characteristics that are mutually exclusive and collectively exhaustive. For example, dimension D₁ may group characteristics C₁₁, C₁₂, and C₁₃ and dimension D₂ may group characteristics C₂₁ and C₂₂. All objects have one and only one of the characteristics C₁j in dimension D₁ and one and only of the characteristics C₂j in dimension D₂. Some dimensions may be dichotomous (e.g., D₂) and some may not be (e.g., D₁). This process is based on the (limited) empirical data that has been gathered about objects and the deductive conceptualization of the researcher. Thus, up to this point the process follows Bailey's empirical to deductive approach.

Now the researcher reviews the first taxonomy to look for additional conceptualizations that might not have been identified or even present in the original empirical data. In the process new characteristics may be deduced that fit into existing dimensions or new dimensions may be conceptualized each with their own set of characteristics. It may even be the case that some dimension or characteristics are combined so that fewer dimension and/or characteristics result. The researcher examines empirical cases using the new characteristics and dimensions to determine their usefulness in classifying objects. Out of this step comes a revised taxonomy. This process follows Bailey's deductive to empirical approach. The researcher then repeats the empirical to deductive and deductive to empirical approaches, as appropriate, until the researcher is sufficiently satisfied that the taxonomy has the attributes of conciseness, sufficient inclusiveness, comprehensiveness, and extendibility. Such closure is subjective and difficult to define; further investigation is needed to clarify it.

Figure 1. The taxonomy development method
Figure 1 also shows how the taxonomy could be used after it is completed. One use could be to identify missing objects in the taxonomy. For example, there may be no objects that have certain characteristics, such as C_{12} in dimension D_1 and C_{21} in dimension D_2. There are several reasons for such a situation. The characteristic or combination of characteristics might be undesirable for the objects under consideration or it might be difficult or infeasible to create an object with the characteristic(s). Alternatively, no one may have thought of creating an object with the specific characteristic(s). In this situation, researchers or practitioners may try to create objects with the missing characteristics, thus expanding the set of objects in the domain. If this process continues long enough new objects may start to appear that do not fit in the taxonomy resulting in the need to start the taxonomy development process over again from the beginning.

5 MOBILE APPLICATION TAXONOMIES

Before using our method to develop a taxonomy of mobile applications we review the literature on the subject. Unfortunately, in mobile commerce and mobile business there exist little research on taxonomies. Only a few taxonomies have been proposed (e.g. Lehmann and Lehner 2002, Okazaki 2005), and there is a lack of a general taxonomy of mobile applications.

Kemper and Wolf (2002) propose a taxonomy based on a three-dimensional classification scheme. The chosen dimensions are degree of innovation, speed of development, and risk, and a set of characteristics focused on mobile application development. The taxonomy is narrowed down to the development process and does not feature the specific characteristics of mobile applications.

Leem et al. (2004) develop a hierarchical classification scheme based on mobile business models. In a first step, they partition into the two dimensions of B2C and B2B/B2E business models. They subdivide each of these dimensions into further dimensions. Their approach enables practitioners and researchers to classify mobile applications from a business model perspective. Thus it can support managerial decision-making and be a basis for mobile business model research. Their approach, however, focuses on a specific perspective and does present general mobile application classifications.

A basic two-dimensional classification scheme of mobile services is suggested by Nysveen and Thorbjørnsen (2005). They group services by type of interactivity and process characteristics. Both dimensions feature binary categories (person- vs. machine-interactive and goal-directed vs. experimental) only. The resulting fourfold scheme provides limited descriptive power. Williams, Chatterjee, and Rossi (2008) also discuss a taxonomy of digital services, but not just mobile services.

Heinonen and Pura (2006) developed a classification scheme for mobile services from a customer-centric perspective that is based on the four dimensions of type of consumption, context, social setting and relationship. Identifying industry specific classification as a limitation of their taxonomy, they suggest further research towards a more general taxonomy of mobile applications.

Dombrovskia and Ramnath (2007) present a taxonomy that focuses on applications they call "mobile, pervasive," which they characterize as "an application integrated with its physical environment and/or aware of its location." Their taxonomy consists of six general dimension – transitionality, time constraints, goal, collaboration, lifetime, and centricity – and seven location and location-related awareness dimensions – absolute location, space, proximity, transition, event, object, and operational.

Finally, Dobson (2004) does not classify applications per se, but instead provides a detailed classification of location in pervasive computing. He gives seventeen types of answers to the question of where someone or something is, ranging from absolute position to unknown, and organizes these answers into hierarchical taxonomy of location.
6 DEVELOPMENT OF A TAXONOMY OF MOBILE APPLICATIONS

With this background we now apply the approach described previously to develop a taxonomy of mobile applications. Again, using the design science paradigm we are validating the artifact (method) built previously by using it to develop another artifact, specifically a model.

We define a mobile application as a use of a mobile technology by an end-user for a particular purpose, e.g., purchase a ring tone, check a weather forecast, transfer funds at a bank, make an airline reservation. Mobile applications are provided by mobile services that have the infrastructure necessary to deliver the application. A mobile service, however, may provide several different applications under the umbrella of one service. For example, a mobile service may provide information about popular music and sell MP3 music files. For this paper we view these as two different applications – one, an informational application, and the other, a transactional application – both provided by one service.

We are interested in the use of applications, not in their hardware/software characteristics such as type of mobile device used or speed of network connection. We want to be able to use our taxonomy to identify how users use applications currently or may use applications in the future. Specifically, the purpose of our taxonomy is to distinguish among mobile applications based on how the user interacts with the application. Such a taxonomy will be able to help us identify whether new applications are truly unique from the user's perspective and where applications do not exist in the taxonomy suggesting opportunities for new applications. Thus the meta-characteristic for our taxonomy development process is the interaction between the user and the application. We start by listing a number generic mobile applications derived from the literature (Varshney and Vetter 2002, Varshney 2007). We note that this list is not exhaustive.

- Mobile communications
- Mobile messaging
- Mobile navigation
- Mobile TV
- Purchasing location-based contents
- Mobile inventory management for a company
- Product location and tracking for individuals
- Mobile auctions & financial services
- Mobile games – individual
- Mobile games – group
- Mobile advertisement user – specific
- Mobile advertising – location-specific
- Mobile entertainment services (contents-on-demand, live events)
- Mobile personal services (mobile dating)
- Mobile distance education – offline
- Mobile distance education – online
- Mobile product recommendation
- Mobile telemedicine and patient monitoring
- Mobile ticketing
- Mobile communities
- Mobile emergency/safety management
- Mobile habitat/environmental monitoring
- Mobile social networking

We begin by noting certain user interaction characteristics of these applications. With some applications users interact synchronously (in real time) with information flowing from the application to the user and with the user engaging in financial transactions. An example of this type of application is purchasing location-based contents such as weather information. With other applications users interact asynchronously (not in real time), they only report information to the application, and they do not engage in a financial transaction. An example is mobile inventory management. We continue by finding other applications with these characteristics. Once we gain confidence that we have identified some distinguishing characteristics, we group these characteristics into dimensions as follows:

- Temporal dimension: synchronous and asynchronous characteristics
- Communication dimension: informational and reporting characteristics
- Transaction dimension: transactional and non-transactional characteristics

At this point we have our first taxonomy, which we have derived using an empirical to deductive approach. Reviewing the taxonomy we deduce that the communication dimension may also have a bi-directional characteristic in which information flows from the application to the user and from the user...
to the application. We find that some applications (e.g., mobile auctions) have this communication characteristic. We call this characteristic interactional and add it to the communication dimension:

- Communication dimension: informational, reporting, and interactional characteristics

We have modified our taxonomy using a deductive to empirical approach. We continue in this way to develop newer versions of the taxonomy, going either from empirical cases to deduction or deduction to empirical cases. For example, we speculate that some applications can interact with anyone, that is, they are public, and some applications can only be used by individuals who have certain privileges such as those who work for a company, that is, they are private. We identify instances of these types of application. For example, purchasing location-based contents is a public application and mobile inventory management is a private application. Thus we add another dimension to the taxonomy:

- Public dimension: public and private characteristics

This new dimension creates the next version of our taxonomy. We continue in this way until we are satisfied that we have identified dimensions with characteristics for a taxonomy with the attributes of conciseness, sufficient inclusiveness, comprehensiveness, and extendibility. Deciding when to end the taxonomy development process is highly subjective.

Rather than continuing with a detailed description of the development process we present in the next section our final taxonomy of mobile applications that we derived by following this iterative process.

7 A PROPOSED TAXONOMY OF MOBILE APPLICATIONS

Our taxonomy of mobile applications is based on the meta-characteristic of the interaction between the user and the application. It consists of seven dimensions, some of which have already been described. We summarize all the dimensions and their characteristics here. For another discussion of these dimensions see Nickerson, Varshney, Muntermann, and Isaac (2007).

Temporal dimension. The user can interact with some mobile applications in real time, meaning that the application services the user’s request almost immediately, whereas the interaction between the user and the application may be deferred in other applications. The temporal dimension identifies when the user and the application interact. It has the following characteristics:

- Synchronous: user and application interact in real time
- Asynchronous: user and application interact in non-real time

Communication dimension. Information may flow uni-directionally between the user and the application or bi-directionally. The communication dimension relates to which way information flows as the user interacts with the application. Its characteristics are as follows:

- Informational: information flows only from the mobile application to the user; uni-directional information flow to the user; information push from the application to the user
- Reporting: information flows only from the user to the mobile application; uni-directional flow from the user; information pull by the application from the user
- Interactional: information flows in both directions between the user and the mobile application; bi-directional flow between user and application; information push and pull

Transaction dimension: Some mobile applications all users to purchase goods or services, normally through a financial transaction, while others do not. The transaction dimension captures this characteristic of the user interaction. This dimension has the following characteristics:

- Transactional: user can purchase goods or services through the application
- Non-transactional: user cannot purchase goods and services through the application

Public dimension: Mobile applications may be available to the general public, or their use may be limited to members of specific groups, such as certain employees of a business. The public dimension relates to whether the application is generally available. Its characteristics are the following:
• Public: application can be used by any user; may be limited to a group but any user may self-select to be part of the group that uses the application
• Private: application can only be used by a pre-selected (by a third party) group of users

Multiplicity (or participation) dimension: Although mobile applications can be used by many users simultaneously, users are often not aware of this characteristic and view their use of the application as singular. With some applications, such as multiple user mobile games, the user knows that he or she is part of a multiple-user community using the application. The multiplicity dimension captures this concept of individual or multiple user interaction with the following characteristics:
• Individual: one user; user experiences the application as if he/she were the sole user
• Group: multiple users; users view use of the application as part of a group

Location dimension: Some mobile applications provided customized information or functionality based on the user’s location, whereas other applications do not depend on where the user is located. The location dimension deals with whether the location of the user is used to modify the interaction of the application with the user. It has the following characteristics:
• Location-based: mobile application uses the user’s location
• Non-location-based: mobile application does not use the user’s location; the mobile application may know the user’s location but it does not use this knowledge to modify the user interaction

Identity dimension: Like the location dimension, some mobile applications adjust their information or functionality based on an awareness of who the user is, whereas other applications do not depend on the user’s identity. The identity dimension relates to whether the identity of the user is used to modify the way the application interacts with the user application based on the user’s identity. This dimension has the following characteristics:
• Identity-based: mobile application uses the user’s identity
• Non-identity-based: mobile application does not use the user’s identity; the mobile application may know the user’s identity but it does not use this knowledge to modify the user interaction

8 USE OF MOBILE APPLICATION TAXONOMY

The final steps in our taxonomy development process are to use the taxonomy to analyze objects of interest. Table 1 lists the generic mobile applications identified earlier (not an exhaustive list) and the characteristics of each application using the dimensions we have identified in our taxonomy. The decisions to identify characteristics for an application in this table are based on our understanding of each application. It is possible that variations of the applications may exist or become available in the future with different characteristics. Such variations can be easily added along with new applications.

We observe several things from this table:
1. There are a large number of generic mobile applications and some applications exist in more than one variation to suit different requirements.
2. The use of seven dimensions and fifteen characteristics appears to be sufficient to describe an application. These dimensions and characteristics also are helpful in differentiating among or classifying diverse applications.
3. There are an approximately equal number of synchronous and asynchronous applications in this list. In the future, new applications could emerge that will be “preferred” to run in synchronous mode, but if network infrastructure is experiencing high traffic load, these applications could adjust to run in asynchronous mode.
4. Only two applications in this list are reporting. Although the current needs of the user for reporting may be met via PC or laptop, it is expected that such requirements will move to handheld devices as many other applications have moved in the past. Thus more research could be done in designing mobile applications that have the reporting characteristic. Many of the reporting applications may be followed by user actions supported by one or more other mobile applications.
5. There is a good balance of transactional and non-transactional applications in this list.
Table 1. Characteristics of selected generic mobile applications

6. There are only four private applications in this list, although several applications can be modified to limit them to private users, including mobile distance education. The overwhelming number of public applications could be because most mobile applications are B2C at this time. As more mobile B2B or B2E applications are developed, there are likely to be an increase in applications with the private characteristic.

7. There are six group applications in this list. There may be an opportunity, however, for more new mobile group applications. It should be noted that most of the wireless infrastructure is not designed to support group-oriented applications due to resource and bandwidth limitations and the difficulty in supporting group mobility. Also, the resource and networking requirements may grow non-linearly with an increase in the group size. These factors may hinder the design and development of group-oriented mobile applications.

8. There is a good balance of location and non-location based applications in this list. In practice, the range of location capability varies among wireless networks, from no location management to
some location management in some places to highly accurate location management. With increased use of location technologies such as GPS, RFID, and sensor networks, there will be more infrastructure support for location-based applications.

9. There are fewer non-identity based, or anonymous, applications than identity based applications in this list. More work should be done in designing non-identity based applications. More specifically, some people might be more comfortable conducting transactions and interactions for services/contents if their identity is not known or kept secret.

These observations are based only on the listed applications, which is not exhaustive. As more applications are analyzed using the taxonomy, additional observations may be made. Also work can be done in studying the demographics of the users of these applications and designing applications to suit different age groups with varying requirements. Most of the mobile applications are designed for working adults, but there are opportunities to design new applications for young adults, older adults, and geriatric populations. For example, an application could be developed to fill a void identified in the taxonomy that is a reporting, non-transactional, private, group application for older adults. This application would allow users to report wellness and interesting tidbits from their daily life to a group of geriatric friends who may not be able to meet face-to-face often.

9 SUMMARY AND CONCLUSION

This paper describes a method of taxonomy development that can be used in the information systems field. The method, developed by following the design science paradigm and based on the taxonomy development literature, uses an indicator model that combines both empirical to deductive and deductive to empirical approaches. The process is iterative and terminates when the researcher is satisfied that the taxonomy has the attributes of conciseness, sufficient inclusiveness, comprehensiveness, and extendibility.

We evaluate the method by developing a taxonomy of mobile applications with seven dimensions – temporal, communication, transaction, public, multiplicity, location, and identity – and fifteen characteristics. We show that the dimensions and characteristics in this taxonomy are useful by analyzing some current and proposed mobile applications. Using this taxonomy, we identify that there are fewer applications among those considered that have the characteristics of reporting, private, group, and non-identity based, and we suggest that new applications can be designed.

The application of our method to the development of a mobile application taxonomy demonstrates that the method can be a useful approach to taxonomy development. We note, however, that one test of a method is not sufficient to show its ultimate usefulness and that further tests are needed. Our analysis of mobile applications using the mobile application taxonomy shows that the taxonomy resulting from our method is a useful tool for analyzing current and future mobile applications.

Further research in this area can follow several paths. One is to continue to explore the effectiveness of the proposed taxonomy development method by applying it to other domains, possibly resulting in improvements in the method. The second path is to refine the mobile application taxonomy by adding, deleting, changing, or combining dimensions, and to continue to test the taxonomy's efficacy by categorizing more mobile applications. A final research path is to use the results of the analysis in Table 1 or similar tables derived by applying the taxonomy to identify voids that could lead to the design and development of potentially more suitable mobile applications or the redesign of existing applications to better support user needs. The work can be extended in the future by elaborating more on user needs and the current context. We continue to explore all these avenues of research.

10 REFERENCES


EMERGING BUSINESS MODELS AND STRATEGIES FOR MOBILE MIDDLEWARE TECHNOLOGY PROVIDERS: A REFERENCE FRAMEWORK

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0494.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Models, Mobile communications, Strategy, Multiple case studies</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
EMERGING BUSINESS MODELS AND STRATEGIES FOR MOBILE MIDDLEWARE TECHNOLOGY PROVIDERS: A REFERENCE FRAMEWORK

Ghezzi, Antonio, Politecnico di Milano, Department of Management, Economics and Industrial Engineering, Piazza Leonardo da Vinci 32, 20133 Milan, Italy, antonio1.ghezzi@polimi.it

Abstract

As Mobile Network Operators are turning their attention to value added services, the need for innovative technology platforms designed for mobile digital contents management becomes evident. Such phenomenon is enhancing the strategic relevance of the “Mobile Middleware Technology Providers” (MMTPs) within the Mobile Content Value Network. The purpose of this paper is to explore which are the most critical choices to be made at a business model design level for a MMTP, to understand how these parameters are interrelated and can be combined to give rise to differential business models, and finally to delineate what are the most significant underlying “strategic patterns” driving the first steps of MMTPs activity within the Mobile Content competitive arena. The research relies on the adoption of a multiple case studies methodology: through 72 semi-structured interviews, 24 MMTPs were analyzed. The research findings show that some key business model parameters identified by the existing literature can be applied to MMTPs’ business model design process, while others were missing or not made explicit. Moreover, three noteworthy business models currently adopted by MMTPs – “Pure Play”, “Full Asset” and “Platform & Content Management” business models – were identified, associated respectively to three underlying strategic patterns – “stay on core”, “grow, wait and see” and “aggressive downstream”.

Keywords: Mobile Communications, Business Model, Strategy, Multiple Case Studies.
INTRODUCTION

Forced to face the gradual leveling off of voice revenues (Nomura, 2005; Arthur D. Little/BNP Paribas, 2005) that lead to a subsequent decrease of Average Revenue per User (Muller-Veerse, 1999; MacKenzie, O’Loughlin, 2000), Mobile Network Operators (MNOs) are to cope with a new dilemma: how to generate revenues for sustaining their future growth. The answer seems to come from the development of a wide and appealing offer of value added, non-voice services, pertaining to the so-called Mobile Content segment (Peppard, Rylander, 2006; Kuo, Yu, 2006; Maitland et al., 2002; Li, Whalley, 2002; Noordman, 2006).

However, the strategic reorientation of MNOs will not be straightforward, and won’t take place overnight. Specifically, on the technology architecture level, MNOs will need to introduce new solutions capable of overcoming the constraints and limitations of legacy systems and of the oversimplified Short Message Service Centers, not suitable for providing carrier-grade performances when dealing with “rich media” digital contents. Such solutions are here named “Mobile Content and Service Delivery Platforms” (MCSDPs), and can be defined as middleware platforms combining a wide set of functionalities – consistently aggregated into different modules, and equipped with network-side and device-side interfaces, thus creating an integrated suite with the purpose of supporting some or each phase of the mobile digital content creation, management & delivery process.

The diffusion of second generation delivery platforms will enhance the strategic relevance of a new player typology: the platform supplier, from now on referred to as “Mobile Middleware Technology Provider” (MMTP). Such players are converging in the Mobile Content market from several neighboring business areas, and their moves can reshape Mobile Content’s Value Network, potentially determining unexpected competitive attritions between these new players and incumbents.

These new competitive dynamics deserve attention from both researchers and practitioners. In particular, questions arise concerning the strategies Mobile Middleware Technology Providers will elaborate to compete in the market, and the business models they will hence design and adopt.

The purpose of this paper is to explore which are the most critical choices – i.e. parameters or “building blocks” – to be made at a business model design level for a MMTP, to understand how such parameters are interrelated and can be combined to give rise to differential business models, and finally to delineate what are the most significant underlying strategies or “strategic patterns” that seem to be driving the first steps of MMTPs activity within the Mobile Content competitive arena.

As a result, a reference model will be created, whose main objective is to provide a description of the key parameters characterizing MMTPs’ business models, to identify the extreme values such variables can assume, and to evaluate and assess the strategic implications of each building block choice. Moreover, the main existing combinations of parameters, which create the business models currently employed by this typology of companies, will be analyzed and interpreted, so to make some inferences regarding the relative overall strategies.

LITERATURE REVIEW

2.1 Mobile Middleware Technology Providers Definition

The literature dealing with technology enablers for Mobile Value Network is quite fragmented, and fails to provide a clear and unified definition of Mobile Middleware Technology Providers. Moreover, such players are also associated with several different sets of roles – i.e. set of distinct value added activities covered within a value system.
This lack of homogeneity in definitions is mainly due to the current complexity characterizing the Mobile Content Value Network itself, which results from the juxtaposition of different major value chains and systems, classifiable as follows: Network transport; Applications operation; Content provisioning; Payment processing; Providing device solutions; Network equipment provisioning; Middleware/platform provisioning (Yankee Group, 2000). As a consequence of the different points of view taken, different definitions and roles arise for Mobile Middleware Technology Providers.

Focusing on the activities strictly related to creation, management & delivery of mobile digital contents, the Value Network here proposed is composed by two parallel but interconnected layers – consistently with the “layered architecture” concept introduced by Huemer (2006):

1. **Content & Service Layer**, covering the activities related to the lifecycle management of mobile digital contents and services;

2. **Platform Layer**, undercurrent to the previous layer, which comprises the activities of designing, producing and operating the middleware platforms destined for mobile contents management and delivery.

![Mobile Digital Content & Service Value Network](image)

**Figure 1. Mobile Digital Content & Service Value Network**

The interconnection between the layers becomes evident with the activity of Content Publishing on the MCSDP. The Content & Service Layer can be divided into an “upstream chain”, encompassing the activities from content creation to its preparation for delivery, and a “downstream chain” considering the stages following the content commercialization.

The main focus of MMTPs resides within the Platform Layer: the middleware technology enablers are active in MCSDP design, manufacturing, provisioning – i.e. supplying the platform to the customers, mainly MNOs and/or Mobile Content & Service Providers (MCSPs), operation – i.e. platform technical maintenance and upgrading, and management – i.e. overall handling of the platform’s functionalities, from content publishing to physical distribution, exclusively from a technological point of view; marketing and selling activities are therefore excluded from this area, and belong to the “Content Delivery & Market Making” activity. Nevertheless, an extension of the MMTPs domain to include one or many overcurrent activities may be plausible: such alternative positioning, deriving
from specific choices made at a strategy definition level, would however potentially determine a competition between MMTPs and MCSPs. The strategic implications of this scenario will be discussed later.

As a result of the Value Network model presented above, and given the range of activities topped by this typology of players, a unified and unambiguous definition of the player typology under scrutiny can be offered, thus filling the existing literature gap: Mobile Middleware Technology Providers players are traditionally positioned on the Platform Layer – the technology enabling Value Chain for Mobile Content market, and their core role encompasses some or each activities related to the development of middleware Mobile Content and Service Delivery Platforms.

2.2 Business modelling design parameters

The concept of business model generally refers to the “architecture of a business” or the way firms structure their activities in order to create and capture value (Timmers, 1998; Rappa, 2000; Weil, Vitale, 2001; Hawkins, 2001). As a literature stream, Business model design has evolved from a piecemeal approach that looked for the single identification of typologies or taxonomies of models, to one searching for the development of a clear and unambiguous ontology – that is, the definition of the basic concepts of a theory – (Osterwalder, 2004), that could be employed as a generalized tool for supporting strategy analysis on firms. In parallel, business model has become an extensive and dynamic concept, as its focus shifted from the single firm to the network of firms, and from the sole firm’s positioning within the network to its entire interrelations and hierarchies (Ballon, 2007).

What is widely accepted by the literature is that a business model shall be analyzed through a multi-category approach, being a combination of multiple design dimension, elements or building blocks. However, the proposed dimensions are quite diverse, and the existing body of knowledge shows a lack of homogeneity.

Noteworthy attempts of providing a unified and consistent framework can be found in Rappa (2001), Weil and Vitale (2001), Osterwalder (2004), Haaker et al. (2004) and Ballon (2007) – this last study showing specific focus on Mobile Telecommunication Industry. The recurrent parameters of their models can be brought back to the general concepts of “Value”, i.e. the way a firm creates actual benefits to its customers and to itself through its value proposition and financial configuration, and “Control”, i.e. the inter-firms or Value Network relationships the firm is involved in and controls over.

The literature review on business model design allowed to individuate a further literature gap: as the Mobile Content segment is a relatively young market, and as the “advent” of MMTPs within such market’s boundaries is an extremely recent phenomenon, only few consolidated theories on strategy creation and business model design in the market context and with reference to the specific player typology under consideration are present.

Therefore, starting from the existing literature on business model design, and taking into account the building blocks so far pinned down, this research attempts to identify the key business model parameters for MMTPs, and to assess the strategic implications of the “parameters mix” actually employed by these players operating in the Mobile Content market.

3 RESEARCH METHODOLOGY

The present research is based on case studies, defined by Yin (2003) as “empirical inquiries that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident; and in which multiple sources of evidence are used”. Qualitative research methodology was chosen as particularly suitable for reaching the research objectives, which aim at understanding the complex phenomenon of business model design
development within a given industry – i.e. Mobile Content – and with reference to a specific typology of players – MMTPs –, and thus at building new theory – or extending existing theories – on it.

To accomplish the previously identified research propositions, 24 in-depth exploratory case studies on MMTPs were performed. Coherently to the research methodology employed (Pettigrew, 1988), the firm sample was not randomly selected, but firms were picked as they conformed to the main requirement of the study, while representing both similarities and differences considered relevant for the data analysis. The main predetermined filters used to discriminate among firms were: the international reach of the firm – assumed if at least two national markets were served –; the presence of a well-defined line of business – if not the core business – dedicated to the commercialization of Content and Service Delivery Platforms or MCSDP modules; and the presence of an offer directed to the Mobile Telecommunications market. The following table provides the full list of analyzed companies.

<table>
<thead>
<tr>
<th>Sample of companies</th>
<th>Sample of companies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcatel-Lucent</td>
<td>Fabbrica Digitale</td>
</tr>
<tr>
<td>Bea Systems</td>
<td>First Hop</td>
</tr>
<tr>
<td>Beeweb</td>
<td>HP</td>
</tr>
<tr>
<td>Converse</td>
<td>IBM</td>
</tr>
<tr>
<td>Dylologic</td>
<td>LogicaCMG/Acision</td>
</tr>
<tr>
<td>Ericsson</td>
<td>Mblox</td>
</tr>
<tr>
<td>Microsoft</td>
<td>Nec</td>
</tr>
<tr>
<td>IBM</td>
<td>Neodata</td>
</tr>
<tr>
<td>LogicaCMG/Acision</td>
<td>Openwave</td>
</tr>
<tr>
<td>Mblox</td>
<td>Polarix</td>
</tr>
<tr>
<td>Microsoft</td>
<td>Qualcomm</td>
</tr>
<tr>
<td>Nec</td>
<td>Reitek</td>
</tr>
<tr>
<td>Neodata</td>
<td>Reply</td>
</tr>
<tr>
<td>Nokia-Siemens Networks</td>
<td>Sybase 365</td>
</tr>
<tr>
<td>Openwave</td>
<td>TXT Polymedia</td>
</tr>
<tr>
<td>Polarix</td>
<td>Xiam Technologies</td>
</tr>
</tbody>
</table>

| Table 1. Theoretical sample of companies interviewed |

A multiple case study approach reinforced the generalizability of results (Meredith, 1998), and allowed to perform a cross analysis on parameters, to pinpoint differentials in terms of parameters combination – to see which variables changed and which remained constant going from one business model to another –, due to the presence of extreme cases, polar types or niche situations within the theoretical sample (Meredith, 1998). The unit of analysis for each case study were the set of decision made at a business model design level.

As the validity and reliability of case studies rest heavily on the correctness of the information provided by the interviewees and can be assured by using multiple sources or “looking at data in multiple ways” (Eisenhardt, 1989; Yin, 2003), multiple sources of evidences or research methods were employed: interviews – to be considered the primary data source –, analysis of internal documents, study of secondary sources – research reports, websites, newsletters, white papers, databases, international conferences proceedings –. This combination of sources allowed to obtain “data triangulation”, essential for assuring rigorous results in qualitative research (Bonoma, 1985).

From January to July, 2008, 72 semi-structured interviews – both face-to-face and phone interviews – were held with 65 persons identified as key participants in the firms’ strategy definition and business model design processes at different levels. The population of informants included top and middle managers – e.g. Presidents Chief Executive Officers, Chief Information Officers, Chief Financial Officers, Marketing & Sales Managers, Project Manager, Software Engineers and Developers –. The semi-structured nature of the questionnaire made possible to start from some key issues identified through the literature, but also to let innovative issues emerge.

Given the explorative nature of the study, the business model variables identified through the literature analysis only constituted a starting point to guide the interviews: the identification of core business models parameters and the disentanglement of their combinations to create a thorough business model will represent a key finding of the present research.
4 MMTP BUSINESS MODEL CORE PARAMETERS

The research carried out through the multiple case studies allowed to shed light on the core business model design parameters for Mobile Middleware Technology Providers. The findings are synthesized in the “MMTP Business Model Parameter Reference Framework” below provided, which identifies three macro-dimensions, in turn divided into 9 parameters.

1. **Value Proposition parameters.** Platform characteristics; Offer positioning; Platform provisioning; Additional services; Resources & competencies.

2. **Value Network parameters.** Vertical integration; Customer ownership.

3. **Financial Configuration parameters.** Revenue model; Cost model.

As it will become clear by analyzing the framework, some building blocks were borrowed by previous models – in particular, Ballon (2007), while others, as not present in the existing literature or not made explicit, were modified or originally created through the empirical research to better express some aspects strictly linked to MMTPs.

For each and every parameter, the “value range” is identified, i.e. the extremes values the variables can assume, which also represent the major trade-off between opposite choices; the main strategic implications deriving from the parameters adoption are also discussed.

- **Platform characteristics.**

  As the MCSDP is the core element of MMTPs’ value proposition, its characteristics are a key parameter to be modeled, for they strongly affect the firm positioning. The main alternatives here are developing a modular and interoperable solution versus an integrated and stand-alone system. Should the platform be modular and interoperable, it would allow an easier and faster market diffusion – such choice being advisable for new entrants, searching for quick consolidation within the market (Blind 2005), thanks to the access to a wider customer base; however, to modularity and interoperability is often associated the risk of easy substitutability. In addition to this, a higher modularity and interoperability of MCSDPs can also give rise to interesting “co-opetition” – coexistence of “cooperation” and “competition” – phenomena among MMTPs, where competitors on a project/product can be partners for the modular development of a different project/product. On the other hand, providing an integrated and scarcely interoperable platform slows down the market penetration process, but if the solution is adopted by MNOs or MCSPs, it strengthens the ties between the customers and the technology supplier, potentially generating lock-in effects.

- **Offer positioning**

  Offer positioning is related to the choice of developing a MCSDP devoted to the management & delivery of “mature” contents – Sms, Mms, logos, wallpapers, ringtones and so on (Bertelè et al., 2008), or meant to deal with more innovative and cutting edge services – like video services or Mobile Tv. While operating in traditional segments grants faster platform diffusion, but forces the MMTP to face a higher level of competition – with a risk of seeing a gradual “commoditization” of its products, the coverage of forefront areas could position the firm in attractive niches, but may even imply higher demand risks, as the uptake of such services is hardly predictable.

- **Platform provisioning**

  The MCSDP provision modality is an emergent parameter, particularly interesting in the case of MMTPs, as it influences the kind of relation the technology supplier creates with its business customers. Installation in MNOs’ or MCSPs’ house is a typical choice for standard, out-of-the-box solutions which only need parametrization, and implies both an increased technical independence on MNO/MCSP side, and a clear separation between customer-supplier businesses. A particular case of housing is represented by the choice of full outsourcing – coming from a cross-fertilization of the MCSDP market from the IT platform and System Integration markets, where this practice is
widespread; in this alternative, the MMTP physically installs the platform within the customer’s structure, and thoroughly takes on its technical management. On the contrary, the hosting or Application Service Provisioning (ASP) option sees the MMTP maintaining the core platform within its perimeter, and supplying it to its customer following the “software as a service” model: this allows the technology provider to keep a greater presidium on the platform, and to exploit both scale and scope economies on the platform provisioning infrastructure.

- **Additional services**
  Another original parameter for MMTP business model design, additional services refers to the complementary offer accompanying the MCSDP selling, which can range from a simple technological management of the platform’s operation – e.g. maintenance, upgrading etc. – to, in some rare case, as discussed in the next paragraph, a commercial management of the contents and services published on the MCSDP itself. While the first choice is a natural consequence of the platform provider’s traditional role, the second implies an atypical evolution of MMTP positioning and market scope, and gives rise to the insurgence of a value network “structural equivalence” (Gulati et al., 2000) between MCSPs and MMTPs, thus determining competitive attrition among the two player typologies.

- **Resources & Competencies**
  As the “research based view” and the “dynamic capabilities approach” state, a firm’s collection of path-dependent core resources and competencies strongly influence its ways of seeking competitive advantage (Hamel, Prahalad, 1990; Teece et al., 1997). As a consequence, if the prevalence of technology oriented R&C makes a firm better disposed towards a mere technological partnership with its potential customers, the unbalance towards content oriented resources and capabilities enhance the MMTP tendency to propose itself as an “editorial partner” to MNOs, that is, a player capable of covering the activities of content creation, management and market making.

- **Vertical integration**
  The level of vertical integration refers to the MMTP coverage of activities in the Mobile Content Value Network. A positioning on the Platform Layer activities denotes a clear choice of self-relegation to a peripheral place in the network, covering a technology enabler role which does not go beyond the MCSDP provisioning and management processes, and stays out the downstream chain that allows direct contact with the end user. Contrariwise, selecting a positioning embracing an integrated technological and commercial management of both the platform and the contents published on it, puts the MMTP in a more central role in the system, closer to the “network focal” – the MNO – and to the primary source of revenues – the end customer – (Gulati et al., 2000; Peppard, Rylander, 2006). Of course, such strategic choice implies a more direct competition with MCSPs.

- **Customer ownership**
  Strongly related to the choices concerning vertical integration, customer ownership deals with the nature of the relationship established between the MMTP and the end customer. An intermediated customer ownership on the Technology Provider’s part implies a higher reliance on MNOs and MCSPs; the MCSDP vendor only receives indirect revenues streams from its business counterparts. Instead, a direct relationship with the end customer enhances the MMTP position in the Value Network, causing competitive attritions with MCSPs.

- **Revenue model**
  The revenue model parameter refers to the kind of revenue streams flowing from the MNO/MCSP to the MMTP, that can vary from mere selling of the platform, to a full revenue sharing agreement on the contents/services delivered through the MCSDP. The choices related to this element, are strictly linked to the platform provisioning parameter, and shall be considered extremely critical,
because of their many implications on the firm’s overall positioning and strategy. While system selling is based on a spot and fixed revenue for the MMTP, and presupposes a clear distinction between its business and the ones of its customers, the full revenue sharing solution rests on a division of potential revenues coming from contents/service selling to end customers. As such, the latter solution is strongly affected by the uptake and the consequent success of the service provided by MNOs and MCSPs; therefore, the MMTP revenues are spread on the whole service lifecycle, and are subject to a higher variance, for the technology provider is sharing not only opportunities, but also risks related to the service commercialization, finding itself in a “business sharing” condition.

The case studies showed that system selling and revenue sharing agreements only represent the extremes of the _continuum_ of solutions available: in between, players can go for hybrid alternatives, like the combination of a “start-up fee” – also known as “set-up fee” or “minimum granted” – covering MCSDP development and installation costs, and a “monthly rent” for the platform provisioning; a “monthly rent” integrated with a “consumption fee” after exceeding predetermined thresholds of usage; or else, a “start-up fee” plus a “revenue sharing” agreement.

- **Cost model**

The cost model refers to the nature of investment undergone for MCSDP development. If the investment are concentrated on the MMTP side, the risks associated to the project are not shared, but the player can benefit from a greater strategic independence after the solution is created. In the case of joint investment between the MMTP and the MNO/MCSP, the risks related to the project are spread on several actors; still, the MMTP enjoys less freedom, as its choices will have to be aligned with the strategic priorities of its partners.

<table>
<thead>
<tr>
<th>Business Model Parameter</th>
<th>Value Range (Trade-off)</th>
<th>Strategic Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Platform Characteristic</strong></td>
<td>Modular, Interoperable</td>
<td>Easier diffusion/Substitution. Co-opetition with other MMTP.</td>
</tr>
<tr>
<td></td>
<td>Vertical, Stand-alone</td>
<td>Increased control. Lock-in/lock-out effects.</td>
</tr>
<tr>
<td><strong>Platform Provisioning</strong></td>
<td>In-house installation</td>
<td>Increased MNO/MCSP technical independence. Separation between MMTP and MNO/MCSP business.</td>
</tr>
<tr>
<td></td>
<td>Hosting (ASP)</td>
<td>MNO/MCSP reliance on MMTP technological infrastructure. Exploitation of scale/scope economies on the MMTP side.</td>
</tr>
<tr>
<td><strong>Additional Services</strong></td>
<td>Platform Management</td>
<td>Focus on technology dimension. Full technical service approach.</td>
</tr>
<tr>
<td></td>
<td>Content Mktg &amp; Sales</td>
<td>Extension on downstream activities. Competition with MCSP.</td>
</tr>
<tr>
<td><strong>Resources &amp; Competencies</strong></td>
<td>Technology oriented</td>
<td>Disposition towards technology partnership.</td>
</tr>
<tr>
<td></td>
<td>Content oriented</td>
<td>Disposition towards editorial partnership.</td>
</tr>
</tbody>
</table>
Vertical Integration

- Platform Layer coverage
- Relegation to technology enabler role.

Customer Ownership

- Intermediated
- Direct
- Increased dependence on MNO/MCSP
- More central role in the VN. Competition with MCSP.

Revenue Model

- System selling
- Full Revenue sharing
- Clear separation between MMTP and MNO/MCSP business
- Business sharing (opportunities/risks) between MMTP and MNO/MCSP.

Cost Model

- Concentrated investment
- Joint investment
- Increased independence.
- Risk sharing.
- Increased dependence on MNO/MCSP.

Table 2. MMTP Business Model Parameters reference framework

In the next section, the noteworthy combinations of business models parameters, as emerged from the case studies, will be disclosed, and the related strategies will be described.

5 EMERGING BUSINESS MODELS AND CORE PARAMETERS

After identifying the strategic implication of single business model parameters, the further step of the study focuses on discovering and interpreting MMTPs’ emerging business models and strategies.

The in-depth analysis on the theoretical sample of 24 firms allowed to identify three main emerging business models currently developed and adopted by these players, corresponding to noteworthy specific combinations of parameters: such business models were then associated to three underlying “strategic patterns” that appear to be driving the players activity in the Value Network.

1. “Pure play” Business Model, determined by a “Stay on core” strategic pattern;
2. “Full asset” Business Model, determined by a “Grow, wait and see” strategic pattern;

![Figure 2. MMTPs emerging business models and strategic patterns](image-url)
The “Pure Play” Business Model is adopted by 14 firms out of 24, and is characterized by: a value proposition strongly focused on technology, in terms of platform provisioning – in-house installation is preferred to ASP or outsourcing, additional services – restricted to platform management, and resources & capabilities mainly technology oriented; a clear positioning on the Platform layer of the Value Network – distant from the end customer, bringing about a sharp distinction between the MMTP and MNO/MCSP businesses; and a financial configuration resting on fixed revenues and concentrated investments. The model is therefore defined “pure play” as the MMTPs employing it have pursued a consistent alignment between internal structure and external positioning, totally focused on the role of technology enablement.

The strategy determining this architecture is called “stay on core”, as all the informants of the firms comprised in the cluster declared that the business model design process was guided by the strategic choices of focusing on the traditional core business, oriented to the simple offer of technology. Other motivations leading to such conservative strategic positioning were the decision to restrain from representing a threat – real competition or even potential overlapping of activities – to their current customers, MNO/MCSP, and the unwillingness to internally develop \textit{ex novo} the structure and know-how necessary for creating and commercializing digital contents.

The adoption of “dirty” business models characterized by a non-transparent positioning towards the customers is explicitly criticized. In particular, the establishment of a full revenue sharing agreement is considered not advisable by the large majority of “pure play” firms – the informants belonging to 12 companies out of 14 labeled it as “way too risky” or “unfeasible”, for the following reasons: the revenue models structure grants extremely low margins to the technology provider – ranging from 1% to 5% of the total revenue; revenue sharing relies too strongly on the delivered services’ performance, and usually turns into a “loose-loose” game for the MMTP – if the service is unsuccessful, a full coverage of MCSDP development and installation costs is not assure; but even if the service proves itself appealing to the market, the MMTP is often forced to renegotiate the contract and reduce its share of margins, due to the higher bargaining power its customers possess.

The “Full Asset” Business Model is adopted by 8 firms. It differs from the “pure play” model in the tendency shown by these MMTPs to acquire and/or develop a wide portfolio of assets, resources and capabilities, not only related to the Platform Layer, but also to the Content & Service Layer. Nevertheless, for the moment these players are not leveraging on their “full asset” portfolio, as their actual coverage is still concentrated on technology activities, not being far from the positioning chosen by “pure play” MMTPs.

Analyzing the interviews, it is possible to argue that these firms are following a “grow, wait and see” strategy, as they recognize the value of creating a know-how on content creation and commercialization, and keep on investing on their pool of assets, but are still reluctant to abandon their traditional business. They would rather wait that the market takes a more defined shape, where they hold a consolidated position as a technology enabler; as soon as “time is right”, they may decide to exploit their high competitive potential, expanding their scope to the market making of contents and services.

The “Platform & Content Management” Business Model is only adopted by 2 firms: still, it deserves attention as its implications for the future development of the whole Value Network can be extremely significant. The MMTPs employing this model have extended their reach to the Content & Service Layer, embracing an integrated technical-commercial management of mobile digital contents. Their value proposition lists to hosting solutions of platform provisioning, to additional services related to content market making, and to content-oriented resources & capabilities; their vertical integration is high, covering activities which grant higher customer ownership; their financial configuration sees the possibility of establishing revenue sharing agreements, as well as joint investments.

Taking advantage of the evolved relationships cultivated with their partners – the MNO Vodafone in one case, and the Media Company Mediaset in the other, these players made innovative and explorative strategic choices, particularly aggressive in the downstream activities close to the end
customer. Their aim is to contribute in creating the commercial ecosystem that represents the main outlet for their technology solutions, and at the same time to place themselves in a more central position in the network, closer to the network focal and to the end user.

The drawback of this new role is related to the competitive dynamics that it could generate. MCSPs could see their business threatened, and start perceiving MMTPs as competitors: to retaliate, they could try to strengthen the ties linking them to Content Providers and Operators, thus isolating the platform providers; the biggest MCSPs could also undertake a process of upstream integration, acquiring the skills to internally develop their MCSDP. However, as the phenomenon of overlapping between MMTPs and MCSPs is extremely recent and not yet generalized, its competitive evolutions are still hardly foreseeable, and shall be subject to future research.

6 CONCLUSIONS

The research allowed to identify the core business model design parameters for Mobile Middleware Technology Providers; moreover, it shed light on how these building blocks are currently combined by MMTPs to give rise to complete business models architectures, and what strategies seem to drive such design choices.

Concerning the first major research objective – the business model parameters identification, the findings shows that some key business model parameters identified by the existing literature can be applied to MMTPs’ business model design activity, while others were missing or not made explicit. With reference to the second research objective – the individuation of the analyzed players’ design choices, three noteworthy business model currently adopted by MMTPs – “Pure Play”, “Full Asset” and “Platform & Content Management” business models – were identified, associated respectively to three underlying strategic patterns – “stay on core”, “grow, wait and see” and “aggressive downstream”. Thanks to the rigor of the methodology employed, and to the width and significance of the theoretical sample analyzed, these research can be replicated, and its findings can be generalized.

The paper’s value for researchers can be brought back to its contribution to Value Network, Business model design and Strategy definition theories. Existing literature on Value Network – with specific reference to the Mobile Content Network – was extended, through the provisioning of a unified definition for the player typology under scrutiny and its role in terms of activities covered. Business model design literature was applied to the study of a new player typology, and original design parameters, as well as their combinations to create a first “taxonomy” of MMTPs business models, have emerged. Moreover, the relation between strategy creation and business model design was made explicit, through the identification of business model design choices’ strategic implications.

The value for practitioners lies in the creation and provisioning of a “reference framework” capable of supporting the decision making process of business model design for a MMTP, as it presents strong ties between business model parameters and strategic implications.

The research represent a significant step towards the development of business model design theory with reference to Mobile Middleware Technology Providers. However, it does not analyze the potential different performances coming from alternative parameters selection. Future works will have to concentrate on the identification of newly emerged strategic patterns, resulting in alternative combinations of business model parameters, and to develop comparative or “benchmarking” analysis among them, in order to explain any differential in firms performances, pinpointing which single parameter or parameters mix may be seen as the origin of such deltas.

References


INFORMATION TECHNOLOGY AND THE FIRST-LINE MANAGER’S DILEMMA: LESSONS FROM AN ETHNOGRAPHIC STUDY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0703.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Ethnography, Management practices, Coordination, Practice</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
INFORMATION TECHNOLOGY AND THE FIRST-LINE MANAGER’S DILEMMA: LESSONS FROM AN ETHNOGRAPHIC STUDY

Joao Vieira da Cunha
Faculty of Economics, Universidade Nova de Lisboa
Email: jvc@fe.unl.pt

Andrea Carugati
Århus School of Business, Århus University
Email: andreac@asb.dk

Abstract

This research builds on the literature on information technology and organizations to suggest an alternative to the current understanding of the production of computer-generated representations of work. This literature sees computer-generated representations of work as automatic outcomes of information technology that managers use to scrutinize employees. We present an ethnography of a desk-based sales unit which suggests that first-line managers can address the tension between the need to enforce prescribed goals and procedures and the need to adapt to and protect employees’ improvisation by forfeiting surveillance and instead use information technology to build a façade of compliance with prescribed goals and procedures. Our results to shed light on the hidden labours behind representations of compliance and place agency in the centre stage of the process of producing computer-generated formal representations of work.

Keywords: Information system use, control systems, managerial work, improvisation.
1 INFORMATION TECHNOLOGY AND THE MANAGER’S DILEMMA

How do first line managers use information technology to address the tension between the pressure to enforce prescribed procedures and goals and the pressure to adapt to local contingencies dictated by customers and employees?

One the one hand first-line managers have a stake in enforcing compliance with prescribed procedures and with prescribed goals. First-line managers’ rewards and their careers depend on their ability to enforce compliance with the organization’s goals and procedures (Drucker, 1993; Ivancevich, 1974; Kamoche, 2000). Compliance with prescribed goals and procedures makes first-line managers more visible to upper-level managers, opening up access to informal networks of power and influence (Jackall, 1989). We refer to the pressure and incentives to enforce compliance as the ‘bureaucratic pull’. On the other hand, first-line managers are the only level of management that is close enough to employees to be able to face directly the ‘improvisational pull’: our label for the inadequacy of prescribed procedures to which employees need to adapt to (Lipsky, 1980; Prottas, 1976; Orr, 1990) to face everyday’s contingencies. This unavoidable inadequacy requires employees to improvise their own procedures for their organization’s sake. A case to the point is the phenomenon of “zeal strike” where employees fully apply prescribed procedures to block work completely.

The management literature has documented three tactics to deal with the first-line manager’s dilemma – the tension between the bureaucratic pull and the improvisational pull: adapting, imposing and compromising (Adler & Borys, 1996; Dutton, Ashford, O’Neill, & Lawrence, 2001). When adapting (e.g. Blau, 1955), first line managers allow employees to adapt prescribed procedures to the contingencies of their everyday work. When imposing (see Gouldner, 1954, chapter 11) first-line managers enforce compliance with prescribed procedures and goals, irrespectively of these contingencies. When compromising, first-line managers attempt to strike a balance by enforcing some prescribed procedures while adapting others to local contingencies (Bartunek, Krim, Necoechea, & Humphries, 1999).

Research treats these three alternatives as equally viable independently of the material aspects of the experience of the first-line manager’s dilemma. However, as Yates (1989) has shown, the use of information technology in organizations, shapes the material experience of the first-line manager’s dilemma. Upper level managers’ assessment of employees’ compliance with prescribed goals and procedures is, in modern organizations, mediated by computer-based formal representations of work. Strictly speaking, in this condition the bureaucratic pull is a pull to produce a representation of compliance with prescribed roles, rules, procedures and goals with the use of information technology. This gives first-line managers another alternative to address the first-line manager’s dilemma: a tactic that focuses on using information technology to create a representation of compliance without enforcing the compliance with goals and procedures that underlies that representation.

This paper explains how this façade-creation tactic can be enacted by drawing on the computer based practices that first-line managers use to address the tension between a strong bureaucratic pull and a strong improvisational pull. To achieve this result we conducted an ethnographic study of a sales unit during the 3 months preceding and the 12 months succeeding the implementation of a CRM system (Siebel & House). We draw on these data to explore first-line manager practices in dealing with the tension and unpack their consequences for theory of technology in the workplace and managerial practices.
2 METHOD

2.1 Research setting

DeskSales, the unit where we did this study, is a business-to-business desk sales unit at E-Tel, a large European telecommunications firm. Its employees (desk salespeople) had two sets of tasks: to sell and to do saleswork. Selling consisted of persuading customers to buy the products and services offered by E-Tel. Saleswork included all the ancillary activities to selling, such as creating sales plans and contacting customers regularly to detect sales opportunities. DeskSales deals with large corporate customers (retail, finance, and tech sectors) working in cooperation with the field sales teams to achieve prescribed sales targets. In practice, each customer has one account team headed by a field sales manager who supervises two to five field salespeople. To each account is allocated a Desk Salesperson who can be tied to max four accounts. Desk Salespeople work in teams of seven to thirteen under a Desk Sales Manager. They report to one of three Senior Desk Sales Managers in the unit who in turn report to the unit’s General Manager. The structure is shown in figure 1.

Figure 1  The structure of DeskSales (example shows Retail and Tech accounts)

E-Tel implemented Siebel to assist Desk salespeople do their work and at the same time, used to assess their performance. Siebel is a sales automation system that is designed to help salespeople throughout the sales process. It allows salespeople to store and access information about their customer contacts; keep track of their present and past saleswork and store and access information about their open sales. Siebel is also a management tool and it was used as such by E-Tel’s managers to supervise desk salespeople’s sales and saleswork. Siebel data were also the basis to calculate the bonuses of salespeople and their managers. Finally the data collected in Siebel were used to justify promotions and account assignments. Siebel was thus at the core of the experience of managers and employees in their everyday work in E-Tel sales units.

2.2 Research procedures

This research focused on documenting Desk salespeople’s and desk sales managers’ work practices. For this purpose one of the authors conducted an ethnography during the first 15 months of DeskSales’ existence. The core dataset for this paper is a set of observations of the managers practices during 51 phone-based desk sales team meetings which we used to infer the practices that desk sales managers enacted in these meetings to address the challenges of prescribed vs. improvised work. These data were complemented with observations (including desk salespeople’s initial and ongoing training), 55 interviews with desk salespeople, sales managers, the unit’s general manager, field salespeople, and the unit’s training staff.
Analysis began by coding full transcriptions of phone-based team meetings, fieldnotes and interviews to articulate the challenges that desk salespeople and their managers faced when attempting to shape the formal representations of their work and by documenting the practices they enacted to do so (Barley, 1990). These challenges and tactics were arranged into narratives (Carlile, 2002) to articulate the processes by which desk sales managers and desk salespeople were able to participate in the production of the formal representation of their work. In what follows, we document the bureaucratic pull and the improvisational pull over desk sales managers and explain the practices that they enacted to address the tension created by these two forces.

3 THE DESK SALES MANAGER DILEMMA

Desk sales managers faced an especially straining experience of the first-line manager’s dilemma. Mariah, who was DeskSales’ General Manager during the period of this study, enforced demanding sales targets creating a strong bureaucratic pull. Desk salespeople’s inability to sell and their unwillingness to report sales and saleswork in Siebel fed a sizeable improvisational pull over desk sales managers.

3.1 Bureaucratic pull

Mariah saw her role as that of promoting DeskSales to the rest of the company. In her own words, she was “the one who sells their [desk salespeople’s] effort. I am the salesperson for the desk role.” Mariah used the unit’s sales target to showcase DeskSales’ value to the rest of the company. In the fiscal year of 2002/2003 Mariah committed to a stretch target of 350 million dollars of sales in 8 months, far exceeding the unit’s initial sales target of 330 million dollars in 12 and thus raising the bureaucratic pull over desk sales managers who were responsible for this target. Mariah also imposed three saleswork targets: desk salespeople had to report making seven calls to customers each day; all of the sales reported in Siebel had to have a complete salesplan; and no desk salesperson could have errors in their sales records in Siebel. Mariah supervised, rewarded and enforced Desk salespeople’s progress towards reaching their sales and saleswork targets. Desk salespeople called her the “Demonic Mistress of Siebel” because of her focus on enforcing reported sales and reported saleswork in Siebel. Mariah claimed that Siebel made her job “much easier,” not only because she could use it to supervise Desk salespeople’s but also because it made DeskSales performance “transparent” to upper-level management: “if you don’t have a shared system, people will just think that you’re making up the numbers – with a shared system it’s totally transparent.”

Mariah did not use Siebel to report automatically the sales and saleswork that desk salespeople logged in the system. Instead she had her desk sales managers prepare these reports. Reporting became an opportunity for compliance. Desk salespeople rarely reported enough calls to customers to meet their saleswork targets for each week. When Desk Sales Managers were creating their team’s weekly report, they had the opportunity to ask their Desk salespeople to report enough calls to customers to hit those saleswork targets. Additionally, this practice allowed desk sales managers to have their Desk salespeople change sales records in Siebel so that, for example, if a Desk Saleswoman had hit her sales target for any given week, she could report sales in excess of that target in the next week so as to also meet her targets then. This effort to use Siebel’s weekly reports to create a representation of compliance for every Desk Salesperson in their team occupied the majority of desk sales managers’ time at work. When asked to reflect on the negative elements of her job as a Desk Sales Manager, Linda complained:

One of the major problems we have had at DeskSales is that we need to keep proving ourselves [...W]e have to keep producing statistics and spreadsheets and things like that and for me personally as a Desk Sales Manager, having to produce these things to feed them up so that [our General Manager] can justify what we’re doing [...] so that [top managemers] are happy. So it
takes a lot of time and it’s unproductive work as well because it doesn’t help you make more sales, it’s just looking back, so that justification process is quite frustrating.

This was but one of many comments of the kind by desk sales managers. Creating reports from Siebel data was thus not only a time consuming activity, but one which desk sales managers interpreted as being of little if any value.

Mariah’s use of reporting shaped desk sales managers’ experience of the first-line manager’s Dilemma in three ways. First it added a set of reporting tasks to desk sales managers’ everyday work, reducing the amount of time they had available to manage and supervise their Desk salespeople. Second it added to the bureaucratic pull: desk sales managers not only had to make sure that their Desk salespeople engaged in the sales and saleswork need to meet their targets, but also that their Desk salespeople reported those sales and that saleswork on Siebel. Finally, it sharpened the challenge created by the bureaucratic pull: strictly speaking, desk sales managers had to enforce the representation of sales and saleswork on Siebel, rather than the actual execution of sales and saleswork.

To summarize, the bureaucratic pull under Mariah was driven by a “stretch” sales target that had Desk salespeople achieve more revenue than E-Tel upper-level managers required, and in less time. This push was strengthened by the increased amount of reporting work that desk sales managers and Desk salespeople had to do but it was partially alleviated because Mariah’s targets were focused on the act of reporting sales and saleswork, not on that of carrying these out.

3.2 Improvisational pull

Desk sales managers had to rely on their Desk salespeople to hit the sales and saleswork targets set by Mariah. Desk salespeople, however, had to hit their sales and saleswork targets under challenging conditions. They did not have sales skills and they were not trained to sell. Desk salespeople were hired from temporary staffing agencies or from E-Tel residential call centres, possessing little work experience and no experience in business sales.

The training that Desk salespeople received upon joining the unit was meant to address Desk salespeople’s inexperience with sales. Mariah’s efforts to enforce sales and saleswork targets, however, deteriorated the quality of incoming Desk salespeople’s training, thus strengthening the improvisational pull over desk sales managers. Mariah enforced saleswork targets based on the number of people allocated to the team, independently of whether they were on training, on sickness leave or on holidays. This led desk sales managers to forbid incoming Desk salespeople to attend most the training sessions. Desk sales managers had desk salespeople spend as much time as possible at their desk, reporting sales and saleswork in Siebel. This reduced desk salespeople’s confidence in their ability to sell because they felt they could not address even the simplest of requests. The week after the training ended, Ted, a newly hired Desk Salesperson, complained to the other Desk salespeople in his cohort that he got an email from his field salesperson asking Ted to price a product. Ted confessed, “I couldn’t understand a word of what was written in it”. This lack of sales skills and sales training was amplified by difficult relationships with field salespeople. Donald, a Desk Salesman in one of the two Finance Desk Sales Teams, voiced a common complaint:

[My job] is quite frustrating, at the moment. I find it quite difficult to get into the account. Quite difficult to sell. Because [...] it’s quite a large account team, they’re involved in sort of building the networks and the relationship with the customer; [...] it’s quite hard for me to get into the account and find a specific role in it. So I find myself just getting on, getting on with [service issues such as] backup for the systems engineers and that sort of stuff so it’s quite frustrating.

Mariah’s tenure coincided with an increased enforcement of Siebel use throughout E-Tel. Desk salespeople took on the increased Siebel work that field salespeople had to complete in exchange for being
able to claim some credit for part of field salespeople’s sales. Desk salespeople also attempted to take over field salespeople service work for their customers. The goal was to increase their interaction with customers to get some of the orders that customers would place with E-Tel’s service units. To this end, Desk salespeople introduced themselves to customers as their “single point of contact at E-Tel.” Desk salespeople’s goal was to have a service relationship with their customers unmediated by field salespeople. This gave Desk salespeople visibility of their customers’ orders – orders that they could report in Siebel as the outcome of their own effort. These two tactics helped Desk salespeople contribute towards their team’s sales targets, partially alleviating the improvisational pull over desk sales managers.

3.3 Addressing the first-line manager dilemma through an electronic façade

Mariah’s demands on desk sales managers’ reporting work left them with very little time to spend actively managing their desk salespeople. Nonetheless, desk sales managers needed to have a high level of visibility and control over desk salespeople’s sales and saleswork because their bonus was only paid if every single desk salesperson on their team met their targets. To address this challenge, desk sales managers focused their interaction with their desk salespeople on a phone meeting with their team (labelled the “team audio-call” at DeskSales), which took place about once every week or once every two weeks – depending on how well the team was doing against its sales and saleswork targets. Although these meetings were conducted over the phone, they were also face-to-face in the sense that the team’s Desk Sales Manager and its Desk salespeople were sitting at their Desk in DeskSales’ open space. Desk sales managers chose to conduct these meetings over the phone to keep them private and to avoid disturbing the other Desk Sales Teams. These calls could range from a five-minute announcement to an hour-long scrutiny of each Desk Salesperson sales targets, but on average they lasted for about 30 minutes. An analysis of all the audio-calls observed during Mariah’s tenure as DeskSales General Manager shows that desk sales managers used these calls to enact four practices to ensure that their Desk salespeople met their targets for reported sales and reported saleswork in Siebel.

The four practices were: monitor Desk salespeople’s representation work (reporting sales and saleswork in Siebel); enforce representation work on Desk salespeople; reward Desk salespeople for their representation work; and advise Desk salespeople in how to hit their targets through representation work. These practices are explained below.

Monitor

Desk sales managers could observe very little of desk salespeople’s everyday work. The amount of reporting work that desk sales managers had to do prevented them from observing their desk salespeople’s work directly. Desk sales managers could only see the sales and saleswork that desk salespeople reported in Siebel. Desk sales managers could not see two aspects of Desk salespeople’s work that were crucial to assess their desk sales teams’ ability to reach its sales targets: their progress towards closing their open sales in Siebel and the sales that they could report as their own in Siebel but had yet to do so. The analysis of the audio-calls observed during Mariah’s tenure reveals that desk sales managers used these calls to monitor these two aspects of their Desk salespeople’s work. Anthony’s team audio-call of February, 24th shows how desk sales managers used these calls to monitor what their Desk salespeople were doing to close the open sales they reported in Siebel. Here we present a fieldnote of a typical sort audio-call:

Anthony’s call started by calling Gregory: "what's the scores on the doors for you today?" Gregory said that his open sales were still far from closing. Anthony then called Kevin, who replied: "I'm just catching up on stuff." Anthony retorted: "that's very specific Kevin" and laughed. Anthony then called Alexander who said that he had spoken with one of his field salespeople and that he had gotten permission to report a sale worth 100 000 dollars under his
own name in Siebel. He said that his 15 million-dollar open sale was now only due to close in March ("sorry Anthony"). Then it was Mark's turn, who said that he was “just following up stuff.” The audio-call ended here.

As Anthony’s audio-call highlights, desk sales managers used these calls to see what they could not see in Siebel – the full process of their Desk salespeople’s representation work. Siebel only showed part of this process: the value of the each Desk salespeople’s open sales. It did not show the progress sales that Desk salespeople had made towards reporting these sales as closed in Siebel. Without their audio-calls, desk sales managers were not able to assess which of their Desk salespeople’s open sales were going to be reported as closed and which would be reported as lost or cancelled.

Desk sales managers also used their audio-calls to monitor Desk salespeople’s use of Siebel to report sales. Once Desk salespeople reached their targets for the quarter, they stopped logging sales in Siebel. They waited for the next quarter to log the sales they were allowed to report as their own in excess of what they needed to meet the current quarter’s target. This allowed Desk salespeople to hit their targets every quarter. If their Desk Sales Team as a whole was lagging behind its sales targets, however, it was in desk sales managers’ interest to have these hidden or latent sales logged in Siebel. This would not help letting Desk salespeople hit their targets but it would allow the team as a whole to do so. In the second quarter of 2003, Roy’s team was lagging behind their sales target, although Karl, one of his Desk salespeople, had already reached his own sales target. In one of his team’s audio-calls towards the end of the quarter, Roy attempted to have his Desk salespeople report all the sales they could in Siebel. During the audio-call, he said, "knowing that we have a gap of 1 million, we need to have a pipeline of 3 million to make sure that we hit that.” With that in mind, he asked if they had "any sales squirreled away under the blankets.” “Do you have anything that you decided, "let's chuck it away for next month or let's keep this one off the radar,” “do we have any sales records below 30% [probability of closing] that we could progress over the value chain in the next month?” He asked everyone individually, by calling their name. Adolph and Barry said that they didn't. Stewart shook his head with a smile, "that was a very shady no" Roy said "do you have one?" Stewart said that he did. Roy said "let's not move everything to 90% percent and expose ourselves, [but] if we're 100% sure put it up to 40%, [Mariah] is going to be looking at where we add value and if she only sees Stewart’s [sales numbers] she may wonder what the hell we're doing here."

Desk sales managers focused their supervisory role on a very specific part of Desk salespeople’s work: reporting sales and saleswork in Siebel. More broadly, what is striking about desk sales managers’ use of their audio-calls to monitor their Desk salespeople was their choice of target. Desk sales managers did not observe, nor did they seek to observe the work that allowed Desk salespeople to report sales and saleswork in Siebel. Desk sales managers only sought to observe the reporting work.

Enforce.

Siebel reports and team audio-calls often revealed a sombre picture of Desk salespeople’s ability to reach their monthly reported sales and reported saleswork targets. The gap between the amount of sales and saleswork that Desk salespeople had actually reported and the amount of sales and saleswork they needed to report for their team to hit its targets was often too wide. Desk sales managers addressed this gap by enforcing representation work – ordering their Desk salespeople to report sales and saleswork in Siebel.

Desk sales managers used their audio-calls to enforce representation work in two ways. One was by exposing desk salespeople’s lack of achievement hoping that they would be shamed into compliance. Desk sales managers announced each Desk Salesperson’s sales and saleswork numbers in audio-calls. In his audio-call of January, 15th, Ferdinand announced the closed sales reported in Siebel in an effort to get his Desk salespeople to report more closed sales. Ferdinand started by announcing the total revenue
figures for each Desk Salesperson. Laura has 5 million, "well done." Oscar "hot on your heels," Ferdinand added to Laura. Kyle got 67,000, "well done." Nancy, has "lots of quotes" but she needs to "translate them into [closed sales] in Siebel.” Nancy replied that quotes would take 2 to 3 months to be in Siebel because she needs to track them down with Service. Christine had 1.5 million: "Christine is really storming up now, well done Christine." Ferdinand said that the team as a whole had brought in 10 million of revenue. "When people from management come down you should know this so that you can tell management so that we'll look good,” Ferdinand added. Desk sales managers made Desk salespeople’s numbers visible to each other was by using their team’s whiteboard. Whiteboards were designed in table format. Desk sales managers used these tables to record their Desk salespeople’s achievement towards their targets. Linda, a Desk Sales Manager in one of the Finance teams explained her rationale for having a whiteboard by stating that, “those who feel embarrassed because they’re not at the top […] quite often come up [to talk to me] and say ‘I’m not doing well here’ and then we’ll have a discussion about that.”

The other way in which desk sales managers enforced representation work was by directly instructing Desk salespeople to report sales and saleswork in Siebel. In his audio-call of May 2nd, Andrew instructed his Desk salespeople to report sales following Siebel’s sales cycle. Siebel had a model of the stages that a sale had to go through. After going through his Desk salespeople’s open sales, Andrew then asked if there were any large sales to be closed that week. Stella said that she would probably have something next week. Michael said that he had his “80K one,” but he was still unsure when he was going to put it on Siebel. Andrew replied, "don't keep too much off Siebel, our team doesn't loose any open sales and that's not good.” Larry added, "if you look on Siebel, we have a very high conversion rate, that's not a true picture.” Andrew said that "we don't want open sales to be logged in Siebel at 100% probability, that looks very reactive.” Larry retorted, "yes, but we get a lot of orders" and Michael replied, "well, we can hold them and log them at a lower probability and close them a few days later.” Andrew said that "we need to follow all steps [stages in the sales cycle] because it puts them [open sales] on the radar early.” Andrew instructed his Desk salespeople to follow Siebel’s sales cycle to create the representation of selling when in fact they only reported sales that field salespeople had already closed.

Thus, when feelings of shame did not ensure that Desk salespeople met their targets for sales and saleswork, desk sales managers ordered them to report more of both. The analysis of audio-calls observed during Mariah’s tenure shows that desk sales managers enforced reporting work, not sales or saleswork. Desk sales managers did not make their Desk salespeople’s sales or saleswork visible to others, only the representation of those sales and that saleswork. Similarly, desk sales managers did not order their Desk salespeople to sell more or to make more calls to customers. Instead, they ordered them to log more sales or report more calls to customers in Siebel. Desk sales managers enforced Desk salespeople’s representation work, not their work itself.

Reward.

Desk sales managers also used their team’s audio-calls to reward their Desk salespeople. Desk sales managers offered good quality and high-priced alcoholic beverages to their Desk salespeople as an incentive to achieve specific, short-term reporting goals. In his audio-call of July 7th, Josh, A Desk Sales Manager in one of the Technology and Media teams, announced an incentive to reward the Desk Salesperson that reported the largest number of open sales in Siebel during the upcoming month. Josh announced pompously that he will give a bottle of champagne to the person that reported the most open sales on Siebel, "I want proper [open sales records] going on there, not a load of bollocks. I want to see them actually converted into [closed sales]." Scott, one of the Desk salespeople on the team, asked "what make of champagne?" Jason replied “Bollinger.” Josh added that "this incentive is just a launch pad. I want this to go along, it's part and parcel of the role, really."
Desk sales managers also used symbolic rewards – plastic Oscar figurines, paper ribbons and tin cups – to reward reporting sales and saleswork. These rewards were not aimed at fostering an increase in sales and saleswork. Instead, these rewards aimed at recognizing Desk salespeople’s achievement towards their reported sales and reported saleswork targets. Andrew had his “Golden Salesman” award for the Desk Salesperson whose sales totalled the highest revenue each week. At the start of his audio-call on August 15th, Ferdinand, a Desk Sales Manager for Retail teams announced that it was time to award "the trophy,” "Oscar is not the best performer anymore, although he might be,” Ferdinand proclaimed. He said that "this week the trophy is going to someone that has been exhibiting the right behaviors, it's Casey!” Casey raised his arms in sign of victory and the team applauded. Ferdinand handed him the trophy while saying, "we got some feedback on your sales plans that you actually took them seriously, while a lot of the others were taking the piss. Also you had a great turnaround in your [reported sales] in the past two weeks." Mild applause followed. Ferdinand said that it was time to take the picture and Casey moved next to the window and Ferdinand used a Polaroid instant-camera to take a picture of him with the trophy. Casey then put the picture and the tin trophy on top of his desk.

Desk sales managers also used formal E-Tel incentives to reward their own team. In his audio-call of July 24th, Daniel, a Desk Sales Manager in one of the Technology and Media team used a bottle of champagne he earned at a sector event the previous day to reward Andy, one of his team’s top Desk salespeople. He said that "Christian [the Technology and Media Top Manager] loves us,” and that "Christian has completely bought into this model [of using Desk salespeople]" because he was “very pleased” with the work they were doing. Daniel pulled out a bottle of champagne and said that "I was dragged on stage and Christian gave me this bottle of champagne and there was a lot of applause.” He added, "you're the ones doing the hard work, so I want to give this bottle of champagne to you.” He said that he realized that he couldn't share the bottle of champagne with everyone, so he was going to give it to the person with whom senior management was most impressed, Kevin. Everyone applauded and Daniel handed the bottle of champagne to Kevin.

Desk sales managers’ rewards thus followed their monitoring and enforcing practices. They focused on Desk salespeople’s representation work, no their actual sales or saleswork. Material and symbolic rewards recognized the use of Siebel to report sales and saleswork, not sales and saleswork themselves.

Advise.

Desk sales managers could enforce and reward their Desk salespeople’s representation work but still see their team lagging behind its reported sales and reported saleswork targets. Desk salespeople could have the motivation to report sales and saleswork but not the skills to sell, the permission to log as their own the sales made by their field salespeople, or the opportunity to call their customers. To address this, desk sales managers advised their desk salespeople to on how to reach their sales and saleswork targets. The analysis of audio-calls shows that desk sales managers advised their Desk salespeople in three ways. First, desk sales managers helped Desk salespeople reach their revenue targets by telling them how to use E-Tel’s information systems to find revenue. In her audio-call on March, 4th, Nina, a Desk Sales Manager in one of the Retail teams, explained how to use one of E-Tel’s information systems to find revenue that Desk salespeople could claim as their own. She explained to her Desk salespeople that "in E-Tel you'll soon find that things aren't like they should be: sometimes [some of the product revenue from your accounts] is not visible on ARTS [E-Tel’s automatic revenue tracking system].” Nina added that "If I spend the whole day [selling] to [one of my accounts] and they're not in the system, I would be wasting all my time because I wouldn't see the revenue.” She continued: "sometimes you get products allocated to your account on ARTS that the [customer hasn’t bought] and they [bring revenue]: those ones you want to keep, so keep quiet.” Nina explained: “basically, the more you look, the more you'll find and it's a way to get revenue...
without having to sell anything. Of course, you will eventually have to sell something, but it's a good way to hit numbers."

Second, desk sales managers also helped their Desk salespeople meet their targets for reported sales by giving them specific advice on how to establish relationships with their field sales teams. In her audio-call of February, 19th, Nina used a question made by Dennis, a Desk Salesperson on her team, to explain to her Desk salespeople how to get field sales people to allow them to log sales in Siebel. Towards the end of the team’s audio-call Dennis raised his hand and Nina gave him leave to speak. Dennis said that there were a lot of large open sales for his accounts in Siebel and asked Nina how he could get some of the credit for those sales. Nina told the story of how Robert, another Desk Salesman on her team, got involved in an 8.6 million dollar opportunity that was about putting E-Tel Broadband access CDs in a retailer’s stores. Nina explained how Robert's involvement was limited to calling the stores to make sure that they got the disks. Nina said that "even making a quote [giving the price of a product] is involvement.” She concluded: "what I would do is ask the account team if there's something I can do to help.”

Thirdly, desk sales managers advised their Desk salespeople on how to log sales and saleswork in a way that helped Desk salespeople reach their targets consistently. Andrew told his Desk salespeople to refrain from reporting sales in Siebel after they had reached their quarterly targets. He advised his Desk salespeople to report those sales in the following quarter to make their targets easier to reach. In his January 15th audio-call, Ferdinand advised his Desk salespeople on how to go about reporting customer calls in Siebel. Ferdinand told them to report "any contact with customer” in Siebel. Laura, one of the Desk salespeople on his team, asked if the "emails that you get" count. Ferdinand said that they did, but "try to keep it relevant,” "I don't want Siebel clogged” with irrelevant information. "You would want to add on a call from a customer clarifying what that quote was.” He was quick to add: "Don't make your numbers plummet because of that." Don concluded "it's the [target] thing" and Ferdinand nodded in agreement. Don said that the point is to "hit 10 and then worry about relevance after [they hit 10].” Ferdinand said that to reach targets the boundaries can be "stretched" but then they should "retract.”

Desk sales managers’ actual advice sharpens the case that they were focusing on reporting work. Desk sales managers might need to monitor, enforce and reward representation work because Desk salespeople were engaging in sales and saleswork without reporting it in Siebel. Desk Sales Manager’s advice, however, focuses on how to scavenge unclaimed sales and how to persuade field salespeople to allow Desk salespeople to take credit for field salespeople’s sales. The tactics that desk sales managers taught to their Desk salespeople in their team’s audio-calls helped Desk salespeople hit their reporting targets without actually engaging in any of the work they reported.

4 DISCUSSION

The way desk sales managers addressed their experience of the first-line manager’s dilemma shows how information technology can be used to address these opposing forces by creating a façade of compliance with the bureaucratic pull, reducing the improvisational pull to employees’ ability and willingness to participate in the production and reproduction of that façade. This suggests an alternative view of the production of computer-generated formal representations of work that complements the dominant view of this process (cf. Ball, 2003; Findlay & McKinlay, 2003; Sewell & Wilkinson, 1992; Townsend, 2005). This dominant view hinges on the assumption that work and its formal representation of information systems are tightly coupled. This assumption has two corollaries. One is that managers use computer-generated formal representations of work to supervise work down the organization. The other is that employees have agency over these representations because these are generated by information technology as employees use it in the course of their everyday work. This view places deviance squarely with workers and establishes access to information as the core element of the struggle for power in organizations.
However, studies of the production of computer-generated formal representations of work that follow a structurational approach to IT suggest that the tight coupling between work and its formal representation is dependent on the everyday practices of managers and employees. This tight coupling is thus not a feature of technology but a consequence of people’s actions (cf. Orlikowski, 1991).

Our analysis suggests that Desk salespeople’s experience explains how a different pattern of production of computer-generated formal representations of work can emerge and be sustained. At DeskSales, desk sales managers and Desk salespeople’s use of information technology kept Desk salespeople’s work and its representation in Siebel were, at best, loosely coupled. Desk salespeople could produce a representation of compliance with sales and saleswork targets without any sales and without engaging in any saleswork (and yet staying legit). Desk sales managers used those Siebel-generated formal representations of work to show compliance upward, forfeiting scrutiny except when Desk salespeople were not representing compliance. When enacting such a set of practices, desk sales managers produced condensed, closed summaries of Desk salespeople’s work. The goal was to make scrutiny more difficult and less likely.

Under such conditions, desk sales managers could focus exclusively on enforcing representation work—the work of producing a representation of compliance with prescribed roles, rules, procedures and goals, without actually complying with any of these. Desk sales managers used their own improvised information system to enforce representation work by supervising down based on informal representations of work. This pattern of production of computer-generated formal representations of work widened the gap between Desk salespeople’s work and its representation.

This view of the production of computer-generated formal representations of work that we inferred from Desk salespeople’s experience at work is not a replacement, but a complement to the dominant view on this phenomenon. Desk salespeople’s experience and that of the research settings featured in the current literature on IT and organizations are two patterns that are produced by the same process, a process that hinges on the tightness of coupling between work and its representation in computer-generated formal representations of work, as determined by managers’ and employees’ practices.

5 CONCLUSION

The first-line manager’s dilemma is a common theme in both empirical and theoretical accounts of bureaucracy. These accounts show that first-line managers can address this dilemma by enacting tactics that can range from an imposition pattern—thus yielding to the bureaucratic pull—to an adaptation pattern, yielding to the improvisational pull. The literature on IT and organizations has argued that in context of computer-mediated supervision only an imposition pattern is sustainable. Desk sales managers’ experience shows that in contexts of computer-mediated supervision there is another alternative—creating an electronic façade of compliance. The way E-Tel’s upper-level managers used Siebel to supervise Desk salespeople’s work suggests that the assumption of tight coupling between employees’ action and its representation in formal information systems is not empirically universal. The disconnect between Desk salespeople’s work and its representation in Siebel allowed desk sales managers to interpret the bureaucratic pull as a push to create a representation of compliance. Desk sales managers limited their imposition of prescribed roles, rules, procedures and goal on their Desk salespeople to the creation of a representation of compliance with DeskSales prescribed goals. Desk sales managers interpreted the first-line manager’s dilemma at DeskSales as a problem of managing their employees’ representation work, instead of their sales and saleswork. To this end, desk sales managers created a parallel set of prescribed rules, roles and processes to manage their Desk salespeople’s representation work. To use desk sales managers’ favorite metaphor for Siebel, IT is indeed a window into everyday work, but instead of being a passive representation such as that allowed by the windows in Foucault’s panopticon, it is an active representation, such as that allowed by a “shop window” … dressed up to impress observers.
REFERENCES


THE RELATIONSHIP AMONG DEVELOPMENT SKILLS, DESIGN QUALITY, AND CENTRALITY IN OPEN SOURCE PROJECTS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0516.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Social Network Analysis, Quality, IS metrics, Skills</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
THE RELATIONSHIP AMONG DEVELOPMENT SKILLS, DESIGN QUALITY, AND CENTRALITY IN OPEN SOURCE PROJECTS

Barbagallo, Donato, Politecnico di Milano, Dipartimento di Elettronica e Informazione, Via Ponzio 34/5, 20133 Milano, Italy, barbagallo@elet.polimi.it
Franca Lanci, Chiara, Politecnico di Milano, Dipartimento di Elettronica e Informazione, Via Ponzio 34/5, 20133 Milano, Italy, francala@elet.polimi.it

Abstract

In a previous paper, we have found empirical evidence supporting a positive relationship between network centrality and success. However, we have also found that more successful projects have a lower technical quality. A first, straightforward argument explaining previous findings is that more central contributors are also highly skilled developers who are well known for their ability to manage the complexity of code with a lower attention to the software structure. The consolidated metrics of software quality used by the authors in their previous research represent measures of code structure. This paper provides empirical evidence supporting the idea that the negative impact of success on quality is caused by the careless behaviour of skilled developers, who are also hubs within the social network. Research hypotheses are tested on a sample of 56 OS applications from the SourceForge.net repository, with a total of 378 developers. The sample includes some of the most successful and large OS projects, as well as a cross-section of less famous active projects evenly distributed among SourceForge.net’s project categories.

Keywords: social networks, software quality, software design skills.
1 INTRODUCTION

Social networks represent social systems characterized by very large numbers of individuals and relationships among individuals. By empirically analyzing the evolution of OS networks modeled as sets of cooperation relationships among project contributors, previous research has tested the following principles:

- The probability with which a new relationship connects to a contributor is exponentially proportional to the number of existing relationships involving the contributor (Xu et al. 2006; Gao and Madey 2007). This law applies to a variety of social networks and is often referred to as the rich-get-richer evolution principle (Albert and Barabási 1999).

- As a consequence of the rich-get-richer principle, OS networks have a few nodes with a number of relationships significantly higher than the network’s average, called hubs. The alternative random evolution principle has been found not to apply to OS networks, i.e. relationships are not uniformly distributed across nodes (Xu et al 2006).

- Hubs have been found to follow a life cycle. Hubs appear, grow, and ultimately stop growing and quickly disappear (Gao et al. 2005).

- The creation of hubs is enabled by the ability of some nodes to evolve considerably faster than the average evolution rate of the network (Gao et al. 2005).

In the OS context, Grewal et al. (2006) have hypothesized that hub contributors have a positive impact on the success of the projects they are involved in. Their claim is that the rich-get-richer principle suggests that hub contributors have the ability to attract further contributions and, thus, positively influence the evolution of their projects. The literature provides numerous metrics that help verify whether a node is a hub, called centrality metrics (see Section 2). Grewal et al. (2006) have tested whether higher values of centrality metrics are positively correlated with the ranking measure of project success. By testing correlation for 12 projects from SourceForge.net written in Perl, Grewal et al. (2006) have found mixed results that only partially support their hypotheses. A possible problem with their approach to testing is the size of the social network that they have considered, which is limited to the contributors of a few projects and their direct connections to other contributors in the SourceForge.net community, while the literature clearly indicates that the laws governing a social network can be observed only if the network is analyzed in the large (cf. Newman et al. 2006).

In a previous paper (Barbagallo et al. 2008), we have tested Grewal et al.’s relationship on a significantly larger sample and actually found empirical evidence supporting the relationship between centrality and success. The relationship between centrality and project success has clear implications for managers. A company that is interested in making business through OS could aim at having its developers become hubs as a way to enhance success. However, in our previous research, we have also found that more successful projects have a lower technical quality. This paper aims at understanding the reasons behind the negative relationship between success and quality. Is success in and of itself detrimental to quality or are there drivers of quality other than success that help explain quality degradation? If so, are these drivers related to centrality?

Answering these questions can help managers understand the implications of gaining a more central position in the network as a way to reach success. The software engineering literature explains that there exists a trade off between quality and costs. Since quality represents an investment in the long term, managers are often willing to accept a lower level of quality if it represents a way to have a marketable output more quickly (Tan and Mookerjee 2005). Explaining the relationship between success and quality can help managers understand whether an OS social network represents a governance structure that allows deliberate quality to cost decisions or it creates new and negative quality degradation effects.

The presentation is organized as follows. Section 2 reviews the literature on social networking centrality metrics and software design quality metrics. Section 3 discusses our research hypotheses, while Section 4 describes the operationalization of variables, the data sample used to verify our hypotheses, our statistical

---

1 A definition of ranking in SourceForge.com is given in http://apps.sourceforge.net/trac/sourceforge/wiki/Project%20statistics
approach and reports the results of empirical testing. Finally, Section 5 provides a discussion of empirical findings and outlines possible directions for future research.

2 RELATED WORK

This section reviews the literature focusing on the concepts of centrality in social networks and software design quality.

2.1 Centrality in Social Networks

The concept of centrality has been defined as the importance of an individual within a network (Freeman 1979). Centrality has attracted a considerable attention as it clearly recalls notions like social power, influence, and prestige. Over time, several metrics have been introduced to formalize and then measure centrality from different points of view.

The first metric of centrality, called degree centrality, discussed by Freeman (1979), is defined as the number of links of a node normalized to the total number of links in the network. Degree centrality still represents the simplest and most widely used indicator of centrality, as it is intuitive and easy to calculate (Choi et al. 2006). A node that is directly connected to a high number of other nodes is obviously central to the network and likely to play an important role (Sparrowe et al. 2001). A node with a high degree centrality has been found to be more actively involved in the network’s activities (Hossain et al. 2006).

Freeman has also introduced the metric of betweenness centrality (Freeman 1979). This metric is defined as the average frequency with which a node is crossed by the shortest path connecting two generic nodes of the network. This metric is widely used in the literature, as it represents the simplest way to measure the ability of a node to reach other nodes in the network and act as an intermediary of the interactions between them. Over time, several refinements of the original Freeman’s metric have been proposed. For example, Newman (2003) has posited that a random walk among all possible paths should be considered as opposed to the shortest path. Although the opposite claim could be put forward too, Newman’s metric has the advantage of lowering the complexity of the algorithm to calculate betweenness centrality.

Freeman has also proposed the metric of closeness centrality, which is meant to extend the concept of betweenness centrality by measuring how far an actor is from all other actors in the network along the overall shortest path (Freeman 1979). This metric is less intuitive and more difficult to calculate than the previous two and has obtained a more limited success. Freeman notes that closeness centrality can be associated with the idea of independence of a node, since high values of closeness involve a lower need to depend on other nodes in order to communicate with other parts of the network. However, the metric becomes meaningless if applied to disconnected networks, as it cannot be calculated for non-reachable nodes. A more recent metric proposed by Stephenson and Zelen (Stephenson and Zelen 1989), named harmonic centrality, represents a measure of closeness centrality that considers harmonic distance in place of shortest path distance.

Previous to Freeman, Bonacich (1972) introduced the metric of eigenvector centrality, which measures centrality as the principal eigenvector of the whole network’s adjacency matrix. The disadvantage of this approach is that the eigenvector of the adjacency matrix must be calculated iteratively and convergence can be very slow. This metric has had limited success, especially due to its mathematical and computational complexity. Furthermore, Borgatti (1995) has noted that eigenvector centrality is conceptually similar to degree centrality. However, it should be acknowledged that the PageRank metric proposed by Brin and Page (1998) is based on the notion of eigenvector centrality. Even earlier than Bonacich, Katz (1953) and Hubbell (1965) introduced two centrality metrics, which, similar to eigenvector centrality, consider a node important if it is connected to other important nodes. However, both indices have series convergence issues that have limited their use in practice.

This paper focuses on degree and betweenness centrality, according to their original definition provided by (Freeman 1979). Degree and betweenness centrality represent the most intuitive and widely used metrics of centrality. We acknowledge that closeness and eigenvector centrality are also theoretically important metrics of centrality in the field of social networks, as discussed in (Borgatti et al. 2006). However, their greater conceptual and computational complexity makes them more difficult to use in empirical research on large social networks.

Proceedings ECIS 2009
Software Design Quality

This paper focuses on the quality of software design, i.e. on the internal quality of software. Previous literature suggests that higher values of software design quality metrics represent drivers of a number of external quality variables, such as testability, correctness, and reliability (Boehm 1976, Brito e Abreu and Melo 1996, Marinescu 2005). In turn, these external quality variables affect user satisfaction and can influence software adoption and actual usage (Bevan 1995). However, the direct analysis of external quality variables, i.e. of software effectiveness variables, is outside of the scope of the present paper.

Software design quality can be measured by analyzing the design properties of source code. There exists a consolidated body of literature focusing on code-based design quality metrics. Traditionally, the measurement of code design quality is based upon i) complexity and ii) design quality metrics. The first research contributions were aimed at providing operating definitions and metrics of software complexity, focusing on the analysis of the code’s information flow. Cyclomatic Complexity (McCabe 1976), Software Science (Halstead 1977), and Information Flow Complexity (Henry and Kafura 1981) represent the most widely used metrics from this early research.

Over time, design quality has become of increasing importance to cope with the continuously growing size of software systems. Research has started to distinguish between the complexity due to poor design quality and the inherent complexity of software due to requirements (Troy and Zweben 1993). The main contribution of these studies has been to show that design quality is necessary to handle the complexity caused by challenging requirements.

With the advent of the object-oriented programming paradigm, coupling, cohesion, inheritance, and information hiding have been identified as the basic properties of software design quality (Emerson 1984, Symons 1988, Chen and Lu 1993, Sharble and Cohen 1993). Based on these four basic properties, a number of metrics have been proposed to evaluate the design quality of object-oriented software. The most widely known metrics have been first proposed by Chidamber and Kemerer (1994) (WMC, NOC, DIT, RFC, LCOM, and CBO) and by Brito e Abreu (1995) (COF, PF, AIF, MIF, AHF, and MHF). These milestone contributions have started a lively debate within the software engineering community on the consistency and generality of such metrics (Harrison et al. 1998). As a matter of fact, metrics such as CBO, NOC, MIF, and DIT represent a standard and are included in most development environments, such as Eclipse and Visual Studio.NET. This paper focuses on these standard metrics.

CBO, NOC, and DIT have been found to impact on software maintainability and, hence, on maintenance effort and costs (Li and Henry 1993). Increasing software design quality is viewed as a costly activity that pays back in the long term by reducing the cost of subsequent maintenance interventions (Slaughter et al. 1998). With proprietary software, companies usually take a short-term perspective and tend to develop code faster at the expense of quality, which, in turn, tends to decrease over time (Tan and Mookerjee 2005). As observed by Tan and Mookerjee (2005), the deterioration of quality over time leads to a break-even time when a short-term perspective becomes economically inefficient and companies should invest in quality. This can be obtained either by replacing old software with new code of higher quality or by launching a maintenance initiative aimed at increasing quality without necessarily developing new functionalities, commonly referred to as refactoring (Fowler et al. 2001).

In OS applications these phenomena are difficult to observe. Some projects become inactive when they reach the end of their lifecycle and, until then, they are continuously maintained. However, projects reach their end for a number of reasons that may not be related to quality deterioration. For example, solo projects, i.e. projects launched and maintained by individual programmers, are often active for a very short period of time and come to an end due to lack of interest from the OS community.

The most successful projects, such as Linux and PostgreSQL, are still active although they are considered mature. Koch (2004) has noted that in OS projects refactoring tends to be a continuous process and developers allocate time and effort to quality improvements when needed. A previous work by Capra et al. (2007) has studied the refactoring process of a sample of 95 OS applications (1251 versions) from SourceForge.net. Empirical analyses have showed that the number of versions between two subsequent refactorings is highly variable. On average, a significant quality improvement can be observed in 40% of the total number of versions, while Tan and Mookerjee (2005) have found that in a sample of closed source applications refactorings occur in about 10% of an application’s versions.
Previous literature indicates that the cost benefits of quality improvements are reaped over time. However, it provides only partial evidence to demonstrate that quality investments have a positive balance (Slaughter et al. 1998). From a theoretical standpoint, Tan and Mookerjee (2005) suggest that quality investments typically represent a zero-sum game. However, the only clear empirical result is that quality involves an investment and, in the short term, it represents a cost. OS projects challenge also this result, since continuous refactoring practices should release similarly continuous cost benefits. A previous work by Capra et al. (2008) has empirically verified that quality and development effort are not correlated in OS projects, supporting the theoretical observations of Tan and Mookerjee (2005). In this paper, we consider software design quality and development effort as independent variables.

3 HYPOTHESES

In a previous paper (Barbagallo et al. 2008), we have found empirical evidence supporting the following hypotheses:

- Projects involving contributors with a higher level of centrality are more successful.
- Projects involving contributors with a higher level of centrality are able to attract a greater development and maintenance effort.
- More successful projects have a lower software design quality.

Overall, these hypotheses indicate that involving contributors that play a more central role in the social network helps projects to attract investments and eventually reach success. However, it does not seem to help the quality of the software artifact. These findings are counterintuitive, since they show that more successful projects (in terms of SourceForge.com ranking) have lower technical quality. In the paper, we discuss the possible reasons behind this inverse relationship between success and quality. A first, straightforward argument explaining previous findings is that more central contributors are also highly skilled developers who are well known for their ability to manage the complexity of code with a lower attention to the software structure (Lakhani and Wolf 2003). The consolidated metrics of software quality used by the authors in their previous research represent measures of code structure. Structure is generally considered a proxy of quality, as it represents the main driver of software maintenance costs (Li and Henry 1996).

This paper aims at providing empirical evidence supporting the idea that the negative impact of success on quality is caused by the careless behaviour of skilled developers, who are also hubs within the social network. Our first research hypothesis addresses the relationship between centrality and developers’ skills. In contexts different from OSS, the literature acknowledges that social networks are a significant enabler of knowledge transfer processes and related effectiveness (Reagans and McEvily 2003, Tsai 2005). In particular, higher levels of betweenness centrality in the network have been found to be correlated to a greater timeliness in reaching new information, while higher levels of degree centrality have been found to influence the ability to gain access to a broader information base (Mehra et al. 2001). In the OSS context, a more central position in the social network surrounding a project can foster the growth of the community by means of the rich-get-richer effect (Gao and Madey 2007). In turn, a more central position can help improve the communication with other network members, increase the number of cooperation opportunities within a variety of projects, help gain new skills, and ultimately develop a broader expertise. (Kidane and Gloor 2005) have found a correlation between group density and its performance and creativity in the case of Eclipse community. These benefits create a virtuous circle, contributing to the growth of developers as hubs of the network. This leads us to our first research hypothesis:

H1: More skilled contributors are involved in projects with a higher level of centrality.

Our second research hypothesis addresses the relationship between skills and quality. As noted before, skilled developers have the ability to solve problems quickly and efficiently, but are less willing to design code according to the software engineering principles ruling structured development (Lakhani and Wolf 2003). It has been demonstrated that organizations tend to apply their most skilled developers in complex tasks, where devising an efficient algorithm is more important than obtaining a well-structured and maintainable software (Faraj and Sproull 2000). These results are interesting in a OS context, since we know from Capra et al (2008) that the governance model of an OS project can resemble that of a closed source project, with a single sponsor company developing the bulk of code. Furthermore, hub nodes tend to experience an exponential growth, as the probability to gain new relationships has been found to be exponentially proportional to the number of existing relationships (Xu et al. 2006, Gao and Madey 2007).
The rate of new change requests (new features, bug reports, etc.) grows accordingly. Hypothesis H1 posits that more skilled developers are involved in more central projects and, therefore, can benefit from all the advantages of their central position in terms of knowledge transfer. However, they are also burdened by a greater information load related to communicating and coordinating with other contributors in development activities and also in a number of non-development activities, including debugging, translation, advisory, and documentation. Hinds and McGrath (2006) have empirically found that highly connected social networks (that is, networks with a high average value of degree centrality) cannot be considered as an effective support to distributed project development. Findings prove that the coordination overhead caused by the extension of the network increases the effort required for software development and maintenance. Therefore, more central developers can be supposed to spend more time in coordination and less time in quality, which, in the short term, represents a time investment that they may not be able to afford given their information overhead. This leads to our second research hypothesis:

H2: Projects involving more skilled contributors have a lower software design quality.

4 METHODOLOGY AND RESULTS

This section presents the operationalization of the variables involved in our testing and the data sample used for empirical verifications.

4.1 Variable Definition and Operationalization

Network model. We model OS social networks as two-mode undirected affiliation networks (Wasserman and Faust 1994) with two types of nodes: developers and projects. A node representing a developer, say d, is associated with another node representing a project, say p, when d is a member of p’s team of contributors. Two distinct one-mode networks can be derived from a two-mode network by considering either developers or projects only:

- Developers network. All nodes represent developers. Two nodes are linked when both developers are members of the same project team.
- Projects network. All nodes represent projects. Two nodes are linked when corresponding projects have at least one developer in common.

Metrics of centrality. The degree centrality (Freeman 1979) \( c_d(n_i) \) of node \( n_i \) is defined as the ratio of the number of edges involving node \( n_i \), \( \rho(n_i) \), to the total number of nodes in the network excluding node \( n_i \):

\[
c_d(n_i) = \frac{\rho(n_i)}{(N-1)}.
\]

The betweenness centrality (Freeman 1979) \( c_b(n_i) \) of node \( n_i \) is defined as the average frequency with which a generic node \( n_i \) crosses node \( n_p \) to reach a different node \( n_k \) trough a shortest path:

\[
c_b(n_i) = \frac{2}{(N-1)(N-2)} \sum_{j\neq i} g_{jk}(n_i) g_{jk}^{-1},
\]

where \( g_{jk} \) represents the total number of shortest paths from \( n_j \) to \( n_k \) and \( g_{jk}(n_i) \) represents the number of shortest paths between nodes \( n_j \) and \( n_k \) crossing \( n_i \). The metric is normalized to the maximum number of shortest paths crossing \( n_i \) in an undirected network with \( N \) nodes. Betweenness centrality is a measure of the ability of a node to control the information flows in the network. A node with a high betweenness centrality can be considered as an important information broker for the network, as it is likely to receive and convey many information flows (Hossain et al 2006).

Degree and betweenness centrality for developers’ and projects’ networks are indicated with apex \( d \) and \( p \), respectively. For the sake of simplicity, we refer to the degree and betweenness centrality of nodes in the developers’ networks as developer degree centrality (\( c_d^d \)) and developer betweenness centrality (\( c_b^d \)), respectively. Similarly, we refer to the degree and betweenness centrality of nodes in the projects’ networks as project degree centrality (\( c_d^p \)) and project betweenness centrality (\( c_b^p \)), respectively.

Metrics of software design quality. Two of the most referenced suites of object-oriented design metrics have been included in our metrics’ set, as suggested by Harrison et al. (1998): the MOOD metrics’ set for the evaluation of quality at the software system level (Brito e Abreu 1995), and the Chidamber and Kemerer (1994) metrics’ suite for the evaluation of quality at the class level. The four chosen metrics (MIF, NOC, MOOD, and MOOM).
CBO, DIT) are the most preferable to provide measures of inheritance and coupling, which are two of the three sets of fundamental quality metrics, and also the most studied in literature (Capra et al. 2008).

**Metrics of skills.** Two metrics deal with skills in our model, skill level and skill range. The skill level metric $SL_k$ of project $k$ is defined as the mean value of the average per-skill level of the developers involved in the project:

$$SL_k = \sum_{i \in k} \frac{SK(i, s)}{MAX\_VAL \cdot ns(i)}$$

where $SK(i, s)$ is the level of skill $s$ of developer $i$, $S(k)$ is developer $i$’s set of skills, $MAX\_VAL$ is a constant representing the maximum skill level (5 in SourceForge.net), $ns(i)$ is the total number of skills of developer $i$, and $nd(k)$ is the total number of developers working at project $k$.

The skill range metric $SR_k$ of project $k$ is defined as the ratio between the number of distinct skills owned by at least one developer in the project and the total number of skills in the project:

$$SR_k = \frac{NDS_k}{NS_k}$$

where $NDS_k$ and $NS_k$ are the total number of distinct skills and the total number of skills in project $k$, respectively. The skill range measures the diversification of skills within a project. For example, if a project has two developers with the same skill set the metric will evaluate to 0.5, while if the skill sets of the two developers are disjoint the metric will be equal to 1.

### 4.2 Data Sample

The data set used for this study has been gathered by analyzing a sample of OS community applications from the SourceForge.net repository. Data on skills have been extracted by analyzing each developer’s profile on SourceForge.net. Since mining on line repositories (such as SourceForge.net) can lead to controversial results because of the varying quality of available data (Howison and Crowston 2004), a first sample of applications (AS1) has been selected according to the following criteria:

- **Project maturity:** active and beta status or higher (inactive and less mature applications have been excluded because of their instability and low significance).
- **Version history:** at least 5 versions released.
- **Programming language:** Java.
- **Domain:** selected applications are uniformly distributed across the SourceForge.net domain hierarchy.

A second sample of applications (AS2) has been considered to allow the correct evaluation of the social networking metrics described in the previous section. Applications belonging to sample AS2 have been selected by relaxing some of the criteria used to select applications of sample AS1. AS2 includes all the active projects of SourceForge.net written in Java. This sample has been used to build a SourceForge.net social network as wide as possible, with the aim of overcoming network size limitations of previous research (Newman et al. 2006). As suggested by Marcoulides and Sounders (2006), confidence intervals at level $\alpha = 0.05$ have been computed for the variables involved in hypotheses testing in order to assess the adequacy of our sample size. Although the sample is not extremely wide, the confidence intervals are relatively narrow for all the considered variables. Data on all the applications of samples AS1 and AS2 refer to June, 30th 2007 to guarantee the temporal consistency of the data sets. Table 1 presents cardinalities of application samples AS1 and AS2.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Dataset AS1</th>
<th>Dataset AS2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of projects</td>
<td>56</td>
<td>29,836</td>
</tr>
<tr>
<td>Number of developers</td>
<td>378</td>
<td>57,142</td>
</tr>
</tbody>
</table>

*Table 1. Cardinity of datasets AS1 and AS2*

Social networking metrics have been derived from the analysis of the social network of the applications included in sample AS2. Social network data has been derived from the SourceForge.net data warehouse.
(Gao et al. 2007) for all the applications in sample AS2, and processed by a tool developed ah-hoc. The computation of social networking metrics has been performed by analyzing the SourceForge.net social network with Pajek (Batagelj and Mrvar 1997), one of the most used and referenced tools for large networks analysis.

Software quality metrics have been evaluated by analyzing the source code of all available versions of each project in our dataset AS1, and considering average values for each metric. Source code has been analyzed with a tool developed ad-hoc. The tool provides data on all the software quality metrics described in Section 4.1, performing static analyses of Java source code. The static analysis engine is based on the Spoon compiler (Pawlak 2005), which provides the user a representation of the Java abstract syntax tree in a meta-model that can be used for program processing.

Statistical analyses and structural equation model testing have been performed with SPSS and AMOS.

4.3 Measurement Model

The measurement model has been defined to verify the assumption that social networking and software design quality metrics actually measure different aspects of the same phenomena. A principal component analysis (PCA) has been performed on both sets of metrics to verify the assumption. Results of such analysis are shown in Table 2.

Figure 1. Measurement model of design quality

Figure 1 presents the measurement model related to the design quality set of metrics. In this case only one factor has been extracted, and has been labelled as Design Quality.

Figure 2. Measurement model of centrality

Figure 2 shows the measurement model related to the social networking set of metrics. As it can be noted, two different factorization variables have been identified, related to the centrality of developers and of projects, respectively. In turn, these two latent variables have been found to be different aspects of the same concept, that has been named Centrality.

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Independent Variable</th>
<th>Standardized Regression Weight (b)</th>
<th>Standard Error</th>
<th>p-value</th>
<th>Composite Factor Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{a}$</td>
<td>Developer Centrality</td>
<td>0.688</td>
<td>1.311</td>
<td>&lt;0.001</td>
<td>0.830</td>
</tr>
<tr>
<td>$C_{b}$</td>
<td>Developer Centrality</td>
<td>0.680</td>
<td>-</td>
<td>-</td>
<td>0.684</td>
</tr>
<tr>
<td>$C_{a}$</td>
<td>Project Centrality</td>
<td>1.079</td>
<td>0.049</td>
<td>&lt;0.001</td>
<td>0.970</td>
</tr>
<tr>
<td>$C_{b}$</td>
<td>Project Centrality</td>
<td>0.896</td>
<td>-</td>
<td>&lt;0.001</td>
<td>0.915</td>
</tr>
<tr>
<td>Developer Centrality</td>
<td>Centrality</td>
<td>1.232</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Project Centrality</td>
<td>Centrality</td>
<td>0.722</td>
<td>1.649</td>
<td>&lt;0.001</td>
<td>-</td>
</tr>
<tr>
<td>NOC</td>
<td>Design Quality</td>
<td>-0.983</td>
<td>-</td>
<td>-</td>
<td>0.911</td>
</tr>
<tr>
<td>CBO</td>
<td>Design Quality</td>
<td>-0.697</td>
<td>0.525</td>
<td>&lt;0.001</td>
<td>0.740</td>
</tr>
<tr>
<td>DIT</td>
<td>Design Quality</td>
<td>-0.973</td>
<td>0.074</td>
<td>&lt;0.001</td>
<td>0.900</td>
</tr>
<tr>
<td>MIF</td>
<td>Design Quality</td>
<td>-0.748</td>
<td>0.099</td>
<td>&lt;0.001</td>
<td>0.894</td>
</tr>
</tbody>
</table>
Table 2. PCA results and estimates of regression weights for the measurement models of Figure 1 and Figure 2

Table 2 shows the results of the PCA performed on the different parts of the measurement model, along with the standardized regression weights of the relationships between latent and observed variables. Results show that all the factorizations should be considered acceptable, since all the composite factor reliability values are greater or very close to the threshold value of 0.700 as suggested by Bagozzi and Yi (1988). All the relationships considered between the set of observed variables (either referred to centrality or software design entropy) and the latent variables are significant with $p < 0.001$: this confirms that the factorizations in the measurement model were performed correctly.

4.4 Structural Model Testing

The estimation results of the research model in Figure 3 used to test the research hypotheses are shown in Table 3. Please note that, for sake of simplicity, the model in Figure 3 is a simplified version in which the factorizations related to the latent variables Design Quality and Centrality discussed in the previous section are not shown. Variables Design Quality, and Centrality have been controlled by project age Age, as suggested by Banker and Slaughter (2000), although the controlling variable Age is not shown to reduce the complexity of the model in Figure 3.

![Figure 3. Structural model for the verification of research hypotheses](image)

![Table 3. Estimates of regression weights for the research model of Figure 3](table)

All the relationships hypothesized between model variables are significant with $p < 0.05$, that is, can be accepted at a significance level $\alpha = 95\%$. Table 4 shows that the overall model fit is satisfactory.

![Table 4. Goodness of fit indices for the research model in Figure 3.](table)

Research hypothesis H1 (more skilled contributors are involved in projects with a higher level of centrality) is represented within our model by the regression relationship between Centrality (independent variable) and Skill Level (dependent variable). Our analysis shows that the regression weight is positive and statistically significant ($b = 0.329$, $p = 0.045$). Consequently, research hypothesis H1 is verified.
Hypothesis H2 (projects involving more skilled contributors have a lower software design quality) is tested by the relationship between Design Quality and Skill Level, where the former is the dependent variable. Since the regression weight is negative (b = -0.556) and the estimation of this relation is significant (p = 0.050), hypothesis H2 is verified. Please note that Skill Level is controlled by Skill Range in order to consider not only the depth of skills but also the heterogeneity.

5 DISCUSSION AND CONCLUSIONS

Empirical results support hypothesis H1, thus confirming centrality as a significant driver of the expertise of developers. Therefore, knowledge seems to represent a first, fundamental benefit of being a hub node in an OS network. Previous to our work, the literature has described the role of social networking as enabler of knowledge transfer processes and related effectiveness. As noted in Section 1, our work is the first empirical contribution providing empirical evidence to support the relationship between network centrality and expertise. In particular, we have made an effort to overcome some of the limitations of previous research by measuring our centrality metrics on the largest network of Java projects that can be built from SourceForge.net, which includes all the relationships among contributors, both direct and indirect. It should be noted that the metrics of centrality have been originally defined to account for both types of relationships (Freeman 1979).

Results also support hypothesis H2, suggesting that more skilled developers will tend to produce a software with a lower design quality. Even if this seems counterintuitive, it should be noticed that hubs in OS communities are involved in several projects and incur a coordination overhead that reduces the share of time that they can devote to quality. The literature indicates that quality represents an investment that requires time and provides its pay-offs over time in terms of lower maintenance costs (Slaughter et al. 1998). The coordination overhead caused by a more central role may reduce the willingness of developers to invest in quality.

It should be noticed that in our model the relation between success and design quality remains significant, thus indicating that development skills are only one of the factors impacting on quality. Future research will investigate the possible influence of other factors that, overall, may have a mediation effect on success.

Our results show that centrality metrics are significant proxies of developers’ skills that should be monitored from the perspective of a project administrator or team manager. However, they also prove that projects with more skilled team members tend to have a lower design quality of software. This has a number of potential consequences that might be visible to users and could cause negative effects over time. From previous case studies such as Mozilla and Eclipse, it is clear that in order for a social network of a project to become a global success, a refactoring intervention is needed, with a consequent infusion of large investments. The natural behaviour of the social network does not seem to be able to cope with these levels of growth. While it can help a project to start growing and reach a significant level of success, above a certain level it may represent a weak management lever. This is consistent with previous literature positing that excessively large social networks have a lower effectiveness (Herbsleb and Mockus 2003) and represents an interesting subject for future research.

References


Do Best Practice Frameworks Fit Open Source Software Customization?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0595.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>FLOSS (Open Source), Management practices, Action research, Distributed IS project</td>
</tr>
</tbody>
</table>
DO BEST PRACTICE FRAMEWORKS FIT OPEN SOURCE SOFTWARE CUSTOMIZATION?

Keßler, Steffen, Institute of Information Systems, Philipps-Universität Marburg, Universitätsstraße 24, 35032 Marburg, Germany, steffen.kessler@wiwi.uni-marburg.de

Alpar, Paul, Institute of Information Systems, Philipps-Universität Marburg, Universitätsstraße 24, 35032 Marburg, Germany, alpar@wiwi.uni-marburg.de

Abstract

Most management models that support software development, information technology (IT) service management, or other IT related tasks have been developed based on extensive experience with these tasks. Their recommendations are, therefore, often named base, good, or best practice. We examine in this paper whether these models are also suitable to support use and adaptation of open source software (OSS) within adopting organizations. OSS is widely used as it exhibits some very appealing features. However, if companies take full advantage of the flexibility OSS affords, especially the possibility to customize it, relating processes get quite complex. Careful control of these processes becomes crucial and use of proper management processes almost mandatory.

In the paper, we first determine the special needs of OSS customization and then examine three popular models as to how well they can support this activity. The examined models are ITIL Version 3, SPICE (ISO/IEC 15504), and V-Modell XT.

Our research shows that none of the models sufficiently covers the needs of OSS customization. While some aspects of OSS customization can be dealt with in the models or require only minimal modifications of the models, none of the models includes best practices for the management of intertwined and concurrent internal and external development.

Keywords: Open Source Software, best practice, customization, adaptation, software development.
1 INTRODUCTION

OSS is developed by many individuals and many organizations. An organization that uses and customizes OSS must carefully manage the development process in order to achieve high quality software (SW) and avoid errors. Frameworks or models for management of SW development were mainly designed for development under single control. Therefore, our research question is: Can models for SW development and use directly be employed by an organization that customizes OSS for its own purposes? In the first subsection, we briefly describe some aspects of OSS that are important in our case. Then, we briefly recall the rationale for software development models and we name three models that we want to examine in detail.

1.1 Open Source Software

The term OSS denotes SW which source code is available under a copyright license in compliance with the Open Source Definition (Open Source Initiative 2008). It has to be distinguished from the term Free Software which is defined by the Free Software Foundation (Free Software Foundation 2008) founded by Richard Stallman. It originated the popular OSS Licence GNU GPL. The Free Software Foundation places an emphasis on the freedom of the SW; for this paper we chose the term Open Source Software (OSS) to place an emphasis on the availability of the source code.

OSS has gained popularity mostly through Linux, an operating system project started by Linus Torvalds, but joined by many developers distributed all over the world (Raymond 2001). Linux was first popular with technically versatile persons but is now available in distribution packages that do not require extensive computer skills. Its use increased with the diffusion of Internet. One example of an OSS commonly used with Linux is the Apache web server [http://httpd.apache.org]. While both Linux and Apache offer general core functions, increasingly OSS for specific applications is gaining popularity, e.g., in the area of business intelligence The Eclipse BIRT (Business Intelligence and Reporting Tool) Project, [http://www.eclipse.org/birt/phoenix/]). Even Enterprise Resource Planning Systems are meanwhile available as OSS (e.g., SQL Ledger [http://www.sql-ledger.org/].

While OSS needs not necessarily to be available at no cost, it often is. Therefore, it is highly relevant for companies trying to lower licence fees. In fact, costs seem to be the main reason for OSS adoption (e.g., Dedrick and West 2004).

The development of OSS by many distributed developers is referred to as an OSS project. The core developers provide an “official” release. Adaptations between releases may be published by the core developers or anyone else. However, there may be also adaptations executed within an organization that are not made available to the public. Adaptations by firms outside of the SW industry seem to be rare. There are different reasons for the reluctance of organizations to modify OSS. First, they may not have a need to do so. Second, they may not possess the skills or capacity to do so (Madanmohan and De’ 2004). There is also a difference in attitude towards OSS adaptation depending on the type of OSS. Organizations are more willing to adapt application-oriented non-infrastructure SW than infrastructural SW (Wichmann 2002, Dedrick and West 2004). Both, adoption and adaptation of OSS may be hindered by licensing intricacies. Some OSS licenses require that all modifications are provided to the public. This is just the opposite of why companies customize standard SW: because they do things differently than others. They usually do not want to reveal their processes to others.

According to a recent citation of a study by Gartner (Morgan 2008), OSS is used in 85 percent of all companies, the other 15 percent planned to do so within the next 12 months. Given this state, whether organizations customize OSS or “just” use it without own adaptations, the question arises whether the models considered below support OSS use. In fact, a possible lack of adequate support may lead to problems or create an obstacle to a wider adoption and use of OSS.
Even if companies do not adapt OSS after adoption, the quick release of new versions of OSS, modifications that are not part of an official release, and possible forking in OSS projects require an active management of OSS use. Therefore, it is worthwhile examining whether the specific needs of OSS are covered by the models.

The deployment of OSS is similar to the deployment of commercial off-the-shelf SW (COTS). Both need to be customized to fit the requirements of the adopting organization. Configuration and parameterization are the approaches of choice in this case. If they are not sufficient, customization of the SW code is necessary. With COTS, customizations are limited to customizations utilizing Application Programming Interfaces (APIs) provided by the vendor. With OSS, anyone can modify the source code. This flexibility offers great advantages but also leads to more managerial complexity. Since research on OSS customization within organizations is rare, we first need to identify the different events in the lifetime of an OSS that can occur before we can analyze how well the models for SW development and use support OSS use.

Research on maintenance of OSS does exist. But it covers maintenance processes within the OSS projects (e.g., Koponen and Hotti 2005). To our knowledge, no research on the case of internal OSS maintenance in companies does exist. As a result, the interactions between the process of adaptations in companies and the ongoing development process within the OSS project have not received much attention yet. This intertwined process may already start with the deployment of OSS within the organization. In this context, it is especially of interest whether the below mentioned models can also support such important tasks as OSS configuration and change management within organizations.

1.2 Frameworks for Software Development and Use

SW development, maintenance, and its continuous use are intellectual tasks that are difficult to manage. Given the economic and legal importance of these tasks, they must be controlled and managed. A number of models or frameworks have been developed towards these goals. They should help to produce high quality SW efficiently (Sommerville 2007, pp. 692). Both of these characteristics have often been badly missed in SW development and SW products (Boehm 2006). Most of the management models have evolved based on the experiences gained in many SW projects in many organizations. Practices that led to desired outcomes have been gathered and further improved so that the resulting frameworks are considered to contain best practice. A best practice process denotes a process, which is the most effective and efficient process for a certain task. A bundle of such processes may become the standard for this specific task and can be applied in diverse organizations.

We chose for our research three frequently used frameworks to represent models for different purposes. They contain guidelines which partly relate to the same tasks (e.g., change management).

Any organization that develops SW systems should follow a system development model. Such models support project management from systems analysis through decomposition and integration including quality assurance. An example of such a model is the V-Model where activities are arranged according to the letter V (Sommerville 2007, pp. 110). The use of its German version, currently called V-Modell XT (V-Modell n.d.), is mandated for systems developed by or for the German federal administration (Lange 2009). Many private organizations make use of it as well.

The use of a SW system in an organization can be considered a service to its internal and external users. The framework called ITIL (Information Technology Infrastructure Library) has been developed with the goals of efficient delivery of such services and high service quality. ITIL is a collection of concepts and guidelines for IT planning and operations that are supposed to represent best practices [http://www.itil-officialsite.com].

A third type of models supports the assessment of the capability of an organization to develop high quality SW. The international standard ISO/IEC 15504 (ISO/IEC 15504-5 2006) constitutes an example of such a framework. This standard, often referred to as SPICE (Software Process Improvement and Capability Determination), defines processes grouped in process categories, capability levels which the processes can achieve, and an assessment model which guides the
measurement of processes with respect to their capability. SPICE is especially relevant to SW vendors but can also be valuable to large organizations developing large programs for their own use.

It is not easy for an organization which wishes to achieve the right level of managerial control of IT use to choose the appropriate combination of frameworks given their number, their overlap, differences in terminology and structure (Paulk 2004). We address yet another question in this context as mentioned above: How well do these models and frameworks support the full use of OSS?

2 RESEARCH APPROACH

Given the scarcity of relevant research and the difficulty to collect detailed data through survey research, we decided to observe an OSS adoption and customization process long enough to be able to induct the managerial needs of these processes. Therefore, we chose to follow the paradigm of action research (Avison et al. 1999). Of course, this approach does not guarantee that we will discover all the events that take place in the life of customized OSS but if we observe enough release changes, we should be able to give an initial answer to our research question.

2.1 Action research

Action research is a qualitative participatory method. The researcher becomes part of the project team in order to study the object of research in all necessary detail. It is widely used in the discipline of information systems (IS) as an interpretive research method (Baskerville 1999, de Villiers 2005). Qualitative research in IS is often applied in new problem domains where empirical data is unavailable or available only in small sample sizes. The process of action research consists of four consecutive sub-processes: plan, act, observe, and reflect (Zuber-Skerritt 1993). Our plan consisted of decisions which information on OSS deployment and adaptations to record and what tools to use. Since proper documentation of code versions and changes is a necessary action in OSS use, this was in the same time our “action” part and observation activity. We archived source code as well as change and feature requests. During our research, seven iterations triggered by official releases of the OSS were examined. One external release was not incorporated into the internal code base by the development team. The collected documents were examined in the reflection phase.

Action research, with its root in the behavioral sciences and founded by Kurt Lewin and other researchers in the 1940s (Peters and Robinson 1984) uses a practitioner-researcher relationship and is participative (Avison et al. 1999, de Villiers 2005). The study of an existing product development over several iterations distinguishes this research method from other interpretative research methods, e.g., research during the creation of a new product (de Villiers 2005). Using action research, researchers participate as well as criticize in the process and, therefore, influence it. As a result, action research enhances the practice and is able to reveal new knowledge.

In addition to the study of the internal code development, we also studied code changes in the external OSS community that did not enter the official releases yet. All internal and all external releases were archived for ex-post examination. Minor internal releases were triggered by internal adaptations or non-official external code used in the internal code base; major internal releases were triggered by the aforementioned official external releases.

Close cooperation with the project team allowed us to gain insights regarding several code changes within the official code that were introduced into the internal code base and led to incompatibilities between the official and internal release. These were examined by the project team manually. Since new releases were never adopted completely, files with changes in the official release as well as in the internal code base needed to be updated manually.

2.2 OSS case

Most OSS projects are in constant flux. They follow the “release early, release often” approach (Golden 2005, pp.19-20) because they are not restricted by business practices and strategies of COTS.
New releases may include bug fixes and new features. The constant changes require formalized support processes (Hoyer et al. 2007). The OSS characteristic of continuous change certainly applies to the OSS system we studied. It is a special-purpose web-based OSS Content Management System (CMS) that includes some typical Web 2.0 features. It was adopted and adapted by a small entrepreneurial company. The OSS system was developed within the Pligg CMS project [www.pligg.com] starting in 2005. It was chosen for use in our case in 2006, when our observation immediately started. The CMS instance we studied is online at http://www.colivia.de.

Users of the system can submit news items that are placed in a waiting queue. The news item is usually a short reference to another website. If other users find the news interesting, they can comment on it and vote for it. News that receives votes above a certain limit is placed on the first page which gives it more exposure. Each vote is not necessarily equal. Its weight may depend on the previous activity of the user who is giving the vote. The exact vote calculation is usually not disclosed. The same is true for the calculation of a reputation score for each user. News items are tagged and assigned to groups. Most common tags within a certain time period are displayed in a tag cloud. Such sites are referred to as social news sites. Pligg has also been used for other applications. Some of the sites built with Pligg are http://www.dealigg.com, http://www.ecofuse.com and http://software.intel.com/sites/coolsw/. Pligg is used all over the world, 40,000 users are registered in its support forums.

3 EVENTS IN OSS CUSTOMIZATION

3.1 Software evolution

As a basis for an OSS evolution process in companies, we use the system evolution process (Fig. 1) proposed by Arthur (1988) and amended by Sommerville (2007, p. 540). The SW evolution process is the result of a sequence of several maintenance iterations. Maintenance can be divided into adaptive, perfective and corrective maintenance (Swanson 1976). While corrective maintenance addresses errors found within the SW product, perfective maintenance is used to increase performance and maintainability as well as minimize inefficiencies. Adaptive maintenance is used to adapt SW to a changing environment. Maintenance often starts with the deployment of a SW product in an organization and stops with its retirement (Bennett and Rajlich 2000).

While maintenance has to be seen as a continuing process throughout the life of a SW product, this is contrary to the usual organization of SW tasks in projects. As a result, SW maintenance is usually organized as a sequence of maintenance projects (Kneuper 2003, p. 84f.).

Studies of SW that evolves over time (E-Type products) led to several research papers often referred to as “Lehman’s Laws” (Lehman and Ramil 2001). They include the proposal of continuing change and increasing complexity during the process of use of a SW product as an E-Type product. Here, we focus on the interdependencies between the concurrent external and internal development of an OSS system as well as the influence of the external releases on the internal code base.
3.2 Events during OSS customization

Even before the deployment of a SW system within the company, required changes are identified, their impact is analyzed, the change is planned and carried out resulting in a new system; change requests after deployment initialize the same steps (Sommerville 2007, p. 540).

Based on our research of the OSS CMS customization, a distinction of two process chains in the evolution of an OSS within an organization can be made. Both process chains use the generic system evolution process from Fig. 1 as the basis but include different process steps to establish a new system release. This distinction is due to the differentiation of internal change requests, triggered from inside the company, and external change requests, triggered by external events, e.g., new official releases or bug fixes available in the OSS project’s support forums. A firm may opt to ignore the second kind of change requests, but it then loses over time the free resources of the OSS developer community. Based on this process model, we identified several tasks in the areas of configuration and change management which are required in the case of entangled internal and external customizations of OSS.

Special requirements within the configuration management

Internal source code changes of OSS in a company require the establishment of an internal code base different from the code within the official OSS release. For OSS projects, a situation with two competing code bases would be called a fork; such a situation should be avoided to minimize inefficiencies (Raymond 2001, pp. 72-73). In OSS customization, the configuration management is required to manage the diverse code branches and minimize inefficiencies by supporting a merging of the internal code and new official releases.

For the internal code branch, the configuration management needs to keep track of both internally programmed code as well as imported external code (e.g., a bug fix available within the OSS project’s support forums). Official code versions of Pligg, like in many other OSS projects, are stored in the version control system Subversion (SVN). External code which becomes available between two official releases in SVN, support forums or in another place, and is used for modifications of the internal code tree, has to be treated as internal code until it is integrated into the official release and the official release is integrated into the internal code base. As a result, a distinction between internal and external code is not sufficient in the internal documentation. The origin of external code that is used in the internal version but has not become part of the official release (yet) has to be recorded as well. Any modification to the internal code base may be the cause for incompatibilities between the internal and official release during the next update.

During the observation period, conflicts caused by incompatibilities of internal changes with changes in a new official release occurred in two iterations, while internal adaptations got obsolete in five iterations. As a result, incompatibilities had to be resolved by modifications to the internal code base if the official release were to be adopted.

Special requirements within the change management

In the life-cycle of the OSS CMS we studied, in case of an internal request for change external sources were always checked first to save efforts. The change management was responsible for the search, assessment and possible adaptation of appropriate code. The search included resources provided by the OSS project, in our case the OSS support forums and the SVN repository. Only if no suitable code from the OSS community was found, internal code was developed. In both cases, this new code remained in the internal release at least until a new official release became available. With a new official release available, internal modifications were checked against the new external code changes.

In general, OSS projects use diverse places to document changes made to the SW. OSS users will, therefore, usually need to monitor several places for relevant changes. In addition to communication tools, e.g., discussion forums and mailing lists, code changes are also entered into bug reporting tools (e.g., Bugzilla). In some OSS projects, users and developers may submit all kinds of source code changes, including updates, fixes, and sometimes even new features, rather than only bugs in such
tools (Michlmayr, Hunt and Probert 2005, Koru and Tian 2004). As mentioned above, OSS projects often utilize revision control systems for configuration management purposes. Usually, all source changes will be tracked with these systems. Read access to revision control systems is usually unrestricted so they are an additional resource to search for relevant code.

The retrieval of information regarding new official releases was also a task assigned to the change management as well as the merging of the code bases in case of a new official release. The merging was supported by the configuration management with information about prior releases; also, the configuration management was responsible for the mapping of the branch merging.

We summarize the observations described above in Table 1.

<table>
<thead>
<tr>
<th>Configuration management</th>
<th>Change management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manage different versions of the same configuration element</td>
<td>Retrieve information regarding new releases from the OSS project (website, news group, mailing list, etc.)</td>
</tr>
<tr>
<td>(separate branches)</td>
<td></td>
</tr>
<tr>
<td>Build and archive baselines of relevant configuration items</td>
<td>Assess external components; adopt if useful and adapt, if necessary</td>
</tr>
<tr>
<td>in case of the adaptation of external components</td>
<td></td>
</tr>
<tr>
<td>Support mapping of branch mergers</td>
<td>Merge internal and external branches in case of new official releases</td>
</tr>
</tbody>
</table>

Table 1. Additional tasks for maintenance in case of OSS customization

Code feedback

A task which may be assigned to the change management as well as to the configuration management is code feedback to the OSS project. It was not performed during our action research study, but code feedback can be beneficial to both the company that gives feedback as well as to the OSS project. If internal changes get integrated into the official release, official code can replace the corresponding internal code making later configuration management easier.

4 ANALYSIS OF FRAMEWORKS

All three models were examined based on original documents. The basic structure of each model is given first. Then, we report the results of our analyses referring to the specific chapters.

4.1 V-Modell XT

The V-Modell XT (V-Modell n.d.) is a SW development model widely used in the German government sector. It is divided into nine parts:

- Part 1: Fundamentals of the V-Modell
- Part 2: A Tour through the V-Modell
- Part 3: Reference Tailoring
- Part 4: Reference Roles
- Part 5: Reference Work Products
- Part 6: Reference Activities
- Part 7: Reference Mapping to Standards
- Part 8: Annex
- Part 9: Templates

Figure 2. Structure of the V-Modell XT based on (V-Modell n.d., p. 1)
Within Part 4, relevant roles for configuration management and change management are defined, e.g., Configuration Management Administrator and Change Request Manager. The defined roles are sufficient for configuration and change management in the OSS context.

The view of the V-Modell XT is not limited to internal developments; it does incorporate the assessment and use of external products, even including market surveys of off-the-shelf products (Chapter 4.3 of Part 5). External SW modules are mentioned in Chapter 4.15 of Part 5. Links to chapters with further information within the standard are provided, including a link to make-or-buy decisions. Another relevant part for OSS customization is product evaluation (Part 6, Chapter 3.5); accordingly, external components can be evaluated before they are integrated into the internal code branch as non-official external code.

The possibility of code feedback is contained in the model, but it is based on the assumption that both the supplier and the buyer of SW products are using the V-Modell XT; feedback is provided in form of a Statement of Acceptance (Part 6, Chapter 5.1), which does not really fit in the case of OSS.

The V-Modell XT does support dividing of projects into partial projects (Part 3, Chapter 5.21), but there is no explicit support of ongoing concurrent development using separated code branches. For external SW units, an adoption procedure is defined in Chapter 3.8.12 of Part 6, which also does not reflect the concurrent development during the OSS life cycle. As a result, the handling of the code branches has to be represented via separate baselines or with the official code base carried within a baseline as a special configuration item (CI) within the configuration management. A baseline is a set of all relevant work products for one specific configuration.

In summary, V-Modell XT does include adequate practices for assessment and adoption of external components. It also offers some ideas which could be used for concurrent development in OSS projects in companies. However, it does not offer a consistent treatment of concurrent development.

### 4.2 IT Infrastructure Library (ITIL)

The IT Infrastructure Library (ITIL) is an example of best practices within IT service management.

They are grouped in functions or processes which are themselves assigned to five core volumes in ITIL Version 3:

<table>
<thead>
<tr>
<th>Service Strategy</th>
<th>Service Design</th>
<th>Service Transition</th>
<th>Service Operation</th>
<th>Continual Service Improvement</th>
</tr>
</thead>
</table>

*Figure 3. Core volumes of ITIL Version 3*

Here, the volume Service Transition is especially relevant, as it “covers the broader, long-term change management role, release and deployment practices, so that risks, benefits, delivery mechanisms and the support of ongoing operational services are considered” (OGC 2008). The subsequent statements are based on the official German translation of the OGC Service Transition volume (OGC 2007).

Chapter 3 presents the Service Transition (ST) Principles; best practices and principles are contained here in several policies. Policy 3.2.2 confirms that all changes of service should be done by the service transition; the second best practice requires that internal and external changes have to be distinguished. While this is essential in case of OSS, according to our research it is not sufficient, as in the case of OSS customization with external code which is not part of the official release, three code types have to be distinguished. Policy 3.2.6 concerns the initiation and maintenance of relations to stakeholders. A best practice suggests building relationships with stakeholders (including suppliers). Since the OSS project can be considered a kind of supplier, this relationship can be the basis for feedback to the OSS project. According to a best practice in Policy 3.2.9, any update of a release has to be recorded within the configuration management system. This is in line with the required archival of all changes to CIs.
by the configuration management determined in our research of the OSS project. The inclusion of information about the origin of code is also essential.

Chapter 4 presents ST processes; several processes can be used in case of OSS use and customization in companies. In Chapter 4.1 (Transition Planning and Support) a release policy is detailed (4.1.4.2), recommending specific IDs for any release. It uses Major Release, Minor Release and Emergency Release as typical examples for release types. This concept needs to be extended for OSS use and customization to include a differentiation between internal and external release.

Several relevant topics are contained within Chapter 4.2 (Change Management). Considerations regarding the design and planning of the change management in 4.2.4.2 include the proposal of both a change advisory and emergency change advisory board as well as competences and responsibilities of stakeholders (e.g., suppliers). With suppliers as stakeholders within the change management, the OSS project may be defined as a stakeholder; as a result, code or other feedback to the OSS project can be covered via this stakeholder relationship. In addition, 4.2.6.7 (Review and Closure of Change Records) explicitly mentions feedback to the stakeholders and 4.2.7.3 includes Sourcing and Partnerships within the change management interfaces. Therefore, ST implicitly includes the option to give feedback to an OSS project. The kind of feedback is not specified, this may include revealing experience (e.g., in form of participation in OSS support forums or submission of bug reports) or actual code submissions to the OSS project. In Chapter 5 (ST Common Operation Activities), the introduction of a Stakeholder Management Strategy with respect to suppliers of services and products is mentioned in 5.3.

Chapter 4 also includes definitions of changes, with standard changes defined in 4.2.4.5. A new official OSS release shares several criteria with a standard change, especially a defined trigger. Without a standard process for merging of the new official version within the external code branch with the internal code branch, this particular change will not share the criterion low risk with the standard change and, therefore, remain a special kind of change every time. ST does not provide actual implementation practices for changes; therefore, a detailed change process for OSS needs to be specified within the company.

ST defines typical roles in Chapter 6. Service Asset and Configuration and Change Management are combined as one responsibility; six roles are defined as typical: Service Asset Manager, Configuration Manager, Configuration Analyst, Configuration-Administrator/Officer, Change Management System/Tool-Administrator and Change Manager. Additional tasks and responsibilities can be assigned to these roles for internal use and customization of OSS without need for additional roles.

The configuration management system is specified in great detail in chapter 4, where examples for configuration splitting are given. ST does not, however, specify best practices for the handling of separate code branches in which the same OSS version is concurrently modified by different development teams. Like in the V-Modell XT, the different code branches can be handled as different baselines, or the official code could be defined as a specific CI. The required support for code merging in OSS customization projects is not included within the ST.

In summary, the adaptation process can in part be structured on recommendations of the ITIL framework, but ITIL does not provide advice for the case of concurrent development of the same base version of SW in different locations (i.e., within the OSS project and the company). Copies of the same CI are only mentioned in Chapter 7.3 (technological considerations of the configuration management system), but it is unclear whether this includes separate copies of the same CI.

4.3 ISO/IEC 15504 (SPICE)

The standard ISO/IEC 15504 (SPICE) is an example of a quality management and assessment standard. It incorporates ISO 12207 AMD1 and AMD2, a process reference model for SW life cycle processes. The assessment model consists of 48 processes distributed over three process categories with one or several process groups within these categories (ISO/IEC 15504-5 2006, pp. 3).
SPICE includes three process categories Primary Life Cycle, Organizational Life Cycle, and Supporting Life Cycle. The last process category covers configuration management and change management which are of special interest here. Individual processes within Supporting Life Cycle Processes are grouped in Support Process Groups (SUP). SUPs can be utilized by processes within the Primary Life Cycle and Organizational Life Cycle. For maintenance, SUP.8 (configuration management) and SUP.10 (change request management) are utilized by ENG.12 (software and system maintenance) within the Engineering Process Group.

Ten base practices (BPs) for configuration management are defined for SUP.8. Several BPs cover topics also found within ITIL ST, e.g., configuration management elements (BP 2), baselines (BP 4) and the archival of changes of every configuration management element (BP 7, BP 10). However, BP 3 and BP 8 cover areas relevant to OSS use and integration not found within ITIL ST. BP 3 relates to the development of a branch management strategy. As part of the development of a configuration management, it covers the handling of more than one code branch resulting in developers working on separate copies of the same configuration element. For this case, special quality requirements are in place. Reviews are mentioned as an example of a possible method; the branch management can be supported by configuration management tools. BP 8 covers reports on the configuration status while BP 9 covers the verification of configuration elements. These BPs are especially relevant in case of integration activities and if parallel development or variants are used (Hörmann et al. 2006, p. 198-199). The demand for a verification of the correctness of a baseline emphasizes the particular relevancy of verification in this case.

The actual changes are reflected in SUP 10 (change management). Change management as a support process includes here BPs that encompass the change management process from the development of a change strategy until the review of the actual change. Several BPs focus on the handling of change requests, which are also defined as one of the work products (ISO/IEC 15504-5, p. 55). Within the BPs, there is no explicit reference to external change triggers. Also, none of the BPs includes the explicit possibility of assessment and inclusion of external code.

4.4 Comparison of frameworks

V-Modell XT provides adequate practices for the assessment and adoption of external components but it is unable to represent the intertwined development processes that occur in case of customization of OSS in companies as a whole. Feedback processes are focussed on the traditional understanding of a buyer-supplier relationship.

ITIL covers the areas of OSS customization that are consistent with the customization of closed-source SW, but it does not provide appropriate best practices for the management and merging of code branches resulting from concurrent development in case of OSS customization. Some practices exist as basic approaches but they need to be modified or expanded to match the requirements of OSS customization, e.g., with the introduction of the OSS project as a new stakeholder.
SPICE is less comprehensive in the areas of change management and configuration management; nonetheless, BP 3 within the configuration management support process recommends developing a strategy for branch management, which is essential for OSS customization. SPICE proposes BPs for the relevant process areas but it does not include examples of typical cases as ITIL does.

5 SUMMARY

The possibility to modify the source code in order to custom tailor it to the needs of the organisation is one of the most unique features of OSS. Through action research of the use and customization of an application OSS over several official releases and internal modifications, we identified differences between the customization process of closed-source SW and OSS. These differences were mainly caused by the concurrent development in the OSS project and within the company and by the possibility of use of externally available non-official code.

Our literature review of selected frameworks on good practices in SW development and use revealed that the process of customization of OSS is not sufficiently covered by these frameworks.

Some tasks that were performed within the OSS CMS customization we studied could not be found in any of the frameworks considered, e.g., the differentiation of code according to its origin (internal code, external non-official code, or official release code).

The overall economic benefit of OSS cannot be disputed (e.g., MERIT 2006). If more companies outside of the SW industry would take advantage of its feature of open source, not just of low costs, the benefits for the whole economy would further rise. Companies may be more willing to choose this option if they get adequate support in managing the process of OSS customization. Such support can come from best practice frameworks that are increasingly being used if the frameworks get extended by practices needed in OSS customization.

While our research of OSS customization is limited by the analysis of a single project, it is sufficient to reveal difficulties if ITIL, SPICE, or V-Modell XT are used as they are to manage OSS customization in companies. According to our results, users who want to customize OSS inside a company cannot rely only on best practices included in the observed frameworks. These frameworks particularly do not cover the concurrent and intertwined process of customization which is characteristic for internal customizations of OSS. Extensions of the frameworks have to be made to cover the complete process. We have indicated in the previous section some directions for extension.

Like in SW testing where a test can only detect errors but not prove the correctness of a program, the experience with one OSS customization can just indicate some problems of the frameworks but not necessarily all of them. Therefore, more research on OSS customization in companies is needed to establish best practices in this area.

References


INFORMATION INFRASTRUCTURE GOVERNANCE AND WINDOWS OF OPPORTUNITY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0181.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>IT governance, IT Strategy, IT/IS management, Health information systems / medical record systems / care delivery /</td>
</tr>
</tbody>
</table>
INFORMATION INFRASTRUCTURE GOVERNANCE AND
WINDOWS OF OPPORTUNITY

Sun, Violeta, University of São Paulo, FEA/USP – Department of Business Administration,
Av. Prof. Luciano Gualberto 908 sala G-122, Cidade Universitária, CEP 05508-900, São
Paulo, SP, Brasil, violeta@usp.br

Aanestad, Margunn, University of Oslo, Department of Informatics, PO Box 1080, Blindern,
Oslo, Norway, margunn@ifi.uio.no

Skorve, Espen, University of Oslo, Department of Informatics, PO Box 1080, Blindern, Oslo,
Norway, espesko@ifi.uio.no

Miscione, Gianluca, Department of Urban and Regional Planning and Geo-Information
Management, International Institute for Geo-Information Science and Earth Observation,
Hengelosestraat 99, P.O. Box 6, 7514 AE Enschede, The Netherlands, miscione@itc.nl

Abstract

In this paper we discuss the challenges of managing large-scale information infrastructures. Various
management models, such as the IT governance model, propose structured approaches for
management of an organization’s infrastructure. This paper argues both theoretically and empirically
that such an approach to information infrastructure governance has its limitations. The paper is based
on empirical material from three change processes in information infrastructures in the context of
health care. We present case vignettes that illustrate how these processes evolved along unexpected
trajectories, subject to factors beyond the control of management. To conceptualize these phenomena
we draw on literature that conceptualize change as emerging from the meeting between multiple
parallel streams of activities. The interactions between the various streams open up windows of
opportunities that affect the information infrastructure development. We argue that such
conceptualizations more realistically depict how large-scale information infrastructures evolve, and
hence how they can be (or not be) managed.

Keywords: IT Governance, IT Strategy, Garbage Can Model, Health Information Systems,
Information Infrastructure

1 INTRODUCTION

The topic of this paper is management of organizational information infrastructures. Often these
information infrastructures are considered to be relatively stable and thus suitable for strategic
planning (Weill and Broadbent, 1998). Our case study indicates that the assumptions of stability and
manageability are not always in accordance with empirical reality, and consequently that the
prescriptive models for strategic planning (such as IT governance models generally) may be somewhat
optimistic.

The field of IT governance has seen a growing interest over the last 20 years. With the fall of Enron in
2000 and the subsequent introduction of regulation that requires auditing and control such as the
Sarbanes-Oxley Act in the USA and Basel-II in Europe, the field of IT governance experienced a
significant boost, and is currently widely touted and deployed. Substantial research findings on IT
governance became available in the early nineties (Loh and Venkatraman, 1992; Henderson and
Venkatraman, 1993) and became prominent in late nineties especially with the works of Brown (1997)
and Sambamurthy and Zmud (1999). In particular the research of Peter Weill and Marianne Broadbent were important in presenting the topic of information infrastructure management for a broader audience (Weill and Broadbent, 1998). They argue that in order to maximize business value from its investments, a company should conceive also its ICT infrastructure as an investment portfolio and subject it to management, balancing risks and benefits like for any other investment portfolio. Weill and Ross (2004) follow up by advocating business executives to not leave the IT decisions to the IT departments, but to “professionalize” IT governance. They suggest a framework for allocating decision rights and accountabilities in the use of IT. Decision responsibilities/rights are allocated to different groups and mechanisms are defined that formalize the relationships and provide rules and operating procedures to ensure that objectives are met. Six metaphors drawing on political archetypes are used to describe the mix of decision rights: Business monarchy, IT monarchy, feudal, federal, IT duopoly and anarchy.

In their book, Weill and Ross also propose a distinction between five different types of IT decisions related to: IT principles, IT Architecture, IT Infrastructure Strategies, Business Application Needs, and IT investment and prioritization. We want to draw the attention to the postulated relations between them. In order to ensure the realization of business value from IT investments, this model for IT governance stipulates an optimal sequence of actions: Initially IT principles should be formulated to comply with the strategic aims. These principles then comprise the basis for designing the architecture of the organizations information infrastructure. This architecture is then again the basis for investments relating to the information infrastructure. Then, the infrastructure capability enables the specific applications to be built based on business needs (to be specified by the business process owners). Finally, the resulting IT investments (in the investment and prioritization process) are expected to contribute to the strategic goals since they have been derived from the preceding definition of the IT principles, architecture, infrastructure and application needs.

In the next section we present alternative studies of information infrastructures. A significant theme in this strand of research is how IT infrastructures often evolve as a result of contingencies and events that were not planned, thus the actual evolution of information infrastructures often exhibit an unpredictable character. This might affect the possibility for following the structured and sequenced approach suggested by IT governance models, and it may seem that the assumptions about the stability and manageability of its object (organizational information infrastructures) are perhaps too simplistic. In the paper we investigate this topic in more depth, and our research aim is to examine the role of unpredictability versus strategic plans in processes of changing the information infrastructure. Thus, in addition to the studies of information infrastructures, we also review literature on the role of emergence in e.g. processes of strategic planning. We briefly present the well-known Garbage Can Model from organizational theory (Cohen, March and Olsen, 1972), as well as adaptations of it within political science that employ the notions of multiple streams, windows of opportunity and punctuated equilibrium. These notions have also been employed within IS, notably by Orlikowski, and constitute the core elements of our analytic framework of this study.

Following the theory section, our methodology section details how we have combined three independent empirical studies from very different organizations (however, all within health care and all describing infrastructural change processes). The case presentation is organised as three case vignettes where we focus on the process of evolution of the information infrastructure in question. The vignettes are crafted to illustrate the different relevant “streams of activities” that go into information infrastructure evolution. The case vignettes describe the different sources and temporalities of the streams, as well as the consequences when these multiple streams intersect, and the “windows of opportunities” that are opened up. This is followed by the discussion where we indicate the relevance of a perspective on IT governance that tries to balance the influence of structured models for IT governance with the influence of contingencies that tend to be beyond the control of IT management.

Proceedings ECIS 2009
2 LITERATURE REVIEW

2.1 Information infrastructures and strategic planning

Some conceptualizations of information infrastructures emphasize not the manageability, but rather the complexity of such large-scale interconnected assemblages of information systems (Contini and Lanzara, 2008). Within this strand of research, an information infrastructure is seen as a heterogeneous social-technical network, not just a technical communication network. The physical connections and equipment, the technical standards, the conventions of use, the technical and organizational support structures, the organization of work and cooperation are parts of the infrastructure (Hanseth and Monteiro, 1997; Star and Ruhleder, 1996). Such complex networks emerge only over time, not instantly. Consequently their existence is a historical result of long-term processes more than a result of single decisions and consequent plans.

These studies also emphasize how the growth of information infrastructures is shaped by the “installed base” that is already in place. Since the existing information infrastructure provides both resources and constraints for further development, radical and abrupt changes in information infrastructures are rare. Any change or intervention needs to take into account the installed base. These information infrastructures are large and open, heterogeneous, socio-technical networks, where various actors have different types and degrees of influence over the complete information infrastructure. This implies that there may not be a single locus of control over an information infrastructure. This distributed nature of mandate, ownership and agency makes conventional, control-oriented management approaches of limited applicability (Ciborra, 2000). Thus, the notion of cultivation has been suggested as an alternative way to conceptualize management or governance of such processes (Hanseth and Lyytinen, 2004). The element of emergence is central in this conceptualization.

We see a parallel in how emergence is perceived in organization studies. For instance Mintzberg (1989) describe how strategies were seldom implemented in the way they were conceived but rather, were emergent. Mintzberg chose “crafting” as a metaphor when he proposes the following postulates about strategies:

- Strategies need not be deliberate – they can also emerge, more or less unplanned.
- Effective strategies develop in all kinds of strange ways.
- Strategic reorientation happens in brief, quantum leaps.
- To manage strategy, then, is to craft thought and action, control and learning, stability and change.

The concept of emerging strategies becomes especially interesting when dealing with uncertainty of the environment; with instability, complexity and change. Top-down strategizing might work under simple and stable conditions, but when conditions are uncertain, instable and complex, strategy making becomes a bottom-up process (Clegg, 2005). We argue that such socio-technical conditions are increasingly a part of organizational lives in a fast moving global economy, and that this is in particular the case for information infrastructures because they often extend beyond the immediate and local context. Thus the changing environment, such as e.g. technological development, or legislative changes, may significantly impact the organization’s internal information infrastructure’s development. However, not just external conditions challenge the straightforward implementation of strategic plans, also internally in an organization the decision-making processes leading to action have been found to be less structured than expected.

2.2 The Garbage Can Model (GCM): multiple streams and windows of opportunity

The Garbage Can Model of organizational choice, proposed by Cohen, March and Olsen (1972), describes decision making in situations that do not meet the conditions of more classical models of decision making, as preferences are seen as problematic, technology as unclear, and participation...
fluid. Here decisions are seen as the results of arbitrary streams of solutions, problems, participants and choices, temporarily coinciding in “choice opportunities”. These streams follow different temporal logics, and only randomly open up windows of opportunities for decision making. The time it takes, and the problem it solves, all depend on a relatively complicated intermeshing of elements. These include the mix of choices available at any one time, the mix of problems that have access to the organization, the mix of solutions looking for problems, and the outside demands on the decision makers.

In the GCM, a decision is an outcome of several partly independent streams within an organization. Cohen et al (1972) outline four main streams as follows:

1. Problems – The concern of people inside and outside the organization, that might arise over issues of lifestyle; family; frustrations of work; careers; group relations within the organization; distribution of status, jobs, and money; ideology.
2. Choice Opportunities – Occasions when an organization is expected to produce behavior that can be called a decision. Opportunities arise regularly and any organization has ways of declaring an occasion for choice.
3. Solutions. Somebody’s product, which is looking for problems to be solved.
4. Participants. Participants come and go. The distribution of “entrances” depends on the attributes of the choice being left as much as it does on the attributes of the new choice.

Although not completely independent of each other, each of the streams can be viewed as independent happenings in a timeframe. Elements of organizational structure influence outcomes of a garbage can decision process by (a) affecting the time pattern of the arrival of problems choices, solutions or decision makers, (b) by determining the allocation of energy by potential participants in the decision and (c) by establishing linkages among the various streams. Organizational structure changes as a response to such factors as market demand for personnel, and the heterogeneity of values which are external to the proposed model.

According to the GCM, although decision-making should be a process for solving problems that is often not what happens. Problems are worked in the context of some choice to be made, but choices tend to be made when the shifting combination of problems, solutions and decision makers happen to make action possible. This aspect of the theory has been taken up also in fields that aim to explain processes beyond the organization, such as policy making. The formulation of the “multiple streams model” by John Kingdon (1984) has been the most well known example. Kingdon proposes that there arise possibilities for policy change when the complimentary streams of problems, policy and politics converge. In themselves they are not effective, but when either two of them or all three meet, this create openings that Kingdon calls “policy windows” where actual policy change can be effected. This point is again reflected in the punctuated equilibrium model for policy change (Baumgartner and Jones, 1993) which explains stability and change in political systems. This model claims that radical change can happen to seemingly stable systems in cases when windows of opportunity allow the opposition to challenge and overturn the established “policy monopolies” (i.e. institutionalised structures).

The most renowned contributions that apply these notions within Information Systems Research are probably those of Wanda Orlikowski. The concept of “window of opportunity” was employed by Tyre and Orlikowski (1994) who drew the attention to the adaptation period following installation of an IS. Their case study showed that there was a brief window of opportunity in which organizational innovation (exploration and modification of the IS) could be pursued. After some time had gone by, the practices had gotten more routinized, and the window for experimentation and change was closed. Later changes could occur, but in an episodic manner and not continuous and ongoing. In a later paper, Orlikowski (1996) discusses situated change as a metaphor that can capture the emergent nature of socio-technical change processes. She finds well known conceptualizations of change problematic because of their assumptions of stability; this includes notions such as planned change, the technological imperative model and the punctuated equilibrium model. While we may not agree with
her on the primacy of change rather than stability in organizational life, we think that her conceptualisation of change has much to offer. She sees organizational change as resulting from “ongoing improvisation enacted by organizational actors trying to make sense of and act coherently in the world.” (Orlikowski, 1996). This acting happens under the conditions of, among other things, temporal structures (Orlikowski and Yates, 1999). The organizational members align their activities to the stabilized-for-now temporal structures, but they also enact them and can thus possibly change them.

Summing up the theoretical framework, we choose to draw on the notion of multiple streams, which we find particularly relevant for information infrastructures, since they have a heterogeneous make-up, drawing together multiple technologies, usages and interests across a potentially very diverse organization. In addition we find the notion of “windows of opportunity” to be relevant in order to identify events and instances where these streams intersect and allow interventions on the information infrastructures. As information infrastructures are large and complex, they require a certain time period to change. A too brief window of opportunity or one that is inappropriately timed vis a vis the infrastructural evolution processes may not perhaps have the same effect as a longer or well aligned window of opportunity.

3 METHODOLOGY RESEARCH APPROACH

Our empirical material is based on the combination of three independent case studies carried out in a Brazilian hospital, in Norwegian a hospital and during an implementation project of a state-level health information system in Gujarat, India. Thus the material has not been produced on the basis of a common research design, but was combined a posteriori as a result of long-term discussions among the authors on the theme of information infrastructure evolution. These three cases span a wide variety of geographical contexts, organizational scales and infrastructural conditions, and the vignettes we present here have their focus on the temporal evolution of the change processes.

The research approach in all three case studies has been qualitative, and the material presented was collected via interviews, document analysis and on site observation. Our research approach has been interpretative in orientation, aiming both at eliciting our respondents’ perceptions and opinions (Klein and Myers, 1999) and understanding the actual doings in which actors are involved. Following we give additional details on each of the case studies.

The Brazilian hospital is a São Paulo State government agency subordinated to the São Paulo State Health Secretariat. It has over 15000 employees, and has on average 2250000 patient encounters each year. The material from this hospital, collected by one of the authors, is based on interviews conducted in 2007 and 2008 with the coordinator of the IT department at the time of the research, a previous IT coordinator of the same department, one medical doctor and interviews with managers at the State Health Secretary with follow-up information exchange via email. Analysis of documents provided by the IT department has also been important. These documents include strategic plans, as well as hardware and software descriptions, and annual IT reports. Also, documents available in the website of the hospital as well as publications related to the IT department were used.

The Norwegian hospital is a highly specialized university hospital with national responsibilities within complex treatments. It has almost 8000 employees, and treated more than 300000 patients in 2006. The material from this hospital has emerged from a long-term collaboration between the Department of Informatics at the University of Oslo (where two of the authors are employed) and the IT department of the hospital. Dating back to 1996, various student projects followed the activities around the development and implementation of an Electronic Patient Record (EPR) system. In the period between 2001 and 2004 the cooperation was organized as a more structured research program, where interviews and on-site work observations were conducted with representatives from all relevant personnel groups in three different departments as well as with key actors in the IT department and hospital management.

Proceedings ECIS 2009
One of the authors did fieldwork in the state of Gujarat in western India. Gujarat is one of India’s most industrialized states, and has a population of around 50 million. The researcher was taking part in a project aimed at implementing a health care information system for aggregated health data (i.e. not clinical, patient-based data, but statistical information for administrative purposes). The data collection was based on qualitative methods and interpretative approach within an action research framework. Also quantitative data were used as empirical material as far as they affected the field of study. Participant observation was complemented by individual and group level semi-structured interviews, and also by observation of a focus group. Shadowing was another method used to follow health personnel, officers, as well as software developers. An important component of the participant observation concerns the online activity of an INGO (Indian Non Governmental Organization) such as email, mailing list, chat lines, and phone calls; which was very important for understanding the activity and coordination. In addition a number of government records and reports were also analyzed to further gain an understanding both of the Gujarat health organization and its ongoing processes of intersection with the INGO.

4 VIGNETTES

The following vignettes illustrate how strategic planning relates to actual information infrastructure evolution. Our focus is on examining to what degree the planned and managed action are influenced by activity streams that are either central or peripheral to, or even not related to the initial strategic plan. The first vignette (from Sao Paulo, Brazil) illustrates that political actors can play crucial role in shaping public health care organizations’ internal information infrastructures. The second vignette (from Oslo, Norway) illustrates how a hospital relocation process opened up a window of opportunity for a redesign of the information infrastructure, but where the actual realization deviated somehow from the plan. The third vignette (from Gujarat, India) illustrates the way that multiple streams of activities related to the development and implementation of a state-wide information system constituted opposing and non-synchronized demands.

Vignette 1. Redesign of Infrastructure (Sao Paulo, Brazil)

In Sao Paulo, Brazil, the period 1995-2006 was characterized by administrative continuity (i.e. a succession of governors from the same political party) and a continuous strong emphasis on the use of IT. IT became a strategic issue due to a large state modernization program started in 1995. Budgeted resources, complemented by international funds for modernization of the financial system, education, health, environment, police and social services, were used to build a basic communication infrastructure, integrate databases and to create multi service government offices that were perceived by citizens as showcases for a modern government (Reinhard et al., 2006).

It was in this political environment that, in the beginning of 1995, the Board of Directors of a Brazilian public hospital decided to renew the hospital information infrastructure, with integration of corporate information as the main goal. A new Superintendent had recently been appointed and plans for the following years had been planned and presented. The hospital, which was a teaching hospital, comprised 6 Institutes, 2 Auxiliary Hospitals, a Rehabilitation Division and one associated hospital in a campus that housed around 3,000 operating beds. The institutes were distributed in a campus of around 352,000 m² area.

The information systems in the hospital consisted, at that time (1995), of two mainframe computers, local area networks of various sizes and technologies, and hundreds of stand-alone PCs. The two mainframe-based systems were unable to exchange information, both between themselves and with the departmental networks. IT services were provided by a government partner organization and a few service providers. Most IT related decisions were depending on the support of the partner organization.
It was understood that a renewal of the IT infrastructure would require the establishment of a common ground of understanding of the involved effort within all hospital units and institutes. It was important to define what information and which processes should be shared and integrated across internal organizational boundaries. So, an Informatics Committee (IC) was created, formed by representatives of every Institute of the hospital, also, partner organizations and some service providers were also invited to become IC members. The IC established the main guidelines for an IT renewal in the hospital. Mainly due to growing dissatisfaction amongst users with the existing systems, the group decided to build a brand new HIS for the hospital from ‘scratch’, rather than upgrading the existing legacy systems (Moura et al., 1998). The easiest way to expand the installed base would have been a continuity of investments in mainframe technologies, as most of the existing applications were based on this platform. Nevertheless, the new platform was built on a very different conceptualization of information infrastructures. This was not based on the main proposal of the previous IT decision makers from the partner organization, so tensions emerged between the new decision makers and the partner organization. There was a strong support from higher administration all through the process as building the information system was considered a priority of the leadership.

This situation opened a window of opportunity where favourable political will stimulated investments in the hospital’s new information infrastructure, despite the previous IT decision makers’ opposition to the new proposal. In this case the political temporality emerged as a strong factor due to the continuity in the political arena that contributed to the implementation of the renewal proposal in the following mandates. In areas characterized by political instability, projects are often terminated without any explicit reasons. As a consequence of shifts in political parties, new projects tend to be started in order to create references for the current political regime, so that credits are not attributed to previous politicians and the regimes that initiated them. Historically this has frequently been the situation in countries, where political instability was the rule, rather than the exception, for many years. In this case, the window of opportunity was not only opened for new investments but also remained open in the following political mandate. A totally different scene would have been likely within a more instable political frame where this project easily could have been terminated without prior warning.

**Vignette 2. LAN INFRASTRUCTURE (Oslo, Norway)**

During the 1980’s and 90’s the IT infrastructure of a large Norwegian hospital had been evolving in an arbitrary fashion. Different units of the hospital were located in separate buildings, with variable access to network infrastructures. Since the early 1990’s a few mainframe systems had served several administrative purposes, such as a Patient Administrative System, a Personnel System and an accounting system. In 1995 a number of projects aiming at the implementation of central clinical information systems were initiated (an electronic patient record system, a radiological information system and picture archive, as well as four laboratory systems). In addition, entrepreneurial doctors in various specialties had purchased or developed specialized applications for their own treatment and research needs. For instance, the heart surgeons had a database application where extensive information on all surgical interventions on heart patients for a couple of decades had been recorded. This information was invaluable for their research, and could not be recorded in or retrieved from standard EPR solutions. As a result of this and similar systems, the ‘corporate’ application portfolio was increasingly problematic to support and the opportunities for integration of the various systems were also limited due to a lack of standardization.

The whole physical network infrastructure was redesigned from scratch before moving into new facilities in 2000. When the bid for tender was announced in 1997, the new network infrastructure was stipulated to serve more than 4,000 PCs, printers and peripheral devices. In addition centralized management of both hardware and software was to be implemented, including the monitoring of the entire network and all connected devices. Also the application portfolio was somewhat reduced during this move, however most of the clinical specialist systems were maintained under a portal architecture that simplified the users’ interaction with the systems. In general the user’s autonomy with respect to
IT was curtailed, as they could no longer install what they desired on their computers or access Internet freely. The protests against this restriction were significant, in particular from researchers using specialized software. The IT department accommodated the users by offering access through another, separate network for the hospital’s university employees (this was a teaching hospital). This alternative network had a more liberal policy than the hospital’s internal network.

The new facilities was expected to be ready in October 1998, but due to various delays during the building process the actual move was postponed three times, until it finally occurred in May 2000. This meant that initially the IT department had expected to be able to leave behind a lot of the issues associated with the Y2K risk, as old servers and most of the old applications were not supposed to be moved along to the new site. When the move was postponed beyond this point, a risk assessment project had to be conducted and necessary measures taken.

In this example, the redesign of the information infrastructure was coupled with the relocation process which opened up a window of opportunity for radical changes. However, this coupling also backfired, when the relocation was postponed, and the IT department had to devise an ad-hoc strategy to cater for the Y2K issues. Moreover, also the users’ protests’ were taken into account and consequently led to non-planned and more improvised adjustments of the initial strategy.

Vignette 3. Reporting System for aggregated health data (Gujarat, India)

An Indian NGO (hereafter INGO), which was part of an international network dedicated to HIS, was responsible for a local tailoring and implementation in Gujarat, with the support of the local health department. The information system was based on a free and open source framework, and the implementation we report from started in February 2006 and went on until April 2007.

The health information system in question was a system that supported capture and reporting of aggregated data for epidemiological and managerial purposes. The healthcare context was bureaucratic and procedural reporting was the main function of information flow. While the ideology of the INGO was oriented towards supporting action primarily at the primary healthcare level, the implementation followed the existing information flow, more vertically channelled, because it was thought to be the best strategy to get acceptance. The health reporting system thus aimed at replicating the existing organizational structure, where each level aggregates data from the lower one, adds new data and produces its own reports as per predefined national formats. Technological and organizational change would be introduced incrementally. The preferred strategy of the INGO was to start pilots at the district level, customize the system, and then scale up.

The state of Gujarat is administratively divided into 25 districts. Within each district there are Block Health Offices (BHO) which coordinates administrative tasks at the sub-district level, and between the levels of the district and community. In the INGO plan, the computers were placed in the Block Health Offices, as all the Primary and Community Health Centres’ personnel would come to these offices for their information processing activities (including data entry and report generation).

The implementation started with a pilot in the southern part of the state. One computer was installed in each of the five Block Health Office (BHO) that together covered a whole district. In July 2006, a positive evaluation was conducted by the state authorities and a prestigious national management institution.

Indeed, in August - September 2006, southern and central areas of Gujarat had been hit by an unusually violent monsoon, which affected a huge number of victims and heavy material damages occurred in such areas. The positive evaluation allowed the “scaling up” of the HIS to four more districts, for a total of more than twenty BHOs. To avoid any extra burden on health personnel in the affected areas, the IT project was pushed to districts where no emergencies had to be prioritized. Those districts were located in the north-western part of the state, where the heavy monsoon rains and floods had been less intense. Thus all these new districts were far from the first pilot site, where IT
capacity had been developed, and far from the state capital, where decisions were taken, and some office space was made available to INGO. Geographical distances, poor transport infrastructure, and weak coordination made implementation in these new districts relatively slow, as five people had to cover an area of several hundred kilometres (many hours by train). The slow implementation was also not synchronized with quick development of negotiations between the INGO and the commissionariat at top level of state administrators, who were also moved by other pressing concerns for the aid for flooded areas.

Beside contingent problems, sharing of information across vertical health programs such as HIV/AIDS, malaria, tuberculosis (TB) was very poor. During an interview, one of the vertical program’s managers declared that “we [the health program managers] have the population in common, only.” The mismatch between the actual flow of information and State Health Commissioner’s need of gaining an overview of the health organization functioning through information was expressed emphatically by him, who declared “I don’t want a system that merely automates the existing inefficiencies”, adding “what we need is a 2-3 pages report, to take action on”. The starting point in that direction was the definition of indicators for the Reproductive and Child Health National Project. This project, indeed, requires pulling data from different health programs (pre-natal care, vaccinations, HIV tests, nutrition, etc). So, a negotiation to define indicators’ formulas and data access started from the top organizational level. Implementation at lower administrative levels, beginning from districts, was discussed, as well. Due to these different ‘streams of events’ (districts rollout of HIS, indicators definitions), different timeframes were manifesting: while computers still were to be installed in some of the five districts, the commissioner started to push for a “dashboard” of indicators to monitor the health situation. By the term “dashboard”, he referred to a visual representation of the state health situation through graphs and maps; the aim was to have health conditions at a glance. From the health department’s side, this was thought to be useful also for the ongoing disaster recovery, but the timeframes (and the geographical areas involved) were mismatching at that point of time.

5 ANALYSIS

Despite their diversity the case vignettes all show that sometimes information infrastructures evolve in ways not directly resulting from strategic planning. Sometimes strategic decisions can be followed through to implementation, but at other times the evolution of information infrastructures does not seem to directly result from the decisions taken to govern them. In the latter case the II evolution is impacted by from circumstances that are neither controllable nor predictable because they are related to issues beyond management’s sphere of influence. Different streams of activities, often beyond the control of the planners, blend in the formation of emergent strategies. These emergent strategies are related to exploiting situations where the confluence between multiple streams of activities constituted ‘windows of opportunities’. This required adaptive and ad-hoc management action for noting and exploiting windows of opportunity.

In the Oslo case, the relocation process - an event beyond the control of the IT department - opened a large window of opportunity for the IT department to go through with a major restructuring of the hospital information infrastructure. Under any other circumstance, such a radical change would probably have been unfeasible. However, the strategy intended to achieve this had to be amended along the way, due to other contingencies: the clinicians’ exercise of their professional autonomy made it impossible to stick to the intended strategy of a single, centrally controlled network. The department’s effort to harmonize the infrastructure was a solution to a long-lasting messy situation, waiting for an opportunity to make the decision to enforce it. The relocation of the hospital provided such an opportunity. But the clinicians’ significance as central actors in the process proved difficult to ignore, and subsequently a second network was implemented. This was not predicted as a problem or as a solution in the original plans, but was still handled along the way in a manner that became acceptable to all parties. The Y2K problem illustrates the significance of the temporalities when we
deal with such windows of opportunities for infrastructural change. Due to the delays in the relocation process it came to have significant impact on the process.

The São Paulo case presented radical changes in the information infrastructure. As for the partner organization that previously was responsible of the IT infrastructure decisions, the influence in decision-making related to the new infrastructure was reduced. Political will and continuity opened up a window of opportunity and contributed to the implementation and continuity of an infrastructure that was completely different from the existing one. Before this change, it would have been very difficult to predict what the new infrastructure would be. Political stability allowed for continuity of investments in the new infrastructure, but even though the infrastructure developed in a fairly “smooth” way the following years, it is hard to know if this infrastructure will have the same support from the top IT decision makers after forthcoming political elections.

In the Indian case multiple distinct streams of activities intersected. The practical activity on the ground, comprising installation of hardware and software as well as capacity building in pilot health facilities, was delayed due to heavy monsoon rains, geographical distances between the pilots and poor transport infrastructure, which weakened coordination. These activities were also influenced by the request for the development and prototyping of a “dashboard” for health conditions monitoring, due to State Health Commissioner needs. The dashboard implied the definition of health indicators’ formulas and access to data sources (from different health programs) and actual provision of data. Each of these flows had its own temporality. The partial match between them and HIS main stakeholders opened for a limited window of opportunity for system implementation.

By employing the notion of “multiple streams” to empirical studies of information infrastructures in health care contexts, we have emphasized the role of multiple streams of activities as well as how windows of opportunities emerged when these streams converge. These streams intersect in multiple ways, not all of which are expected. These intersections or interdependencies can have quite dramatic impact on assumedly unrelated activities. We argue that models that suggest structured approaches to information infrastructure governance (such as IT governance models) do not emphasize these aspects adequately. We have shown in the case studies that the information infrastructure’s development was influenced by activity streams emerging outside of the organization. Sometimes these streams interact to further management’s plans, such as when the hospital relocation served as a golden opportunity to redesign the LAN infrastructure or when the political stability and investments in IT in Sao Paulo enabled the redesign from scratch of the hospital’s information infrastructure. Sometimes, however, the intersection or interdependencies between the different streams generate unforeseen problems.

The implications of this view is that models that advocate structured approaches to governance need to be complemented with learning-based or adaptive strategies for decision-making and operational action. For instance Hayes and McGee (1998) applied the Garbage Can Model in a study of IS in organizations and acknowledging the constructive power of these processes, they concluded by recommending us to “Look out for and use those garbage cans!” (ibid p.33). Similarly, Fardal and Sørnes (2008) claim that Garbage Can processes are productive at organizational level, and they conclude that:

“[…] acknowledging that decision-making isn’t always a rational process could improve the fit between intended and actual decision-making structures. At last, at the organizational level, GCM decision-making provides sensible IS strategic decisions, meaning that CIOs can focus on implementing the decisions rather than imposing a decision-making rationality that participants do not adapt to” (ibid p. 565).
6 CONCLUDING REMARKS

In this paper we have emphasized the role of multiple streams of activities as well as how windows of opportunities emerged when these streams converge. These intersections or interdependencies can have quite dramatic impact on assumedly unrelated activities. From this perspective, it is no wonder that organizations often experience that they have limited control over internal streams and limited influence on external ones. While the streams can not be managed, organizations can try to manage their outcomes. And they are more apt to be able to do so if the inherent power and importance of these phenomena in organizational decision making is acknowledged. Our major message in this paper is that dealing with the dynamics of information infrastructures requires a balance between planned and emergent strategies and thus continuous learning. Based on this we could formulate the following guidelines for organizational members involved with IT governance:

• Since the intersection of the streams may emerge unpredictably, an organization need to strengthen its ability to predict and detect the windows of opportunities opened. Generally this requires a good understanding of the challenge that information infrastructures pose, as well as a sensibility towards the character of organizational processes. The theoretical streams and concepts presented in this paper can be seen as contributions toward this aim.

• Since the intersections of such streams can represent unique opportunities for change, as exemplified in the empirical cases, the organization also needs to develop agility to exploit them. This requires agility both in the organization and in the IT governance structures:

• Any IT governance model in use should incorporate the required flexibility to leave aside the standard plan it has sketched. Ideally it should have an inbuilt robustness towards adaptive handling of change. For instance, the advice of the IT governance models to distinguish between different types of IT decisions may be sound in general. Sometimes, however, the various areas intersect in unpredictable ways. New application needs may challenge the previous decision of architectural aspects of the infrastructure, for instance. Sometimes external change, e.g. in legislation, may overturn the carefully designed strategic IT principles. How is then this handled by the governance structure in place? If the IT governance model does not incorporate the flexibility required for this kind of deviations, it needs to be overruled and set aside in periods of radical change occurring from the kind of phenomena we have described.

• Usually this kind of flexibility and agility needs organizational slack resources and/or redundancy. Be prepared that the organizational hierarchy may need to be softened up to enable the right participants to exploit the windows of opportunities that may emerge.

The baseline of our message is that the garbage can is there, and windows of opportunities can stay open or close without prior advices. It is thus up to us to choose if we will regard it as a disturbance to our ready made plans, or as a resource in order to reframe plans with unpredicted events.

References

CONFIGURATION OF ACTORS AND ROLES IN ESTABLISHING ICT

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0249.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>IT governance, IT innovation, IT-enabled social innovations, Theory Building</td>
</tr>
</tbody>
</table>
CONFIGURATION OF ACTORS AND ROLES IN ESTABLISHING ICT

Ariel La Paz, Universidad de Chile, Diagonal Paraguay 257, Santiago, Chile, lapaz@fen.uchile.cl

DongBack Seo, University of Groningen, Nettelbosje 2, 9747 AE Groningen, The Netherlands, D.Seo@rug.nl

Abstract

Establishing technologies has brought mixed socio-economic impacts across nations and regions. Researchers have studied the relationships between the establishment technologies and its impacts through identifying innovative processes, major actors, and available resources. However, the challenge to this literature is how less resourced countries have achieved greater prosperity than better resourced countries by establishing Information and Communication Technology (ICT). To understand and analyze this phenomenon, we propose a typology of the configuration of roles and actors in establishing ICT based on an innovation framework. The proposed typology can be used not only to explain different socio-economic impacts among countries or regions, but also to suggest a constructive way in establishing ICT through reconfiguring involved actors in the key roles.

Keywords: ICT establishment, Innovation, Triple Helix Model
1 INTRODUCTION

Technology has been identified as a means to increase economic growth (Fagerberg 1987) and potentially can originate virtuous circles that catalyze factors in the productive process and accelerate the generation and use of new knowledge (Garud & Kumaraswamy 2005, Bresnahan & Trajtenberg 1995). Consequently, the successful establishment of technology can improve standard of living and create positive socio-economic impacts. By establishment of technology we mean development, adoption or adaptation of technology as well as its implementation (Orlikowski 1992).

Certain types of technologies have been called General Purpose Technologies (GPT) due to their pervasiveness, technological dynamism, and innovative complementarities (Bresnahan & Trajtenberg 1995). Information and Communication Technologies (ICT) are one type of GPT that have been widely distributed throughout managerial, productive and communicative processes in the commercial industries, academic organizations, and governments.

ICT is principal to today’s knowledge-based economy and a backbone of commerce (Powell & Snellman 2004, Tripathi 2006). The abilities to manage, store, retrieve and transmit knowledge have positively affected societies and become catalyzing factors in productive and knowledge-intensive processes (Garud & Kumaraswamy 2005). Negative impacts in establishing ICT can be mentioned as well, like large investments without social net benefits or the severe imbalance of technology benefiting only a few parties instead of the whole community (Tyler 1962, Hanseth & Jacucci & Grisot & Aanestad 2006). Some nations or regions have enjoyed more positive socio-economic outcomes from the establishment of ICT, while others have struggled between positive and negative effects or suffered more negative consequences. Because of the central role of ICT and the potentially conflicting consequences for society, we are interested in learning about its establishment process and socio-economic impacts. To do this, we look for an alternative explanation in the configuration of roles and actors involved in the process of establishing ICT.

For the establishment of ICT, we can identify sets of actions, attributes and duties that need to be fulfilled, which we call ‘roles.’ Different actors can play various roles, forming configurations between roles and actors for the establishment of ICT. In establishing ICT within the socio-economic system, three major actors (government, universities and ICT industry) have been identified to perform different roles according to their special abilities and skills (Leydesdorff 2005, Williamson 2005). However, their performance can be defined by political forces or cultural mores instead of efficiency (Hamlett 1991).

The interesting question is why different regions have different socio-economic impacts, even when they develop or adopt the same or similar technology (Fagerberg 1987). According to Resource-Based View (RBV) theories (Barney 1991, Grant 1991), more endowed countries or regions have more resources to invest in developing or adopting technology than poor countries or regions, therefore, countries or regions with more resources can enjoy more positive socio-economic impacts than less resourced countries or regions; nevertheless, we have witnessed instances where one country with similar or fewer resources than others experience more positive socio-economic impacts from the establishment of ICT. For example, relatively less resourced South Korea has enjoyed mostly positive socio-economic consequences from adopting and implementing Code Division Multiple Access (CDMA) technology as its second generation wireless communication technology standard, while countries with more resources at the time like Japan, Taiwan, and Singapore have not had similar positive impacts.

To explore alternative theories to explain the variation in socio-economic impacts among regions, this paper rests on a conceptual framework to propose a map of configurations between actors and roles in establishing ICT. The map can be used to analyze the relationship between ICT establishment and its socio-economic impacts in a given region over time, or to compare different regions. Furthermore, the findings can provide pragmatic ideas about reconfiguring actors and roles in order to decrease the
differences and gaps in socio-economic impacts among countries and regions created by the establishment of ICT.

2 LITERATURE REVIEW

The establishment of innovations in ICT and its socio-economic impacts have been analyzed from a historical perspective (Dudley 1999); changes, interactions and consequences of technology establishment have also been studied (Feinson 2003, Lipsey et al. 1998, Niosi 2002). Dudley (1999) describes how the principal characteristics of ICT (storing, transmitting and reproducing information) have positively impacted economic growth in the period between 1000 Common Era (CE) and 1975 CE. He showed that the decentralization and distribution of information reduced the costs of disseminating information and increased literacy levels. As a result, this phenomenon led to the decentralization of economic activity, and more communities benefited from sharing or generating knowledge and accessing financial support and technological assistance.

Unintended negative effects of technology establishment have been reported as well. Tyler (1962) points out that the presence of technology in productive processes, products, and services affects people’s values, habits, and interests. As a result, we observe more powerful, expensive and destructive weapons, compulsive consumption, or decisions based on ‘efficiency,’ when the problem may not lie in the realms of economics, science or technology. Other negative effects have come as a result of automation, like the replacement of the workforce (Noble 1984), and the concentration of production has had its greater impacts on the less resourced and less skilled companies and workers (Flamm 1988). In general, social, economic and ecological problems have arisen as unintended side-effects during and after the establishment of various technologies (Kingdon 1995).

At the organizations level, mixed results have been reported as outcomes of ICT investments (Dehning & Richardson 2002), nevertheless it has been claimed that such paradoxes can be explained by measurement problems and outdated data (Barua & Kriebel & Mukhopadhyay 1995, Brynjolfsson & Hitt 1996). Seo & La Paz (2008) analyze the dark side of Information Systems, and recommend firms to be aware of the negative impacts of IS and balance both positive and negative side effects for organizational agility. Similarly, Carr (2003) recommends spending less on IT investment and thinking more to generate advantages by having superior understanding of the use of new technology.

Different types of actors play various roles in establishing a technology. For example, in the development of the semiconductor and computer industries in the U.S., the government was a principal actor in the early stages of invention and improvement for the first computers and telecommunications networks by fostering and funding research on defence (Flamm 1988, Ceruzzi 1998, Noble 1984, Denning 1997). History also confirms that the government has been not only the first promoter of innovations, but also was the principal customer of computer technology for statistical data analysis for a long period (Flamm 1988, Ceruzzi 1998). Academia was typically the recipient of research grants and carried out projects to develop and utilize new technology for improvements in productive processes (Noble 1984). Some successful academic projects transformed into venture companies by transferring innovative technologies into commercial products or services, becoming prosperous companies like SUN Microsystems (Ceruzzi 1988). Companies recognized the significance of research and development and kept investing in computer technology. For example, IBM invested strongly in R&D to enhance the processing and storing capacity of the first primitive computers, so IBM could rapidly commercialize and customize its products to meet market needs (Flamm 1988). In this case, IBM as a company in the IT industry played a fundamental role not only in research and development but also in establishing and commercializing the outcomes of applied research (Pugh 1995).

In order to discover the main categories that shape the configurations of roles and actors for the introduction and establishment of ICT, we surveyed the literature of innovation, economic growth, and
the actors of innovation. Our survey is focused on the identification of both roles and actors for the establishment of ICT. We acknowledge that other types of analysis may lead to explain the optimal configurations based on efficiency (see for example Williamson 2005), however, we attempt to set up a framework to describe the actual configurations that can potentially explain the variance in the establishment of ICT across countries or regions, and explain how different countries and regions have similar socio-economic impacts even though their ICT development paths are different (Doty & Glick 1994).

2.1 Innovation and economic growth

Society-wide economic growth is influenced by many factors such as culture, social events, economic, political, and technological forces (Spencer & Murtha & Lenway 2005). Focusing on technological and political factors in economic growth, Lipsey et al. (1998) propose a structuralist growth model. They criticize models of growth for being abstract and intangible in describing how a production function can transform input into economic output. The authors open up the black-box of production function and break it up into four elements affecting economic output: 1) Technology (stock of existing technological knowledge), 2) Facilitating Structure (physical and human capital, production facilities, managerial and financial organization, industrial concentration, financial institutions and financial instruments), 3) Policy Structure (means to achieve public policy such as public institutions), and 4) Policy (expressed in legislation, rules, regulations, procedures and precedents). The social facilitating structure transforms inputs into outputs affected by technology and policy structure. Policy structure is influenced by policy itself. Following this rationale, technological innovations would increase economic growth by adding input into the facilitating structure, and effects would be appreciated on the policy structure and legislation to increase economic performance.

Innovation plays a significant role in economic growth (Denning 1997, Mahdjoubi 1996). The Linear Model of Innovation describes four factors to generate innovation (Godin 2006) – 1) Basic Research conducted by the academia and funded by government or firms; 2) Applied Research, where firms use and test the basic research for practical usage; 3) Development, with firms further developing and customizing products and services for end users; and 4) Commercialization, where firms distribute products and services. According to this model, investment as input in basic research can generate more innovation as the output of the four factors, impacting economic growth (Godin 2006). As Mahdjoubi (1996) points out in his Holistic Innovation Model, innovation is not only the disruptive result of basic research, but it can be the product of continuously improving processes in designing or manufacturing, enhancing supply chain networks, emulating other technology, and acquiring new technology by purchasing licenses or patents. The Holistic Innovation Model specifies more functions in each stage and recognizes feedback between stages.

These models provide a basis in identifying necessary roles in establishing ICT. In the following section, we review the existing literature to categorize actor groups that can play roles in the ICT establishment.

2.2 Actors of innovation

forcing a transformation of each actor and their role (Han et al 2001, Leydesdorff & Meyer 2003, Shinn 2002).

The Structuralist Growth Model, Linear Model of Innovation, and Holistic Innovation Model emphasize various functions to establish ICT and show how different degrees of socio-economic effects can occur depending on those procedures (functions). On a different bracket, investigators concentrate on identifying actors without indicating their roles in conducting different procedures (functions) in establishing ICT (Cohen & Levinthal 1990, Etzkowitz & Leydersdorff 1995, Mahmood & Rufin 2005, Spencer et al. 2005). These two branches of literature focusing on roles or actors in establishing ICT and their socio-economic impacts open up intriguing gaps: (1) what are the major roles played by the various actors and (2) who plays which role(s) in establishing ICT. The latter question brings up the very significant perspective that the same role can vary in its outcome depending on the actor performing it, and such variation may have an impact on ICT establishment.

By reviewing various theories and models, we are able to recognize roles and actors. From the literature, most innovation theories focus on the first role of the academic and private sectors in detecting the needs and solutions of technology innovation. Meanwhile organizational theory and the triple helix model emphasize the role of government as an actor who can play the similar role, which can be achieved mainly through national labs and close linkages with academia. The second most popular role is creating a market and regulations accordingly. This role can be performed by the public or private sector. In order to support technology innovation, the theories of innovation systems, and innovation and economic growth identify the third role – educating potential developers and users, so they can use and apply the innovation. This role can be played by the industry, the academia, or both. Finally, the commercialization and production of a technology becomes the last role to diffuse the technology for mass users. Only the industry actor group can conduct this role according to the holistic model of innovation and innovation systems. Table 1 summarizes these four identified roles and three actor groups.
### Roles and Actors identified in Innovation and Socio-Economic theories

<table>
<thead>
<tr>
<th>Theory</th>
<th>Publications</th>
<th>Roles</th>
<th>Actors identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear model of innovation</td>
<td>Godin 2006</td>
<td>Detection of needs and solutions</td>
<td>University, Industry</td>
</tr>
<tr>
<td>Holistic model of innovation</td>
<td>Mahjoubi 1996</td>
<td>Detection of needs and solutions</td>
<td>University, Industry</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Development, production and commercialization</td>
<td>Industry</td>
</tr>
<tr>
<td>Innovation systems</td>
<td>Denning 1997</td>
<td>Detection of needs and solutions; Learning ICT</td>
<td>University, Industry</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Development, production and commercialization</td>
<td>Industry</td>
</tr>
<tr>
<td>Innovation and economic growth</td>
<td>Lipsey et al. 1998</td>
<td>Market creation and regulation</td>
<td>Government, Industry</td>
</tr>
<tr>
<td></td>
<td>Smit 2007</td>
<td>Market creation and regulation</td>
<td>Government, Industry, Learning ICT</td>
</tr>
<tr>
<td>Organizational theory</td>
<td>Nelson 1959</td>
<td>Detection of needs and solutions</td>
<td>Government, University, Industry</td>
</tr>
<tr>
<td></td>
<td>Sang-Young et al. 2002</td>
<td>Market creation and regulation</td>
<td>Government, Industry</td>
</tr>
</tbody>
</table>

**Table 1.** Roles and Actors identified in Innovation and Socio-Economic theories

## 3 CONCEPTUAL FRAMEWORK

From our review, we identified three main relevant categories of actors – Government, Universities, and ICT Industry – and four major roles for the establishment of ICT in the socio-economic context: 1) Detection of users’ needs and solutions, 2) Development, production and commercialization, 3) Market creation and regulation, and 4) Learning.

We acknowledge that there are sub-categories with heterogeneous actors, for example, various firms in ICT Industry. However, this study focuses on the macro level in establishing ICT. It means, for instance, that we emphasize the difference between a government systematically leading the detection of needs and solutions and the situation in which companies in the ICT industry compete freely for the detection of needs and solutions in establishing ICT.

We also recognize that the four identified roles (functions) are a simplification of the process of ICT establishment. This process is much more complex and includes many specific roles. As mentioned,
again, our research focus is not how the roles should be completed at the micro level, but what the major roles that should be accomplished at the macro level are, what the possible configurations between actors and roles are in establishing ICT, and whether those configurations have an impact on the measures of ICT establishment. Consequently, this theoretical framework can help researchers, practitioners and policymakers to better explain, analyze, and improve the existing configurations in establishing ICT, so that they can reconfigure the patterns of actors and roles in order to maximize positive socio-economic impacts for the region while minimizing the negative side.

Our overall perspective for the establishment of ICT is a compound of processes connected in a non-linear manner (Mahjdoubi 1996) to first detect users’ needs and solutions, subsequently develop, produce and commercialize ICT products and services in a market created and regulated to foster the use of ICT. One of most effective to complete this process is through education and training developers and users, which ultimately, not only improves performance, but also expedite the technology diffusion. We will examine now each role with the special abilities of the actors who can offer and accomplish the corresponding function.

3.1 Detection of users’ needs and solutions

It is critical to identify the user’s needs and especially to detect the technological opportunities with great potential to generate value (Bond et al. 2003). Once market needs and opportunities are identified, actors think about possible solutions to fulfil these needs and opportunities. The value of adopting technology can be estimated by analyzing the surrounding situation and technological possibility for a solution in the particular socio-economic system (Mahdjoubi 1996, Mahmood et al. 2005). If solutions have been successfully developed and established in other regions, then adoption can be one way to satisfy the need if the culture and technologic-economic platforms allow imitation (Spencer et al. 2005). If no available technology fulfils the needs, innovation is compulsory. This decision is critical, because the investment in innovations may denote an impossibility to shift to alternative solutions when the new technology is in the process of development or production due to high switching costs (Tyler 1962). Government, Universities, or the ICT Industry are able to act as a detector of needs and solutions by conducting studies in national labs, university labs and companies’ R&D divisions (Lazzeretti et al. 2005, Pradhan 2002); the strengths and weaknesses of each actor are:

- **Government**: Governments can have first-hand information from the community when they conduct projects. They sometimes directly survey the population to identify needs. Many governments in developing countries look for solutions in technologies that can be imitated from developed nations, which can be useful in the early stages of innovation (Mahmood & Rufin 2005). Conversely, in a more advanced innovative setting, paternalistic governments may engender passiveness and restrain innovation when making decisions (Spencer et al. 2005). The main weaknesses of governments are the bureaucratic processes that slow procedures while consuming resources and efforts. Additionally, political forces may act based on factors different from efficiency or social welfare (Hamlett 1991).

- **ICT Industry**: Opportunities to open new markets by innovation or imitation are the main incentives for the ICT industry to detect needs and solutions, so companies in ICT industry can be the first mover. However, due to their nature, companies will ignore signals to meet some of the needs unless they expect profit opportunities in the future. Large investments can be risky for companies because technology research and development consume time and resources without guaranteeing a return on investment (Dehning & Richardson 2002). If no economic benefits are expected, the opportunity is not attractive to companies.

- **Universities**: Their traditional primary missions are teaching and research (Etzkowitz 1989). Rigorous research methods can help in identifying needs, solutions and impacts (Nelson 1959, Mahdjoubi 1996, Godin 2006); however universities may be biased in favor of developing basic research rather than applied works. Some universities with entrepreneurial vision have engaged in applied projects, which can present an advantage for this role (Etzkowitz & Leydesdorff 1995, 1998, 2000, Lazzaretti & Tavoletti 2005, Shinn 2002).
3.2 Development, production and commercialization

The mass production of ICT involves the coordination of productive processes to create and deliver value, including the purchase of raw materials or equipment, the use of a skilled labor force, product design, and strategic financial and administrative management (Riches 2003). In democratic systems, production is mainly conducted by private sectors; this may be different in centrally planned economies in which all the productive decisions come from the central government. In addition, we need to remember that the commercialization of technologies can be very specialized in high-tech markets, in which complex technology products require a skilled workforce (Foxall 1988).

- **Government:** Many innovative technologies in ICT have been born in not-for-profit research centers or projects sponsored by governments, such as the Internet and other technologies that were first designed for military purposes (Noble 1984). Governments can be directly or indirectly involved in coordinating and formulating the value network that keeps control over the nation’s resources based on different socio-economic governing systems, which may or may not be more efficient than an open market approach (Mahmood 2005). The perspective of paternalistic governments may be involved in not only production, but also the distribution of social benefits.

- **ICT Industry:** The Industry is commonly the main player for this role. Companies have consistently developed or imitated (adapted or adopted) new ICT to protect and extend their market share. From the market perspective, configuring value networks for production and commercialization is about business objectives, so the successful configuration of value networks in this role is directly related to companies’ prosperity.

3.3 Market creation and regulation

The creation of markets is referred to as the construction of a structure where buyers and sellers can trade (Burton Jr. 1993). In this structure, the public sector plays a significant role in defining policies and rules, which is an important factor in the establishment of technology (Lipsey et al. 1998). This rule will oversee the market to foster transparency and efficiency, keep the market away from monopolistic practices, so it provides justice in cases of policy violation. In this role, the authority with the attributions to create laws and supervise their observance is the government; however, in practice, informal regulations and lobbying forces can also be found. The regulations that a government authorizes can hopefully encourage positive effects while discouraging negative impacts; however, companies in the industry also create invisible self-regulating rules and practices that impact market performance (Williamson 2005).

- **Government:** Governments should pursue a transparent function of markets through the creation and implementation of adequate regulations and policies (Williamson 2005). A well designed structure under the proper regulations and policies will attract more investors and foster good practice, so it will consequently speed up the ICT establishment (Lipsey et al. 1998). We should acknowledge, however, that excessive regulations and policies by the government can slow the process and self-evolving industrial dynamics in establishing ICT.

- **ICT Industry:** ICT Industry can provide feedback to the government to improve regulation that can bring more positive dynamics in the industry. However, this procedure can be influenced by dominant companies in the market for their own advantage rather than for overall socio-economic benefits.

3.4 Learning ICT

The process of learning contributes to establishing ICT in different ways (Smit 2007). One is educating customers to adopt a technology more efficiently and effectively when users search for. Another is training specialists to develop or adapt ICT. The specialists usually learn basic knowledge and gain more sophisticated knowledge through directly developing or adapting ICT (Arrow 1962). The complexity of ICT requires professional skills and training that people can acquire in universities.
or companies they work for (Ross 1967). Governments can foster the info-alphabetization or influence the design of educational curricula; however governments do not directly provide education or training to people to establish ICT.

- **ICT Industry**: ICT development or adaptation is directly related to companies’ fates. Thus, companies invest in training their personnel and supporting research centers to maintain or achieve technical capabilities. Companies also educate and support customers to adopt and use their technology more efficiently or effectively through support centers, conferences, or educational institutes.

- **Universities**: One of traditional university missions is teaching people to become professionals who are able to use and improve technologies (Denning 1997). Universities provide highly skilled personnel not only by teaching them through courses but also by giving them opportunities to conduct applied research in conjunction with companies.

### 4 TYPOLOGY OF THE ROLES AND ACTORS

Doty and Glick (1994), argue that typologies are complex theories that frequently misinterpreted. A typology is different from classification in representing series of logical arguments that connect concepts, constructs or variables, and incorporating multiple levels of theory that can predict variance in dependent variables.

Having identified the fundamental roles and actors needed to establish ICT in socio-economic system, we now introduce the typology of the possible configurations between the roles and actors that shape the establishment of ICT. By studying the configurations of different socio-economic systems, we may find some patterns that help in creating strategies to improve the rates and impacts of establishing ICT, or to identify configuration benchmarks that reveal the most effective organization (Niosi 2002).

In order to build the typology, we will consider one leading actor for each role, so there will be one, two, or three possibilities to fill the leader position for each role. For example, in the role of Detection of user’s needs and solutions, there are three actors (Government, ICT Industry, and Universities) as a potential leader. Table 2 presents the configurations in the form of a menu with four columns that represent the roles; for each role, the potential actors to perform the role are listed in the corresponding column. A total of 24 different configurations \((3 \times 2 \times 2 \times 2 = 24)\) can be formed by identifying one leader for each role. The configurations would identify various patterns, for example, a government performing most roles in the establishment of ICT (Government leading the detection of needs and solutions, the development, production and commercialization, and the market creation and regulation, with academia or IT Industry leading the learning process for ICT establishment). In another instance, ICT industry plays major roles in the process of ICT establishment when the ICT industry leads in all four roles. Other configurations can represent mixed distributions of roles played by actors influenced by market and political forces (Denning 1997).

<table>
<thead>
<tr>
<th>Detection of users’ needs and solutions</th>
<th>Development, production and commercialization</th>
<th>Market creation and regulation</th>
<th>Learning ICT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government</td>
<td>Government</td>
<td>Government</td>
<td>ICT Industry</td>
</tr>
<tr>
<td>ICT Industry</td>
<td>+ ICT Industry</td>
<td>+ ICT Industry</td>
<td>Universities</td>
</tr>
<tr>
<td>Universities</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2. Configurations Map for ICT Establishment**

For example, in many developed western European countries like Denmark, all roles are led by ICT industry except the role of market creation and regulation (Haekkerup 1999, Shapira & Rosenfeld 1996), while in developing countries like Indonesia, governments has involved in not only the role of market creation and regulation, but also the role of detection of users’ needs and solutions.
(Madanmohan & Kumar & Kumar 2004, Sharif & Sundararajan 1984). Is it coincident or natural? Can poorly performed nations in term of establishing ICT improve their performance by changing the configuration of roles and actors?

The proposed configuration will provide a theoretical basis to answer these questions. This configuration can be used 1) to compare countries and regions that have different socio-economic impacts in establishing the same ICT, and 2) to analyze the historical development of single country or region. By studying the configurations that countries have adopted over time, we are in a better position to explain how various actors can perform differently on a given role and how the performance of actors on different roles impacts the process of establishing ICT. In this way, we may identify patterns that help to explain why and how some countries can sustain their development and have more positive socio-economic impacts than others. In addition, we can find and propose more appropriate configuration for a country through observing and analyzing the evolving configurations of countries.

5 DISCUSSION, LIMITATIONS AND FUTURE RESEARCH

Although this study is a simplified representation of the ICT establishment process, it contributes to knowledge by identifying the major actors and roles and suggesting a holistic way to analyze their intertwined relationships which can be further used to describe patterns and benchmarks in establishing ICT. The proposed framework provides an alternative explanation that complements existing theories about technology adoption, diffusion, innovation and the socio-economic impact of establishing ICT. We acknowledge as a limitation that empirical information may not always identify one leader for each role, but roles may be performed collaboratively (or competitively) by many actors; however, this paper introduces the concept of the configuration that affects the establishment of technology.

The 24 theoretical configurations introduced in this study are based on all possible combinations of actors and roles defined; however, we may find a smaller subset of configurations in practice, which opens up a whole new avenue for future study in recognizing and describing what are the factors that produce the different configurations and their impacts. In this vein, researchers can survey nations and regions to identify how many configurations of roles and actors there actually are, whether there are patterns that can be categorized by geographic regions or economic levels, and how the configuration of roles and actors can be reshaped over time in order to improve the establishment of ICT. Researchers, practitioners and policymakers can use the proposed framework from evaluating current configurations to designing strategies for reconfiguration that stimulate positive impacts while repressing negative effects in the ICT establishment.

For future research, we can test, confirm, and improve this proposed framework through empirical studies based on quantified and qualified data. Based on further empirical analysis, researchers and practitioners can develop recommendations for the strategies of ICT establishment that can provide more specific and practical roles for governments, universities and companies.

References


Proceedings ECIS 2009
Dehning, B., and Richardson, V.J. (2002) Returns on Investments in Information Technology: A
134.
619.
laboratory for knowledge-based economic development. EASST Review, 14 (1), 14 - 19.
University-Industry-Government Relations, London.
"Mode 2" to a Triple Helix of university-industry-government relations. Research Policy, 29 (2),
109-123.
(2-4), 87-99.
Feinson, S. National Innovation Systems: Overview and Country Cases. Rockefeller Foundation, 13 -
38.
Management of Knowledge: The Case of Infosys Technologies, 29 (1), 9-33.
Hall
technology innovation. Paper presented at the Portland International Conference on Management of
Engineering and Technology PICMET '01.
Development: The Case of the Entrepreneurial University of Twente. European Planning Studies
13 (3), 475-493.
systems. International Journal of Contemporary Sociology 42 (1).


Shapira, P., and Rosenfeld, S. An Overview of Technology Diffusion Policies and Programs to Enhance the Technological Absorptive Capabilities of Small and Medium Enterprises. Chapel Hill, NC.


Tripathi, M. (2006) Transforming India into a knowledge economy through information communication technologies—Current developments. The International Information and Library Review (38), 139 - 146.


THE BUSINESS INTELLIGENCE COMPETENCE CENTRE AS AN INTERFACE BETWEEN IT AND USER DEPARTMENTS IN MAINTENANCE AND RELEASE DEVELOPMENT

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0526.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Intelligence, Agile computing, Empirical study, Governance</td>
</tr>
</tbody>
</table>
THE BUSINESS INTELLIGENCE COMPETENCE CENTRE AS AN INTERFACE BETWEEN IT AND USER DEPARTMENTS IN MAINTENANCE AND RELEASE DEVELOPMENT

Baars, Henning, University of Stuttgart, Breitscheidstraße 2c, 70174 Stuttgart, Germany, baars@wi.uni-stuttgart.de
Zimmer, Michael, University of Stuttgart, Breitscheidstraße 2c, 70174 Stuttgart, Germany, zimmer@wi.uni-stuttgart.de
Kemper, Hans-Georg, University of Stuttgart, Breitscheidstraße 2c, 70174 Stuttgart, Germany, kemper@wi.uni-stuttgart.de

Abstract

The growing relevance, scale, and complexity of Business Intelligence (BI) entails the need to find agile and efficient solutions for the coordination of maintenance and release processes – under consideration of the heterogeneity of the involved units on the IT and the business side. The finance industry with its mature BI infrastructures and its highly turbulent business environment is a forerunner for these developments. Based on a survey among BI users in the finance sector, relevant problem areas in the BI service provision are identified and structured. A series of qualitative interviews among banks and insurance companies is used to gain further insights into approaches for dealing with the related issues. The studies uncover several advantages of a central “BI Competency Centre” (BICC) as well as levers for effectively structuring the interfaces between BICC, IT, and user interface.

Keywords: Business Intelligence, Empirical Study, Governance, Service Orientation
1 RELEVANCE

In recent years, Data Warehouse (DWH) based infrastructures for management support have grown tremendously in relevance, scope, and scale. The integrated approaches they are built upon are usually referred to as Business Intelligence (BI). In more and more cases, BI reaches out to all business functions and managerial levels and is turning into an integral segment of the business (Baars and Kemper 2008). On the one hand, these developments open up significant business potential. On the other hand, however, they come at the price of increasing costs and a constantly rising technical and organizational complexity. These challenges need careful management that includes a thorough reconciliation of technical and application oriented requirements (Hughes 2008, Miller 2006). The issues are more severe in the domain of BI than in the operational IT because BI applications are required to embody ever changing business semantics and versatile analytical needs (Finger 2008, Moss 2006). Many organizations have reacted by establishing specialized units for running and supporting BI solutions – central BI “centres of competence” (BICCs). In general, BICCs act as linking pins between the user side and the infrastructure provision (Miller 2006, Unger & Kemper & Russland 2008). However, it is yet unclear under which conditions a central BICC is to be preferred over a decentralized approach with BI units spread across different user departments, how to distribute competencies among the user, IT, and the BICC side, and how the respective interfaces need to be crafted under consideration of the constant stream of system modifications and re-designs (Eckerson 2006, Geiger & Hill & Ton 2007, Strange & Hostmann 2003).

When tackling this issue, the literature on IS (de)centralization provides a rich body of rationales for deriving the spectrum of basic organizational options. However, the respective sources are not geared at the particular characteristics of the BI domain and are thus hardly applicable without further modifications (Boddy 2005, Gordon & Gordon 2004, Haag 2003, Inmon 2005, Inmon 2006, Kimball 2002, Laudon & Laudon 2005). It is yet unclear what the core priorities are in the BI domain with respect to the involved users, the IT and the BI units (Kleese & Winter 2007). There are also several relevant BI characteristics to be considered, e.g. the need to built data structures that feed both unit specific analyses and cross-departmental reporting applications, or the blurry distinction between “routine” changes (e.g. modifications of established data sets, often in form of so called “data cubes”) and more far reaching requirements (e.g. in case such modifications involve the integration of new source systems or require a higher load frequency).

This contribution deals with these issues by deriving a service-based concept for the distribution of tasks within the systems maintenance and the release development phases of BI solutions. It gains insights from application oriented research conducted with insurance companies and banks. Financial organizations have been pioneers in the realm of BI for a long time and thus come both with a rich experience base and comparatively mature solutions (Kleese & Winter 2007). Furthermore, external pressures are especially salient in this industry due to the volatility of the business environment. As will be discussed in the final section, most of the core conclusions are expected to be invariant to industry peculiarities because the underlying qualitative rationales are not affected.

Due to the novelty of the phenomenon in discussion, the research is of an explorative nature (Schwab 2005, Yin 2009). It encompasses an explorative quantitative study on the prioritization of BI maintenance and release development processes as well as a series of interviews in the finance industry that provides insights into the options, possibilities, and challenges of their organization. The results motivate the draft of a concept that aims at overcoming remaining shortcomings of a BICC approach by distinguishing three classes of services.

2 CONCEPTUALIZATION

The presented results focus on the financial sector. The financial sector is commonly understood to encompass all those organizations that are primarily dealing with the allocation, management, acquisition and investment of monetary resources and thus with all kinds of matters related to money and the financial markets (Howells 2007). Among others, companies in the finance sector include full service banks, insurance companies and investment trusts (Moormann 2007). The relevance of BI in
these organizations is driven not only by the fact that information is usually a core product but also by a number of pressures both from the regulatory and from the market side. Recently, developments like the subprime crises (Blundell-Wignall 2008) or the growing number of natural disasters are additionally fostering the need for an agile reporting and analysis infrastructure that provides a comprehensive overview on internal and external developments – based on integrated data pools. (Moormann 2000). It does not come as a surprise that finance sector organizations, and banks and insurance companies in particular, are among the most avid users of advanced BI infrastructures. The finance sector can thus be treated as an epitome for current developments in the BI domain (Kimball 2002, Moormann 2007).

This contribution deals with the organizational interfaces between the participants involved in the maintenance and release development for BI solutions. Here “maintenance and release development” is understood to encompass all types of alterations on an existing solution (IEEE 2008). Maintenance is used for minor changes that can be dealt with in near-time within the established reporting structures. Release development adheres to the definition of a project. In the light of the presented aims, the scope of this research is confined to changes that are involving the user department and that lead to modifications relevant for and perceivable from the user side. Among other activities, this excludes backend optimizations, the versioning of data management tools or redesigns of transformation logic that is not visible from the frontend (Baars & Horakh & Kemper 2007).

The studies that led to this contribution are embedded in a larger research context that tackles BI organization and its IT support (Baars & Horakh & Kemper 2007, Horakh & Baars & Kemper 2008, Unger & Kemper & Russland 2008). The fundamental research objective was to identify viable options for the design of the interface between users, the BI units and/or the IT department in order to effectively cope with perceived shortcomings or to pursue desired enhancements. Due to the dearth of research into the subject, it was tackled according the paradigms of exploratory research (Schwab 2005) and addressed with two complementary studies that have been conducted in the first half of the year 2008: A descriptive, quantitative survey to get a first overview on fundamental priorities and a series of qualitative interviews that was geared at a more in-depth understanding of the subject matter.
shortcomings in the BI area on the one hand and demanded enhancements on the other. The results are used to structure and guide the further exploration which is achieved by the interviews.

A pivotal factor that needs to be introduced for a valid interpretation of the interviews is the type of requirement that is dealt with in a particular system change. It was expected that different types of requirements would lead to different ways of dealing with them and gaining an understanding of those relations was a major part of the study. The second group of insights deemed to be relevant pertained to rules and regulations that had been imposed and their respective relevance. To grasp relevant causal and temporal relationships, questions on the concrete experiences behind the chosen design options were introduced. All this should condensate in the actual design of processes and projects for maintenance and release development. The resulting conceptual framework is depicted in Figure 1.

3 PRIORITIZATION – QUANTITATIVE EXPLORATION

The questions on the prioritization were covered as part of a wider online survey. As depicted in the conceptual framework, they encompassed an investigation of the type of BI organization, and of shortcomings and demands for enhancements regarding BI service provision. From about 11,000 addressees who were contacted via personalized emails, 1,519 followed the link to the survey web page. A total of 97 persons from the finance sector filled out all questions relevant for the discussed research. The sample was gathered from a list of experienced BI practitioners from IT departments (19 respondents), dedicated BI units (19 respondents), and BI user departments (59 respondents). Due to the explorative nature of the study, the results were primarily approached with descriptive statistics and used as a means for hypothesis generation rather than hypothesis testing (Schwab 2005). All results where checked with Chi-square tests for significance – based on a significance level of 10%. Due to the qualitative design and the sample size, differences in the results will also be discussed when they are not statistically significant but nevertheless lead to insights into potential differences among the groups. The shortcomings and demanded enhancements that were identified to be most relevant with respect to the overall study will be discussed in the following two sub-sections.

3.1 Shortcomings

Two questions addressed the shortcoming side of the framework: The participants were asked to select the three most pressing issues in for BI maintenance and release development respectively. The lists of 9 choices each were compiled based on BI literature and prior field experience. The differentiation of the results according to the type of organizational unit exposed diverging views on the shortcomings in BI maintenance.

Striking is the difference in the evaluation of data quality which was – in line with results from similar studies – the issue that heads the overall list of chosen shortcomings: Only a minority (16%) of the respondents from IT units considered data quality to be among the top-three relevant issues, while participants from user departments chose this item in almost half of the cases (46%). In contrast, the number for the BI units topped 60%. When evaluated with a Chi-square test, these differences also prove to be significant with respect to the chosen significance level (p=0.077). This result can be interpreted as an indicator for the distance of the IT staff from the use context on the one hand and the underestimation of data quality related efforts by the users on the other. In a BI unit, however, data quality related issues converge and culminate. A possible similar discrepancy might exist regarding the diversity of the analysis systems. This item was declared slightly more often to be a problem in the business departments than in the BI and IT units. A reason could be that tool integration is primarily driven by handling the frontend.

Interestingly, about a third of the respondents in the BI and business units selected missing functionality as a core issue while the result for IT is zero (significant with p=0.054). Again, this reveals a lack of insight of IT into the application side.

The IT seems to be a little more conscious of an insufficient coverage of informational needs (42%), which a slightly lower percentage of users (32%) and only a minority of the BI unit respondents (12%) considered to be an issue. A possible explanation could be that IT has an overview on all potential data sources while the user department is not yet fully aware of those possibilities. Being situated between
IT and user department, the BI unit has the best vantage point over the actual information supply and demand. A different result comes up for the lack of integration into the IT-backend which neither for the IT (16%) nor the user department (24%) ranks among the most frequently selected issues while still 32% of the BI unit saw this as a challenge. Although the results are not statistically conclusive yet, they again might hint at the tendency of the IT and business departments to underrate the conceptual challenges that go along with the integration of heterogeneous business semantics.

Regarding the release development for BI applications, 61% of the participants from the user side complained about an unsatisfactory time budget [of the BI providers] to support the user departments. In contrast, 28% of the IT and BI unit participants considered this to be of particular relevance (significant with p=0.067). This can be interpreted as another indicator for diverging priorities: The providers act according to a defined prioritization and under consideration of available resources. As lower prioritized requests have to wait, business departments are left unsatisfied. This goes in line with the occurrence of development backlogs which are considered critical by all three groups with item selection rates that vary between 30% and 40%.

Not surprisingly, only a small group (16%) of participants from IT considered lack of maintenance and release knowledge in IT to be a problem. When contrasting this with the values for the user departments and for the BI units (together 36%), one can suspect that the IT might indeed overrate its abilities. Similar, about 20% of the BI and user units complained about a lack of models for agile development, while the number is a little lower for the IT side with 16%. Regarding the item lack of management support, the results are consistent across all groups (about 30%).

It is obvious that the views of the participants vary – and in some cases (data quality, functionality, time budget to support user departments) even significantly. Several answers indicate that the full impact of BI is most comprehensively grasped in the BI unit. This can not only be explained by the hinge function of the BI unit between the user and the IT side but also because of its degree of specialization. Ideally, the different priorities that surface in the answers are reconciled by flexible organizational concepts that – where possible – leverage the different awareness of issues among the groups.

3.2 Demanded enhancements

The second aspect of interest pertained to opportunities for enhancing the value of BI. The questionnaire covered this in two questions with 9 choices each.

With respect to the maintenance side, a functional requirement holds a top rank: 63% of the dedicated BI units demanded more degrees of freedom in the data analysis applications. This wish was not shared uniformly: On the user side, only 47% selected this item (and only 21% in IT). A possible interpretation is that the BI side rated this issue from the information production perspective (generation of reports) while the answers of the users reflect a more passive, consumption oriented view. This would also explain why this result is not corroborated when it comes to the question of data mining: About half of the users (47%) demanded such systems but only 21% of the BI units and only 11% of the IT. On the content side, there was a perceived need for analyses across business function that was seen by both the BI and the user side (each about 40%) – but not in IT where only 16% selected this item. Obviously, BI users see demand for more in-depth and more far-reaching insights – although not necessary in tandem with more degrees of freedom or even a higher number of tools.

Shared wishes are a higher degree of integration of external data (around 35%) and portals as single point of contact (around 30%). Portals have a slightly higher importance for the BI while data integration is more important for the user departments.

With respect to organizational measures, a better data quality management dominates the answers of all three types of organizations. This result is consistent with the shortcoming side.

Regarding release development, the item business departments should be given more room to implement their own solutions is of special importance. This option was demanded by 49% of the participants from user side, and surprisingly still by about 25% of both the BI and the IT participants (a significant difference with p=0.050). Obviously, despite all proclaimed trends to concentrate IT...
development, there is a remaining need for decentralized activities that is especially felt in the user departments. Mirroring this against the issues of backlog and unsatisfactory time budgets, a potential to find a shared solution to issues of fast and flexible development is discernable. It is equally noteworthy that rules between business department and IT is not once selected by respondents from IT units (16%) while user participants chose this in 44% of the cases (significant with p = 0.053).

On the structural side, an integration of a BICC for future projects was demanded by about 40% of the BI units and by slightly fewer participants from user departments. The fact that a BI unit has been established does not necessarily mean that it also takes over the role of a BICC and is tightly integrated into the company. The answers from IT only reached a mark of about 20% - a result that might again signal an underrating of problems by the IT. A similar situation can be found when it comes to the wish for faster implementation which is conceived to be a problem by 35% of the participants from business and BI units but only for 10% of the IT participants (significant with p=0.075).

Service orientation and better data loading processes were of some relevance, although the numbers are not particularly illuminating. Yet, the service orientation (in the sense of delineating, providing, and managing customer-oriented services in the sense of IT service management) is seen as a building block for the implementation of the concept that is derived here.

In conclusion, the results provide some ideas for designing the interface between the units: They indicate that there is a need for a differentiated, carefully governed division of labour that allows bringing together seemingly conflicting requirements regarding the management of data quality, fast reaction times, more responsibility for the user departments, and providing cross-departmental solutions. The BICC stands at the core of this.

4 DESIGN OPTIONS – QUALITATIVE STUDY

For the qualitative case interviews (Yin 2009), a total of ten companies from the finance sector with mature BI infrastructures were selected. Their business model, type of BI organization, number of users, and BI experience (measured in years of usage) is presented in Table 1. In each case, interview partners from dedicated BI units were selected – to capture the more holistic view of BI units that became visible in the survey. This encompasses both centralized and enterprise-wide BICCs (4 cases) as well as specialized BI units that report to a user department and are part of a decentralized BI approach (6 cases). Each of the interviews lasted between 60 and 110 minutes. The interview questionnaire contained qualitative, open questions that were structured around the conceptual framework that has been introduced in section 2.

<table>
<thead>
<tr>
<th>Business model</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
<th>Case 6</th>
<th>Case 7</th>
<th>Case 8</th>
<th>Case 9</th>
<th>Case 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>International group of Banks</td>
<td>100</td>
<td>12</td>
<td>100</td>
<td>10</td>
<td>90</td>
<td>80</td>
<td>30</td>
<td>70</td>
<td>50</td>
<td>300</td>
</tr>
<tr>
<td>German group of Banks</td>
<td>1000</td>
<td>1000</td>
<td>12</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>90</td>
<td>70</td>
<td>50</td>
<td>300</td>
</tr>
<tr>
<td>Automotive Bank</td>
<td>100</td>
<td>1000</td>
<td>12</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>90</td>
<td>70</td>
<td>50</td>
<td>300</td>
</tr>
<tr>
<td>Insurance company</td>
<td>100</td>
<td>1000</td>
<td>12</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>90</td>
<td>70</td>
<td>50</td>
<td>300</td>
</tr>
<tr>
<td>Global Bank</td>
<td>100</td>
<td>1000</td>
<td>12</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>90</td>
<td>70</td>
<td>50</td>
<td>300</td>
</tr>
<tr>
<td>BICC</td>
<td>100</td>
<td>1000</td>
<td>12</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>90</td>
<td>70</td>
<td>50</td>
<td>300</td>
</tr>
<tr>
<td>Decentralised BICC</td>
<td>100</td>
<td>1000</td>
<td>12</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>90</td>
<td>70</td>
<td>50</td>
<td>300</td>
</tr>
<tr>
<td>Decentralised BI</td>
<td>100</td>
<td>1000</td>
<td>12</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>90</td>
<td>70</td>
<td>50</td>
<td>300</td>
</tr>
</tbody>
</table>

Table 1. Characteristics of examined cases

As illustrated in Table 2 and Table 3, the results were aligned with the identified challenges and opportunities for enhancements. Marked with a “-” are cases in which the identified challenges still remain unsolved or where the opportunities are not yet grasped while the “+” indicates the availability of an approach to deal with the various subjects. It is noteworthy that most of the companies with a centralised BICC (cases 1-4) have found at least partial solutions to most of the problems. In fact, it can be derived from the interviews that centralised and decentralised BI departments follow different strategic trajectories: The centralised BI departments show a strong focus on questions of efficiency while their decentralised counter-parts are primarily occupied with the fulfilment of the business departments’ requests. Units that report to user departments naturally have weak points when it comes to the consolidation of the diversity of analysis systems or tackling a lack of integration into the IT-
Moreover, centralised BI departments have also come up with ways to turn the discussed demands into reality. In the following sections, the different approaches are discussed in further detail.

### Table 2. Challenges of BI and the availability of solutions in the cases

<table>
<thead>
<tr>
<th>Shortcomings of running the system</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
<th>Case 6</th>
<th>Case 7</th>
<th>Case 8</th>
<th>Case 9</th>
<th>Case 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diversity of analysis systems</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>O</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Lack of integration of BI-solutions into the IT-backend</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shortcomings of developing the system</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
<th>Case 6</th>
<th>Case 7</th>
<th>Case 8</th>
<th>Case 9</th>
<th>Case 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory time budget of the BI providers to support the user department</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Development backlogs</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Lack of maintenance and release knowledge in IT</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>No models for agile development</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of BI-organization</th>
<th>central</th>
<th>decentral</th>
</tr>
</thead>
</table>

*Legend:* unsolved, 0 not relevant, + approach available

### Table 3. Demands for enhancement and the availability of solutions in the cases

<table>
<thead>
<tr>
<th>Demands for running the system</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
<th>Case 6</th>
<th>Case 7</th>
<th>Case 8</th>
<th>Case 9</th>
<th>Case 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Options for business departments to implement their own solutions</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Demands for developing the system</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
<th>Case 6</th>
<th>Case 7</th>
<th>Case 8</th>
<th>Case 9</th>
<th>Case 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Better processes for data loading</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Integration of a BICC</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Faster implementation</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Service orientation</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>Rules between user department and IT</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-50</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Advancement of data quality management</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of BI-organization</th>
<th>central</th>
<th>decentral</th>
</tr>
</thead>
</table>

*Legend:* unsolved, 0 not relevant, + approach available

### 4.1 Challenges and approaches to deal with them

#### Diversity of analysis systems

In the study, an unhealthy diversity of analysis systems only occurred in companies with a decentralised BI organization. The central BICCs were able to impose regulations that constricted the range of tools to a defined set. Furthermore, they guided or supported the user departments when it came to developing new releases or adapting existing ones – and in the process managed to curb or prune uncontrolled growth.

Cases 1 and 2 are prime examples for imposing a defined and confined set of tools. The interview partners from these companies also emphasized the resulting opportunity for systematically building up knowledge about the maintenance and the development of the respective applications.

The picture is totally different for decentralised BI units which primarily adhere to the wishes of the separate departments: This approach fosters heterogeneity between different solutions. The company of Case 9 can act as an example for the resulting issues: The interview partner described experiences with handling the constant stream of new tools introduced – tools which all needed to be connected on the IT side and that brought about massive issues regarding the built up and maintenance of the required knowledge base.

#### Lack of integration of BI-solutions into the IT-backend

A centralised BICC also has advantages when it comes to the integration of BI-Solutions into the overall IT architecture. Because of their size and role they are able to react swiftly to BI-relevant
changes on the IT-backend. This can be illustrated with Case 2: Although BI and IT units were independent subsidiaries they established well defined processes to cope with changes on the IT-backend. Conversely, decentralised BI-Departments are usually neither important nor big enough to be integrated into the workflow of the IT-backend. This can even lead to a situation where relevant changes on the IT-backend are not communicated to the BI units at all. Examples are the companies in the Cases 8 and 10 which both struggle with insufficient information on changes of the backend infrastructure. Independently of the organizational variant, the integration into the workflows of the IT department is crucial.

Unsatisfactory time budget of the BI providers to support the user departments

A best practice for allotting time and space can again be found in the centralised BICC: Implementing separate budgets for release development and ongoing maintenance tasks. This dual-budget approach has been implemented in all 4 BICC cases. Especially in Case 10, the lack of such a distinction was a root cause for all sorts of problems – even minor changes to the systems had to undergo the steps for the set up of a formal project. In consequence, dual budgets support flexibility.

Development backlogs

Backlogs have been dealt with successfully both in some centralised and decentralised units. As the answers on the experiences and the resulting rules and regulations show, a well defined requirements engineering alongside a communicated prioritization is the practice of choice. Such an approach had been implemented in all four BICC cases as well as in the companies 5, 7 and 9. In Case 9, this was complemented by the use of agile development methods that actually embedded developers in the department.

Lack of maintenance and release knowledge in IT

The case studies support the assumption from the quantitative exploration that a centralised BICC is – due its very design as a mediator – best suited to get a comprehensive overview of both the user and the IT needs. BI units in user or IT departments usually do not see both sides of the coin. This leads to problems like insufficient budgets for creating sustainable concepts or the tendency of IT to prioritize changes on the BI systems very low, even if they have a huge impact on the company. The companies described in Case 8, 9 and 10 face these problems.

No models for agile development

Only the company in Case 9 applied a model for agile BI development that guided the release development. It is noteworthy that this company had introduced several standardized processes between the user department and the IT to curb efficiency losses and foster cost reduction. An issue that still awaits a solution is the handling of data integration processes. At best, a combination of agile BI development with defined processes and a BICC driven coordination is leveraged.

4.2 Opportunities to realize demands for enhancements

Options for business departments to implement their own solutions

As discussed above, the idea of supporting professional system development within the user departments can be achieved in a setting like Case 9 which is based on a high degree of decentralization. In case of a more centralised approach, a central BICC can open up carefully crafted options that allow trained members of the user department to design their own reports or analysis cubes. This solution relies on extending BICC competencies to trained power users. It has to be mentioned that only the BICC in Case 2 was able to offer training services for such purposes.

Data quality management and data loading processes

All centralised BICC had already been able to systematically introduce rules and guidelines to improve data quality management and data load. As discussed above, the decentralised BI departments in four out of six cases had not yet been equally integrated into the IT workflows. This aggravates the problems as it leads to a frequent lack of information.
**Integration of a BICC**

Even though they were not always familiar with the term, the centralised BI units found in the study conformed fully to the understanding of a central BICC. These units had created well-structured processes and responsibilities for data ownership or requirements engineering. A remaining problem for these BICCs is the time it takes to deliver solutions for the user side.

**Faster implementation**

The acceleration of changes and the increasing rate of release development projects still pose problems for all but two of the companies. Small and specialized BI units like the decentralised BI department in Case 9 with its own DWH have an edge in this regard because they can act immediately and without much coordination effort. This situation can barely be transferred to enterprises which are in need of a core data warehouse that spans department borders. Changes to the core DWH naturally take time as they usually involve a larger number of stakeholders. Up to now there is no solution implemented in the examined companies for churning out such modifications in a short time frame.

**Service orientation**

The central BICCs have all started to plan, develop or deploy dedicated BI services. The company in Case 1 has already developed several services, while the company in Case 2 is right in the process of service specification. In contrast, most of the decentralised BI departments have not even thought about introducing services – which is not surprising as they are woven together with user departments and will by design not push a higher degree of bureaucracy. The exception to the rule is, again, Case 9. Here BI development services are used as building blocks in the agile development model.

**Rules between user department and IT**

All those companies that have implemented a BICC have also defined a set of rules. Actually, implementing a BICC as an organizational entity naturally involves the definition of rules and responsibilities – the establishment of a central BICC fosters regulation. In contrast, companies with a decentralised approach struggle much more with the definition of rules because too many parties with heterogeneous systems and requirements have to be factored in. That does not mean that this is an impossible endeavour: two decentralised companies succeeded in that regard and came up with rules that govern the interface between IT and BI. The company in Case 9 only has rules for ramping up new applications while in Case 7 strict rules had been dictated by the IT side that had to be followed by all decentralised BI units. There is a catch: In Case 7, the development times are as slow as if they would be when following a centralised approach.

**5 CONCLUSIONS**

The presented results show that there is a particularly strong case to establish a central BICC which coordinates the BI tasks of maintenance and release development and that acts as a mediator between IT and user departments. Although it is certainly not impossible for decentralised BI organizations to enforce enterprise-wide stipulations that govern tasks like data quality management, clear requirement prioritization, a definition of allowed analytical systems or the integration with the backend systems, this is much harder to achieve in decentralized environments. Likewise, the IT cannot easily take over the role of a BICC as it lacks its “dual-view”. Actually, the very set up of a central BICC is not possible without the parallel introduction of at least some rules that define the interfaces between IT, user, and BI unit. The exception to this logic is Case 9 where the solution is facilitated by a narrowly confined scope of the supported solutions and a separated single-purpose DWH. Notwithstanding, the definition of agile release development processes based on defined services achieved in this enterprise has to be regarded as groundbreaking. This is of particular interest for changes that are required to be inserted relatively quickly – which is especially relevant for the finance sector.

Ideally, the need for agility can be reconciled with the rationales that foster centralization and regulation. Both Case 9 and the four BICC cases offer a powerful lever to tackle this seemingly impossible task: The definition of BI services and the design of BI service oriented management concepts. A map of self-contained services helps to break down the complex tasks of BI maintenance from a service-user perspective and to distribute the services across different units. By grouping the
services based on their different needs for centralization, a hybrid of a decentralised and a centralised approach can be designed. As the experiences with the backend-systems highlight, it is recommendable to also add “informational veins” which transport messages on events at the backend (e.g. changes in the source system) or at the application side (e.g. new analytical requirements) – implemented e.g. by supporting infrastructure workflows or interfaces that are adhering to the Service Oriented Architecture (SOA) concept. Flexibility on the technical side supports agility on the application side.

Based on these considerations, it is proposed here to distinguish between three classes of services that are differentiated by the criteria participating partners and possibility to be based on well defined and technologically supported processes. This draft of a service based concept aims at exploiting the potential of decentralized self-services as far as possible without impeding the potential of a BICC to foster integration, governance, and efficiency. It is visualized in Figure 2.

<table>
<thead>
<tr>
<th>Automatable predefined services</th>
<th>Not automatable predefined services</th>
<th>Individual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modification of report layout</td>
<td>Adoption of reports for several departments</td>
<td>Creation of individual prototypes</td>
</tr>
<tr>
<td>Ramp-up of a front-end tool</td>
<td>Integration of analysis results into the backend (&quot;closed-loop&quot;)</td>
<td>New types of data load processes (e.g. near-time)</td>
</tr>
<tr>
<td>(from a set of defined tools)</td>
<td>Integration of new data sources that are potentially not unit specific</td>
<td>New types of source systems (e.g. with unstructured data)</td>
</tr>
<tr>
<td>Ramp-up releases or fixes</td>
<td>Tracking of changes in source data structures</td>
<td>Development of concepts to speed up processing times</td>
</tr>
<tr>
<td>New indicators based on existing data fields</td>
<td>Integration of new analytical tools</td>
<td></td>
</tr>
<tr>
<td>Coordination and calibration of existing ETL processes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integration of aggregates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integration of unit specific data sources</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Service classes and responsibilities

Class 1 refers to technical changes that can be tracked with workflow software and executed jointly by the IT department and the BICC in a semi- or even fully automated fashion. Examples encompass modifications to the report layout or in the calculation of performance indicators. This type of service needs to be handled at the provider side, but – due to its standardized nature – can be triggered from users in a self-service fashion. Streamlining such activities supports fast response times for standard changes while leaving data quality and backend integration tasks in the hands of the BICC. Furthermore, when system supported, it also fosters transparency e.g. regarding prioritization and front-end-integration – all of which are core concerns of the user side.

Class 2 addresses tasks that have repercussions through the whole system and involve more than one user department. Even in case of being well defined, these tasks cannot be equally automated as they involve negotiation and bargaining activities. Also, the initiative for starting a Class 2 service might just as well come from the BICC, e.g. in case of identifying new opportunities regarding data usage. Separating Class 2 from Class 1 services further increases a transparent prioritization and backlog handling while at the same time utilizing the role of the BICC as a coordination and communication hub between the different departments and the IT side. A careful handling of Class 2 type services can be expected to facilitate data quality management, the integration of external data sources, and the systematic built-up of cross departmental solutions.

Class 3 involves the maintenance and release development for individual solutions which are totally evading any kind of standardisation, e.g. the design of a prototype for a new class of analytical systems. These are separate projects which are naturally dealt with by a team of representatives from the user department and the BICC side. They might later trigger Class 2 or Class 1 services – when a designed system is going to be embedded in the overall BI infrastructure. Such services need to be dealt with altogether manually as there is no template to apply.
Figure 2 visualizes the three classes and gives some examples.

6 DISCUSSION AND OUTLOOK

Trailing the ever growing size and relevance of BI infrastructures, the subject of “BI organization” is likely to come further into focus as well. The presented results provide explorative insights into a particular variant of a BI organization: A dedicated and central BICC – a “Business Intelligence Competence Centre”. As has been shown, the establishment of a central BICC can for several reasons support a more professional development and maintenance of BI solutions. Its weaknesses primarily lie in transparency and reaction times – as known from other centralized IT approaches. To address them, a concept has been drafted that distinguishes between three classes of BI services which allow for a differentiated and staged approach to professionalization.

The concept is understood to be a first building block to an agile and IT supported BI development. It is conceived to be built on a defined service management concept and to be technically supported by a SOA architecture. In further steps, the presented results need to be fleshed out, coupled with concepts for BI service management and SOA and be validated by the use of prototypes.

Limitations of the study lie in the need for deeper and more solid statistic evidence, in the requirement for additional qualitative validation steps, and in the confinement of the exploration to the finance sector. Regarding the empirical data, the presented results are understood as a rather broad first step which is not meant to provide final and comprehensive results. However, together with the conclusions derived from the qualitative study they deliver some grip for the design of a much more sophisticated quantitative empirical instrument. Besides, additional interviews might help to gain further insights into possible applications areas, particular restrictions and additional options regarding the implementation of the concept.

The sectoral limitation should also be scrutinized. However, while the finance industry indeed brings a specific set of BI requirements (e.g. real time data feeds from trading systems) and unique BI applications (e.g. for risk management or data mining), the presented arguments are equally valid for other businesses as well. An exception is the need for very fast development cycles. But in times of complex global turbulences it can be expected that on a mid term time horizon this type of setting will become relevant for most other businesses as well – with the finance sector again becoming the forerunner in the domain of BI solutions.

References


Proceedings ECIS 2009


AN ONTOLOGY-DRIVEN TOPIC MAPPING APPROACH TO MULTI-LEVEL MANAGEMENT OF E-LEARNING RESOURCES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0315.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-learning, Information management, Knowledge Management Systems, Education</td>
</tr>
</tbody>
</table>
AN ONTOLOGY-DRIVEN TOPIC MAPPING APPROACH TO
MULTI-LEVEL MANAGEMENT OF E-LEARNING RESOURCES

Bhavani Sridharan, Hepu Deng, Brian Corbitt, School of Business Information Technology,
RMIT University, GPO Box 2476V, Victoria, 3000, Australia, {bhavani.sridharan;
hepu.deng; brian.corbitt}@rmit.edu.au

Abstract

An appropriate use of various pedagogical strategies is fundamental for the effective transfer of
knowledge in a flourishing e-learning environment. The resultant information superfluity, however,
needs to be tackled for developing sustainable e-learning. This necessitates an effective representation
and intelligent access to learning resources. Topic maps address these problems of representation and
retrieval of information in a distributed environment. The former aspect is particularly relevant where
the subject domain is complex and the later aspect is important where the amount of resources is
abundant but not easily accessible. Conversely, effective presentation of learning resources based on
various pedagogical strategies along with global capturing and authentication of learning resources
are an intrinsic part of effective management of learning resources. Towards fulfilling this objective,
this paper proposes a multi-level ontology-driven topic mapping approach to facilitate an effective
visualization, classification and global authoring of learning resources in e-learning.

Key words: Ontologies, Topic Maps, E-learning, Knowledge Management
AN ONTOLOGY-DRIVEN TOPIC MAPPING APPROACH TO MULTI-LEVEL MANAGEMENT OF E-LEARNING RESOURCES

1 INTRODUCTION

E-learning is “the delivery of educational content via electronic media” (Tastle, White & Shackleton, 2005) (p. 242). To improve the effectiveness of e-learning, the constructivist learning paradigm (Papert, 1990) proposes utilizing multiple student-centred learning strategies including collaborative learning, explorative learning, adaptive learning, and concept mapping for facilitating the smooth transferring of knowledge in knowledge construction. While these strategies do assist learners with their knowledge construction, they also create a problem of information overloading due to an enormous amount of learning resources generated (Vivek, Steven, Dennis, Gretchen, 2007). Such information overloading prevents learners from effectively wading through this superfluity of information in an e-learning environment. As a result, a well-organized representation of course concepts, learning resources, learning tools for facilitating the effective and efficient use of learning resources is essential for successful e-learning.

Numerous developments have emerged using ontologies and semantic web technologies for addressing these problems (Jin et al., 1999, Nejdl et al., 2002, Anderson & Whitelock, 2004, Sampson et al., 2004, Dicheva & Dichev, 2006, Gasevic & Hatala, 2006). These developments, however, have not yet been fully incorporated into the existing e-learning scenario due to the challenges involved in implementing them. Existing e-learning websites are still predominantly based on the first generation of learning management systems (LMS) such as blackboard, WebCT, Top class, Learning space, Virtual Laboratory, FirstClass, Moodle and others. These systems do not support for effective management of e-learning resources (Hatem, Ramadn & Neagu, 2005). The reason for this is that these systems are created for human and machine readability but not for computer understandability. To effectively manage learning resources, improving machine understandability and enabling intelligent information processing are critical.

Semantic web is a promising solution to enable intelligent web-based information processing (Berners-Lee, Handler & Lassila, 2006). Realizing the potentials of semantic web technologies in education, initiatives using semantic web technologies in e-learning started in late 90’s (Jin et al., 1999, Nejdl et al., 2002, Nilsson, Palmer & Brase, 2003). The major argument for this is that the availability of massive information is of no use, unless the ‘right information in the right context with the right level of details to the right person at the right time’ (Holger, 2003) is delivered. There are further developments towards addressing the problems of information representation, retrieval and reuse including topic mapping (Mizoguchi, 2004, Vatant, 2004, Dicheva et al., 2006) and resource description formatting (RDF) (Garshol, 2004, Yang et al., 2006).

Regardless of these developments in using topic mapping and ontologies in e-learning, very little or no evidence is presented to confront specific issues that online learners face. These critical issues include simple and flexible visualization (Dicheva, Dichev & Dandan, 2005b) and representation of e-learning resources, single point access to all learning resources (Dicheva et al., 2005c), and quality assured sustainable resources for reusability (Brase & Nejdl, 2003). The need to address these issues is further reinforced by the recognition of the stakeholders on the critical importance of a proactive use of diagrammatic representation of course content and effective management of learning resources for successful e-learning (Sridharan, Deng & Corbitt, 2008).

This paper proposes a novel approach for topic visualisation, representation and authoring features to fulfill specific e-learning requirements by extending our earlier work on the identification of critical success factors for sustainable e-learning (Sridharan et al., 2009). In particular, this paper introduces a new dimension to knowledge representation of learning resources based on pedagogies employing topic mapping paradigm. This helps learners create their mental models by choosing the learning path
based on their learning styles and preferences and allows them to follow a step-by-step process in their knowledge accumulation processes. Furthermore, the proposed approach enables a sustainable learning object repository, proactive contribution of learning resources with authentication facility.

The remainder of the paper is organized as follows. In Section 2, the research questions and significance of the research are covered. A description of the theoretical background followed by a brief review of related work is presented in Section 3. An ontology-driven topic mapping approach is presented in Section 4. Finally a summary with future research is included in Section 5.

2 RESEARCH QUESTIONS

A massive amount of learning resources is generated through various pedagogic strategies in the current e-learning environment. Some of these resources are generated using disparate learning tools such as discussion forums, e-mails, chat, open courseware materials, useful links etc. Such a rich collection of learning resources provides learners with numerous opportunities for reinforcing and strengthening their learning through effective knowledge construction. It, however, comes with various shortcomings including (a) ineffective use of available learning resources, (b) absence of facilities to capture valuable tacit and explicit knowledge generated, (c) scarcity of time and effort required by domain expert to capture and maintain these resources, and (d) deviation of learners from their mainstream learning activity while trying to interact with multiplicity of tools and technologies associated with an online course.

Against this background, the main question for this research is ‘how to realize effective transfer of knowledge through efficient capturing, eliciting, organizing, authenticating, retrieving and reusing learning resources in an e-learning environment?’. Following this question, several subsidiary questions can be defined as follows: (a) How to represent resources in a way that is easy and fast for learners to intuitively comprehend the course concepts? (b) How to provide a single point access to course resources and tools supporting various learning strategies? (c) How to represent the retrieval results in an organized manner based on pedagogic strategies? (d) How to create sustainable learning resources through global authoring of learning resources with authentication facilities?

To answer these questions, this paper proposes a multi-level course visualization and representation approach through considering an ontology-driven single point entry to e-learning resources through topic mapping. The proposed approach has the following specific features including (a) simple multi-level representation of course material both graphically and textually using topic mapping in a taxonomical manner with provision to expand and collapse based on the selection of a branch by the learner, (b) provision of one point access to both course related resources and tools associated with various learning strategies to prevent learners lost in information ‘space and technology’, (c) provision of easy and fast access to context-specific quality learning resources and tools through navigational topic maps or query-based topic maps, (d) presentation of retrieval results in a pedagogically classified style with freedom for learners to choose their own learning path, and (e) facilities for populating new resources and associated ontologies by participants, pending approval from domain experts using simple easy to use interface through topic maps-based graphic or text navigation.

The proposed approach mutually benefits both learners and teachers in many ways. Learners benefit through effective knowledge absorption through step-by-step course visualization, effective presentation of context specific annotated learning resources and freedom to choose the right resources for learning based on their level and style. A single-point access to all course-related resources prohibits learners from being side tracked from the mainstream activity. Learners can derive a sense of satisfaction to see their contribution appreciated and added to the authenticated learning object repository. Maintenance and sustainability of the learning object repository are made easy as this approach facilitates global contribution with quality check. This ensures that domain experts can focus their efforts towards authenticating rather than populating the repositories for reuse.
3 RELATED STUDY

The constructivist learning paradigm promotes the use of intuitively visual interfaces such as concept mapping (Novak, 1998) and topic mapping (Pepper, 2000) for effective learning. Visual representations are powerful mechanisms for materializing meaningful learning process (Gershon & Eick, 1995, Le Grand & Soto, 2000) in learning environments. Concept mapping is a “technique for representing knowledge in diagrams called ‘knowledge graphs’, which are networks of concepts with relationship between them represented respectively by nodes and links” (Novak, 1998). Concept maps are invented with the objective of externalising the understanding of a domain by learners and experts.

Topic maps represent a collection of topics, their relationships and their information sources. Topic mapping provides a mechanism for intuitively representing course concepts and their relationships between the concepts. In addition, to enable fast and easy retrieval of resources, topic map facilitate merging of electronic indexes similar to back of book index. These indexes are however, derived from multiple sources and anchored to actual resources, but kept separately.

Three key concepts are involved in topic mapping including topics, associations and occurrences (Pepper, 2000). Topics representing concepts are the syntactic constructs processable by machine. Relationships between concepts are determined by the associations with grouping of concepts without any implied direction. Occurrences are a way to binding topics to relevant resources. Such occurrence can be a resource reference such as the uniform resource locator, pdf, doc, audio, video or a string value such as population in a country. Figure 1 gives an example of a topic map in a database course.

![Figure 1. Example of Topic Map with topic, association and occurrence](image-url)

Knowledge management is emerged to enhance the organizational performance through the synergy of people, process and technology and to address the problems of information overload. It is commonly referred to as a systematic process of acquiring, eliciting, organizing, representing and retrieving organizational information and knowledge (Nonaka, 1998, Duffy, 2001). In the current e-learning scenario, learners are often confronted with more learning resources than they can deal with or handle using various pedagogic strategies. To tackle this problem of information overload, embracing the principles of knowledge management in conjunction with the semantic web technology is imperative.
for successful management of e-learning resources. The semantic web is an extension of the current World Wide Web, which envisions intelligent information processing (Berners-Lee, Hendler & Lassila, 2001).

Ontologies are “an explicit specification of a conceptualization” (Gruber, 1995) (p.908). It is the “vocabulary for expressing the entities and relationships of a conceptual model in a domain along with semantic constraints” (Obrst and Liu, 2003). Ontologies are “the metadata schema providing a controlled vocabulary of concepts” (Maedche et al., 2002). Ontology is a document that formally defines the relationship among terms (Berners-Lee et al., 2006). It provides a mechanism to communicate between people and computers through a shared understanding of resources in a domain (Davies and Fensel, 2002).

There are two popular technologies for ontology construction including RDF and topic mapping. Topic mapping and RDF technologies, however, are developed with slightly different flavours for fulfilling similar objectives. Garshol (2004) asserts, “topic maps were created to support high level indexing of sets information resources to make the information in them finable. RDF, on the other hand, is intended to support the vision of semantic web through providing structured metadata about resources and a foundation for logical inferencing”. Both technologies aim at realising the semantic web vision of machine processability by annotating, associating between concepts and attaching more semantics. Topics maps are an ISO (International organisation for standardization) standard developed in 2001. RDF is a W3C (world wide web Consortium) standard developed in 1997. Web ontology language (OWL) built on RDF is the new W3C recommendation for ontology construction with facilitates for effective reasoning capabilities by consistency checking through inference rules such as transitivity, symmetry etc.

Much research in topic mapping has been done in knowledge visualization (Le Grand & Soto, 2001, Dicheva et al., 2006), knowledge retrieval (Yang, 2003; Garshol, 2004; Shin et al., 2006), knowledge capturing (Dicheva and Dichev, 2006), and managing and reusing repositories (Ahmed, 2000; Dicheva et al, 2005). Several interactive graphic navigating tools and techniques for visualization through topic mapping have been developed (Mondeca, 2001, Fluit, Sabour & van Harmelen, 2002, Alani, 2003, Dicheva et al., 2005b). For instance, Topic Maps for Learning (TM4L) (Dicheva et al, 2006) provides support for graph view, text view and tree view of learning repositories. Ontopia navigator (2001), Mondeca’s topic navigator (2001), techuila’s (2001) Topic Maps for Java (TM4J) dynamic visualization (2001), and Alani’s (2003) visualization tool (TGVizTab) provide variations to existing knowledge visualization methods. Addressing the readability of a graph in a complex domain, Munzner (1997) proposes hyperbolic geometry to display large numbers of nodes in the screen. Le grand and Soto (2000) propose an interactive topic mapping visualization tool to represent semantic graphs in three dimension spaces. Nevertheless, automatic generation of graphic visualization using these tools becomes rather cumbersome and incomprehensible for a novice learner, especially in a complex course domain. In this regard, Le grand (2001) recommends topic mapping metrics to deduce efficient positioning of node to enable quick scrutiny of topic and explore the relevant ones in detail. Fluit (2002) proposes grouping instances in clusters according to their classes, while Ahmed (2000) suggests different layers of details in a topic map to display general information and to explore into details once the specificities are identified. The proposed approach adds a new dimension to knowledge representation by including topic maps based presentation of learning tools and resources based on the pedagogic classification.

Yang (2003), Shin (2006), and Vivek et al. (2007) use topic maps to enhance the retrieval of learning resources for overcoming the limitations of key-word based search engines. Yang (2003) presents mechanisms to improve the retrieval results from collaborative learning repositories in semantic retrieval of concept model (SRCM). Shin (2006) proposes conversion of topic maps and RDF metadata for effective retrieval online. Vivek et al. (2007) investigate ways to improve information retrieval capabilities in e-learning. Nevertheless, presentation of retrieved results is equally important for learners to get the best out of the resources through effective classification of retrieved results.
Dicheva (2006) builds an ontology-driven learning repository using topic mapping in TM4L ontology editor, whereas Yang et al. (2006) propose a topic map-driven visual authoring tool (XRVAT) for constructing ontologies. Ahmed (2002) examines the potentials of topic maps as an interface to multi-user document repository. Realising the problems of identification of topics and relationships between them, Dicheva et al. (2005a) highlight the importance of identification of minimal ontology using topic maps. However, facilitating global authoring with quality check is crucial aspect for creating a sustainable learning object repository, which is not given sufficient thought in the literature.


It is evident from the literature that a concerted effort is directed towards tackling the problem of information overload in managing e-learning resources, employing the next generation of web. Visualisation, retrieval and authoring of resources and associated ontologies using topic map is a significant development towards effective management of resources in a distributed environment. Nonetheless, generation of topic visualisation through these approaches become very cumbersome and incomprehensible, especially for a novice learner seeking to acquire knowledge in a complex domain. Besides, not much attention is paid to the representation of retrieved resources generated from multiple pedagogic strategies, which is equally important. Furthermore, global authoring, with authentication facilities, is very critical to create a sustainable learning object repository by capturing both tacit and explicit knowledge for reusability. Finally, single point access to not only to learning resource, but also learning tools will enhance the effectiveness of learning by enabling the learner to focus on the learning activity. Towards filling these gaps, this paper extends the existing research for tackling these specific aspects in an e-learning context. These aspects are simple multi-level presentation and representation of retrieved results through visual graphs, authenticated global authoring of learning resources and single point access to all learning resources including tools associated with various pedagogies.

4 AN ONTOLOGY-DRIVEN TOPIC MAPPING APPROACH

4.1 Motivation

The support for this study comes from our earlier work on evaluating the critical success factors for sustainable learning (Sridharan et al., 2009). This qualitative study entailed interviewing 29 academic staff involved in the e-learning domain from a wide range of function areas including the strategy and policy division, the library resource management division, the technology and media division and, the teaching experts in an e-learning environment. The objective of this study is to identify the critical success factors intertwined within pedagogies, technologies and management of learning resources in an e-learning environment. The critical success factors identified in the study include appropriate choice of pedagogies, proactive use of technologies supporting pedagogies, understanding the pedagogical principles behind technologies, appreciating and exploiting the full potentials of various technologies supporting pedagogies, effective management of learning resources through selected metadata and ontologies and, more importantly synergising pedagogies, technologies and management of learning resources for sustainable e-learning. However, it was apparent, that there are some
practical difficulties which prevent teaching staff from incorporating these factors in reality for enhancing the e-learning effectiveness. Some of the prohibiting factors identified in the study include ineffectiveness of LMS, lack of time and effort, lack of interest, problems associated with knowledge sharing etc. in proactively employing the identified critical success factors in reality. This paper addresses the specific problems related to proactive use of concept mapping technique as a pedagogical strategy, associated technologies and effective management of learning resources through ontology-based topic map approach.

In this direction, this research aims to propose a novel approach for extending the topic visualisation, representation and authoring features to fulfil specific e-learning requirements. The proposed approach consists of five modules to facilitate effective management of learning resources, namely student module, domain expert module, ontology module, authentication module and learning resources module. To comprehend the process involved in each of these modules, a motivating scenario is given in the following with a detailed explanation of functionalities of each of the modules.

4.2 Typical Scenario

To illustrate the proposed ontology-driven topic mapping approach, we take a typical learning scenario, where the learner is required to accomplish a task for fulfilling the course objective. Let’s take an example where the learner is expected to finish an assignment on normalization in a database course before a short deadline which is approaching. The learner is required to gather as much knowledge as possible to complete the assignment before attempting or posing any queries or doubts related to assignment. In the current e-learning set-up with LMS, the learner has to wade through multiple sources, multiple paths, and multiple modes to accomplish the task. In this scenario, despite spending a lot of time and effort, it is highly likely that learner either got side-tracked or missed some important aspects due to ineffective organisation and presentation of learning resources.

4.3 Student Module

Student module facilitates a systematic presentation of topics, resources and tools in an organised manner using topic maps to enable step-by-step knowledge acquisition process for fulfilling the task. This module contains facilities for viewing the entire course in multi-levels with more details in a particular sub-topic shown on demand by the learner. Basic association types namely pre-requisite, co-requisite and follow-up concepts are represented. Learners have a choice of navigating through textual and graphical modes. Both text view and graphic view would contain the semantic overall view of the course to construe the main concepts represented by nodes in a course at a glance. Each node will have four options namely expansion, view resources, search, append within the chosen sub-topic.

By choosing the expansion option, the chosen branch will expand. The next level of details will be provided in both text and graph view. Alternatively choosing view resources option will generate a page with three frames as shown in Figure 2. The top frame of the page will contain a topic map of classified resources (such as basic learning material, collaborative learning resources, interactive learning resources, concept maps creation facilities etc.). The middle frame contains an annotated version of all relevant resources related to the chosen node from the learning object repository enabling the learner to decide their learning path. The bottom frame will contain all related tools associated with the chosen pedagogic strategy. The generated result will contain aggregated resources of all the topics based on the semantic relationship between topics represented by ontologies in the chosen node. Search option will generate the same result with classified representation of resources using both the key-word and semantic relationships from ontology structure. In addition, if learners find the resources presented to them have not clarified their understanding, they can access the query tool based on a simple web interface from the same knowledge representation sub-section to pose queries. This will automatically be directed to domain experts and peers without the need to go outside the learning space. On the other hand, if learners either found a very good learning resource or have...
some queries relating to the chosen node, they can populate the learning object through a similar interface along with metadata and ontologies.

This module is linked with the ontology module and the learning object repository module for facilitating effective retrieval of context-specific learning resources. In this approach, the course structure is pre-determined and remains relatively static with the objective of keeping it simple and comprehensible to a novice learner. The ontologies and learning resources at the back end is dynamic with facilities for uploading new resources with associated ontologies. Considering that topic maps are powerful for visual representation and navigation capabilities, the use of XML (extended make-up language) based topic maps (XTM) technology is proposed for this module deriving basic knowledge structure from ontologies.

Figure 2. A multi-level learning resources visualization and representation

In the given scenario, the learner chooses the relevant node namely the normalization concept using this approach and then selects the option of viewing aggregated results or going into further details. By clicking on the view option, the system presents a categorized list of all resources on normalization and the supporting tools. The presented resource result is classified into basic resources, resources from collaborative learning such as frequently asked question, interactive learning resources such as self-test quiz and problem-based learning resources, explorative learning resources (such as open course ware material, additional reading material, useful links with annotation), tools for constructing mental maps such as concept map tools and tools for practicing problems (intelligent tutoring systems) or self-testing their knowledge. The learner can either choose aggregated result to show all resources in the chosen node and choose a specific category based on the stage of learning and preferences and styles of learning. The chosen resources will be rendered in a new page in detail with facilities for tracking of completed sections.

4.4 Domain Expert Module

The domain expert module consists of simple web 2.0 based friendly interfaces for manually creating, updating, deleting topic maps along with learning resources and ontologies. The same interface is used for posting queries and answers to queries, which will automatically be added to the learning repositories after quality check by domain experts. Apart from this domain experts can create some inference rules to create automatic authentication of learning resources either based on the author, source and other selected key elements. This module is constrained by ontology module and authentication module.
4.5 Ontology Module

The ontology module is at the back end of the system anchored to both the visualization module and the learning resources module. This module comprises of four types of ontologies such as context ontologies, structure ontologies, domain ontologies and pedagogy ontologies. When a node is chosen or a keyword-based query is entered, the ontology base is activated to get additional information for both retrieving and classification of resources before presenting it to the learner. Using Protégé OWL, the representation of disjoint classes to assert a concept can not be instance of more than one class, consistency checking through a reasoner, and adding value restrictions can be effectively executed. Also, OWL is powerful for inference due to its expressivity and stricter rules such as transitivity, symmetry, functional properties etc. Due to these reasons, use of OWL is proposed for this module. A sample ontology construction using Protégé Owl is given in Figure 3. All the complex relationships between concepts and inference rules are included but concealed to the learner. This complex semantic relationship between concepts enables effective retrieval based on the semantically related concepts and representation of organized results.

![Figure 3. A sample construction of ontologies using Protégé OWL](image)

4.6 Authentication Module

The authentication module is anchored to the domain expert module and the learning resource module. It facilitates either an automatic quality check based on the inference rules schema created by experts or sending alerts for domain experts to check for quality before committing the resource in the learning object repository.

4.7 Learning Resources Module

Finally the learning resources module contains heterogeneous learning resources which are anchored to student module, domain expert module, ontologies module, and authentication module. Use of
XTM technology is recommended to represent this module. All relevant quality learning objects are included, irrespective of the type and source of resource. The resources could be an audio, video, outside link, discussion material, question and answer, example, definition etc.

5 CONCLUSION AND LIMITATIONS

Having built on our earlier work in evaluating the critical success factors for sustainable e-learning (Sridharan et al., 2009), we in this paper have introduced a novel approach to overcome some of the practical difficulties in the process of embedding some of the critical success factors identified in our earlier research for successful e-learning. Specifically, we have addressed the problems associated with using concept mapping techniques, effective representation of learning resources, sustainable and reusable management of learning resources through ontology-driven topic maps in e-learning websites. This leads to some key contributions to the domain of e-learning including multi-level topic visualisation, resource representation using topic mapping, and authenticated global authoring of learning resources with critical ontologies. Furthermore, we have proposed a novel approach that is capable of recommending a single point access to not only learning objects, but also course related tools using a single interface, wherever applicable, to avoid learners losing track of their mainstream activities.

In this paper we have proposed to blending topic maps and OWL for presenting the course in a simple and comprehensive manner through topic mapping and concealing the complexities of semantic relationships in a course behind OWL ontologies. To fulfil the objective of effective presentation of course structure in a simple and comprehensible manner, a multi-level presentation of course through topic maps is pursued. To help learners grasp the basic knowledge before exploring the details of a course, high level overview of domain knowledge is proposed with expansion and collapsing facilities. To effectively represent retrieval of learning resources, this paper proposes a topic mapping based classification of resources based on pedagogies used in a course incorporating semantic annotations. To enable successful and sustainable management of learning resources, this approach proposes mechanisms for populating new resources and associated ontologies within the topic-maps based framework pending quality check by experts.

The study is limited in a number of ways. Firstly, this study is based on the outcome from a qualitative study based on a small sample size. As a consequence, generalization across the learning community is not possible. Secondly, evidence of support for such an approach is one sided as the learning community, other key stakeholders in e-learning environment, are absent. Thirdly, the success of this approach requires the identification of critical metadata and ontologies for creating a sustainable learning object repository, which has not been identified. Finally, various options and mechanisms for either developing a stand alone system or integrating with LMS and content management have not yet been explored.

Future direction for this research requires more in-depth study incorporating learners’ view is critical for wider acceptance of this approach. In addition, identification of critical ontologies from teachers and learners perspectives to effectively implement a system-based on topic maps is crucial. Based on the results, development of a prototype with functionalities for presentation, representation, retrieval, authentication, inference and populating mechanisms is necessary to implement in reality. In addition, identification of mechanisms for keeping track of individual learners’ progress by representing completed topics in one colour and yet to complete in a different colour within the topic map paradigm will enhance the use of the tool.

References


Exploring The Development Of Social Alignment Within An Innovation Context

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0506.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Models, Information exchange, Shared knowledge, IT innovation</td>
</tr>
</tbody>
</table>
EXPLORING THE DEVELOPMENT OF SOCIAL ALIGNMENT WITHIN AN INNOVATION CONTEXT

Tadhg Nagle, University College Cork, Ireland
William Golden, National University of Ireland, Galway

Abstract
There has been increasing attention in IS literature around the technology-business relationship and the concept of social alignment. However, little attention has been given to how or why social alignment should be developed. Further examining the concept this paper explores the development of social alignment within an innovation context. Using canonical action research the study is based within an innovation network. Primarily focusing on four organisations, a lack of social alignment was diagnosed. In particular, it was found the business executives did not fully understand the potential of the new technologies available. An intervention was then planned to facilitate the organisations in utilising a business model framework for developing social alignment by creating a common understanding between the business-technology executives. The study which took 24 months to complete involved over 110 hours of primary data. The results highlight the positive impact of the business model framework in developing social alignment. However, it was found that the business model framework alone is not enough to develop social alignment within an innovation setting. Careful thought should be put into who facilitates the development of social alignment and a shared understanding, as CTO’s are often too caught up in day-to-day operations to fulfill the role.

Keywords: Social Alignment, Business Models, Shared Understanding

1 INTRODUCTION

The importance of alignment has been frequently noted (Avison et al., 2004). So much so, it has been viewed as a core component to the theory and practice of IS strategy (Galliers & Newell, 2003). Using a framework set out by Horovitz (1984) for the process of strategic business planning, Reich and Benbasat (1996) converged alignment research into two dimensions of business and IS/IT alignment: (i) intellectual and (ii) social. The intellectual dimension, which was developed from the work of Henderson and Venkatraman (1989) refers to the existence of a high-quality set of interrelated business and IS/IT plans (Reich & Benbasat, 1996; Tan & Gallupe, 2006). However, given that the majority of research is positioned within the intellectual dimension (Tan & Gallupe, 2006), it is evident that the lack of senior executives in the formulation of IS planning and the presence of internal inhibitors to use IS strategically, demonstrates the deficiency with this approach to IS/IT alignment (Kearns & Lederer, 2000). To aid understanding on how to overcome this problem, this paper focuses on the social side of alignment, which has been highlighted as a potential solution (Feeny et al., 1992).

This paper begins by defining the concept of social alignment and subsequently proceeds to highlight the compatibility of the business model as a tool for creating a shared understanding; thus developing social alignment. Furthermore, the paper describes the business model concept as defined by Osterwalder et al. (2005) and in particular notes the untested proposition by the authors, which puts forward the business model framework as a platform for common communication and shared understanding. Justifying the use of canonical action research the paper proceeds to highlight the suitability and relevance of the Industry Led Research Project (ILRP) that bounds the innovation network within which the study is based. In addition, along with the description of the implemented action/intervention, the findings collected from the diagnosis and evaluation stages of the research
methodology are outlined and analysed. Finally the findings are discussed highlighting the main
conclusions from the study.

2 THEORETICAL GROUNDING

2.1 Business Models

The use of the term ‘business models’ within academic literature became particularly noticeable
during the dotcom era (Osterwalder et al., 2005). This association has been explained by the business
model concept taking a central role in describing how organisations should compete in the digital
economy (Lee, 2001; Seddon et al., 2004). This link is further supported by views that Internet
technologies have challenged conventional methods of value creation and also generated a wide
knowledge gap between IT developers and organisational stakeholders (Schmid, 2001; Gordijn &
Akkermans, 2001). However, persistent confusion over the domain during the period earned it the
label of being the “most discussed and least understood aspect on the web” (Rappa, 2001).

Since then a strong vein of research has developed around using business models in communicating
the “core logic”, “business system” or value creation within an organisation (Linder & Cantrell, 2000;
Petrovic et al., 2001; Auer & Follack, 2002). More process-oriented definitions declare that a business
model is a “story that explains how an organisation works” (Magretta, 2002) and the first step in
gathering requirements for business information systems (Gordijn & Akkermans, 2001). In addition to
these broad comprehensive descriptions, there has been an accumulation of business model definitions
detailing its primary components and possible interrelationships (Pateli & Giaglis, 2003). One such
eample is the highly recognisable definition given by Timmers (1998), stating that a business model is
“an architecture for the product, service and information flows, including a description of the
various business actors and their roles; and a description of the potential benefits for the various
actors; and description of the sources of revenues”.

However, like the definition stated by Timmers (1998), many descriptions of a business model list
numerous components which has nullified their attempts to clear confusion and inversely added to the
ambiguity around the domain (Pateli & Giaglis, 2003). As a result, more recent approaches at creating
a universal definition of a business model have incorporated a methodology of synthesising large
quantities of past research. In an effort to find the most common components in a business model, two
individual studies create a list of components using the criterion that the component had to be
mentioned by at least two authors (Osterwalder et al., 2005; Shafer et al., 2005). Between the two
studies the framework put forward by Osterwalder et al. (2005) has been more widely accepted by the
practitioner community. Summarising the framework, Table 1 highlights the nine components, which
are grouped by Product, Customer Interface, Infrastructure Management and Financial Aspects.

<table>
<thead>
<tr>
<th>Grouping</th>
<th>Business Model Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product</td>
<td><strong>Value Proposition</strong> – overall view of an organisation’s bundle of products and services</td>
</tr>
<tr>
<td>Customer Interface</td>
<td><strong>Target Customer</strong> – segments of customers an organisation wants to offer value</td>
</tr>
<tr>
<td></td>
<td><strong>Distribution Channel</strong> – various means of an organisation to get in touch with its customers</td>
</tr>
<tr>
<td></td>
<td><strong>Relationship</strong> – kind of links an organisation establishes between itself and its different customer segments.</td>
</tr>
<tr>
<td>Infrastructure Management</td>
<td><strong>Value Configuration</strong> – arrangement of activities and resources</td>
</tr>
<tr>
<td></td>
<td><strong>Core Competence</strong> – the competencies needed to execute the business model</td>
</tr>
<tr>
<td></td>
<td><strong>Partner Network</strong> – network of cooperative agreements with other organisations.</td>
</tr>
<tr>
<td>Financial Aspects</td>
<td><strong>Cost Structure</strong> – monetary consequences of employing a business model</td>
</tr>
<tr>
<td></td>
<td><strong>Revenue Model</strong> – revenue flows through which an organisation makes money.</td>
</tr>
</tbody>
</table>

Table 1: List of Business model components (Osterwalder et al., 2005)
2.2 Social Alignment and Shared Understanding

Evidence has shown that while technology executives view both alignment of business with technology and alignment of technology with business as being very important, non-IS executives only deem the alignment of IS with business planning as being important in gaining competitive advantage (Kearns & Lederer, 2000; Byrd et al., 2006). This is a serious misjudgement, as both types of alignment ensure that senior executives from business and technology give greater commitment to strategic IS resources and gain a strong understanding of each other domains. This results in more effective use of IS resources towards improving organisational performance (Reich & Benbasat, 1996; Kearns & Lederer, 2000). Thus, in order to develop and maintain a mutual/shared understanding between business-IS executives, research efforts began to focus the social side of alignment in the early 1990’s (Feeny et al., 1992).

Social alignment is defined as “the level of mutual understanding and commitment to the business and IT mission, objectives, and plans” (Reich & Benbasat, 1996). Further developing their research framework, Reich and Benbasat (2000) divide the social alignment dimension into short-term alignment and long-term alignment. Short-term alignment is defined as the “state in which business and IT executives understand and are committed to each other’s short-term plans and objectives”. The key factor in determining short-term alignment is the shared knowledge between business and IT executives. More recently, researchers have used the terms convergence, mutual understanding, shared understanding, and shared cognition synonymously to indicated a state where a common ground of comprehension is reached on a certain topic (Johnson & Lederer, 2005). Moreover, with similar emphasis on a shared understanding, long-term social alignment is defined as the “state in which businesses and IT executives share a common vision of the ways in which IT will contribute to the success of the business unit” (Reich & Benbasat, 2000).

With respect to this paper, a number of studies have shown that strong social alignment yields multiple benefits for innovation. Firstly, research has shown strong social alignment has a significantly positive effect on innovation (Jansen et al., 2006). In particular Landry et al. (2002) highlighted that marginal increases in social alignment contribute more than any other explanatory variable in increasing the likelihood of innovation in organisations. In addition, Jansen et al. (2006) discovered that managers might develop “densely” connected social relations within their organisations to increase its ambidexterity and found that informal coordination mechanisms are more important than formal coordination mechanism for developing both exploratory and exploitative innovation.

Moreover, the findings of Subramaniam and Youndt (2005), found that an organisations social alignment enables its ability to develop incremental and radical innovations. In fact, without good social alignment the knowledge and skill held by individuals had a noted negative effect on radical innovation. Practical implications discourage the employment of highly independent experts who are reluctant to share their ideas with their colleagues, as this is counterproductive. In addition, as specialists within a function tend to have similar backgrounds, experiences and career path opportunities, they usually have a specialised language and perspective that may not be easily understood by members outside their department or research unit (Hutt et al., 1995). However, instead of eliminating specialisation that is necessary for focusing on different elements of the technology or market (Dougherty, 1992), innovation management must build communication bridges or shared knowledge between different departments. In a study done by Mom et al. (2007) it was found that information flows from managerial peers positively related to their exploratory activities. From this, numerous unique views of technology can come together to form a vision of how the technology can meet market requirements (Bond & Houston, 2003). Nonetheless, a paucity of such tools that enables clear and shared communication around innovation projects has been highlighted (Christensen et al., 2008).
One of the major aims of alignment methodologies is to get IT and business managers to the same table (Booth & Philip, 2005). This may increase social alignment by increasing the communication between the different departments, however, recent studies have shown that these communication links need to be frequent to create a common understanding within an organisation (Johnson & Lederer, 2005). Furthermore, it has been known that frequent communication alone is not enough. Using a strategic management quote, “one can brief a reluctant manager endlessly without accomplishing anything, unless one comes to realize his hidden resistances and strives to bring them up to consciousness in some way” (Churchman & Schainblatt, 1965). Heretofore, there is a need for a common language of words or symbols that create a positive social influence as different environments, most often than not, use different languages (Reich & Benbasat, 1996). Once this is achieved a synergy of shared knowledge is created (Nelson & Cooprider, 1996), overcoming the problem of having executives on different levels of understanding (Reich & Benbasat, 1996; Bassellier & Benbasat, 2004). Unfortunately, outside the practical suggestions of increased collective encounters (Tan & Gallupe, 2006; Feeny et al., 1992) or formal reporting systems (Preston et al., 2006) there is a scarcity of research in examining the mechanisms and process through which social alignment can be achieved (Cohen & Toleman, 2006).

However, linking the use of business models in developing social alignment, Osterwalder et al. (2005) specifically posits “the business model concept helps increase the mutual understanding between the business and IT/IS domain. It creates a common language and shared comprehension”. However, the Osterwalder et al. (2005) paper is purely theoretical and even though the framework has been widely adopted by the practitioner community, there has been no rigorous validation of any of propositions put forward in their paper. As a result this paper focuses on:

**RQ: Investigating the effectiveness of a business model framework in developing social alignment, within an innovation context?**

3 RESEARCH METHOD

Analysing the different research methodologies, action research stands out as a strong choice as it not only merges theory and praxis but also uniquely “associates research and practice, so research informs practice and practice informs research synergistically” (Avison et al., 1999). In summary “action research strives to marry rigor to relevance by conducting scientific research in the setting of a real world problem”. Taking this into account, action research has been deemed as being very suitable for the IS domain as: (i) IS is a highly applied field “almost vocational in nature”, (ii) action research is highly clinical and puts the researcher in a helping role, (iii) the aim of action research is to produce highly relevant findings which is an important factor in the significance of IS research (Baskerville & Wood-Harper, 1998). Throughout the twelve different types, canonical action research is the most widely accepted form of action research (Kock, 2004). Using the Susman and Evered (1978) five step model of: (i) diagnosis, (ii) action planning, (iii) action taking (intervention), (iv) evaluation, and (v) reflection, the researcher collected data over a 24 month period.

Another reason why action research was chosen is that the researchers were externally involved in the Industry Led Research Project (ILRP) within which the study is based. This project enabled an existing network of primarily four eLearning firms to gain access to substantial R&D resources. Ranging in size from 20 to 150 employees, the organisations had individual revenues of up to €15 million around that period. In addition, all of the organisations are primarily involved in the industry (compliance and regulation eLearning sector) on a European and/or global basis and are in existence between 10 and 20 years. Their customers primarily involve the pharmaceutical, high-tech, financial and hospitality industries. Starting in 2006 the ILRP partnered the organisations with a research institution, which specialised in Web 2.0 and semantic technologies. As a member of the institution the researcher was able to collect primary data for over 24 months of the project, in which the eLearning organisations faced the technological disruption of Web 2.0. Summarised in Table 2, over 110 hours of primary data was collected from multiple sources and multiple informants using
interviews and participant observation. In order to answer the research question the action taking (intervention) was carried out over six months and involved the researcher demonstrating how the business model could be used as a communication platform. This was done at a number of open days but also on a one-to-one basis with the organisations. In total, 17 semi-structured interviews were undertaken, lasting between 45 – 90 minutes. Participant observation was used throughout the 24 months in gathering data from inter-organisational meetings and open days. Even though the meetings were not as formalised as interviews the researcher was still able to collect data by asking questions or noting points of information that complied with the questions asked in previous interviews. Furthermore, on more than one occasion meetings were conducted over a full business day. Open days also consisted of one to two day events where all of the organisations in the innovation network were invited to one location to discuss current issues and topics associated with the network. Finally, analysis of company/industry reports and press releases that applied to period and organisations in question were also used to triangulate data used in the study.

Table 2: Summary of data collection during the study.

<table>
<thead>
<tr>
<th>Company</th>
<th>Primary Data</th>
<th>Secondary Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Interviews</td>
<td>Meetings</td>
</tr>
<tr>
<td>Company A</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Company B</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Company C</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Company D</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Research Institution</td>
<td>2</td>
<td>N/A</td>
</tr>
<tr>
<td>Enterprise Ireland</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Total (Total in hrs)</td>
<td>17 (20 hrs)</td>
<td>17 (39 hrs)</td>
</tr>
</tbody>
</table>

4 FINDINGS

Primarily focusing on the data gathered in the diagnosis and evaluation stages of the action research methodology this section highlights the findings of the study. In particular, the cause and effect of poor social alignment is outlined in the diagnosis section. Furthermore, the use of the business model as a communication platform is evaluated with regard to its impact on social alignment between the business and technology domains of the organisations.

4.1 Diagnosis

While collecting data during the diagnosis stage of the project, it became apparent that the participation of the eLearning organisations was significantly lacking. When the project started it was assumed that the eLearning organisations would become actively involved in developing the innovative technologies researched within institute. However, this was not the case. To encourage participation from the organisations open days were conducted. These consisted of presentations and demos of the research being accomplished. However, after they had taken place there was no observed increase in participation between the eLearning organisations and the institute. Such was the lack of participation, the Programme Director (Enterprise Ireland) felt that after the ten months of the project, the likelihood of any of eLearning organisations commercialising any of the research was exactly same as when the project started.

Analysing why the open days did not work it was identified that all the attendees to the open days were all people involved in the technology side of the organisations. In particular, it was highlighted by the Director of Informatics (Enterprise Ireland) who had the responsibility of overseeing 3-5 similar ILRP’s. With in depth experience of such initiatives the Director of Informatics pointed out that “the typical person that shows up to the open days 2-3 times a year, tends to be the technical people from the companies. That’s all well and good and they maybe getting value from it. But at CEO level, in order to keep them supporting their techies ..... it’s really important and useful for them to see this is
how it (technology research) might change how they do things.” This in turn “allows the technical guys the space and the time for them to stay involved in the project”. This clearly shows that there was a breakdown in communication between the technology and business sides of the organisations. Furthermore, it demonstrates that the business executives did not have enough information to fully grasp how the technological research could potentially impact on their business. As a result they had no reason to commit resources to the innovation initiative.

Originally, the fact that the participants were technologists was not seen as a problem, as it was assumed that they were in the best position to communicate the opportunities back to their counterparts or CEO of their organisation. However, the early open days were dominated by discussions on deep technological topics such as: SCORM (most widespread eLearning standard) compliance, privacy, security and the incompatibility of the new compared to those used in industry. Moreover, the majority of these discussions focused on the immediate obstacles rather than exploring potential opportunities through the new technologies. Furthermore, with all the discussions of the open days focusing on technological issues, very little business issues were discussed. In addition, it was found that even though the attendees found the open days useful when they reported back to their respective managers they got little or no response. Highlighting this fact, one of the regular attendees stated “the open days have been great but when we go back and report to our managers, they tend to glaze over it” (R&D Engineer, Company A). This is not surprising as the report would have consisted mainly of technological issues with very little information of the business opportunities being afforded to them by the project. In a sense, the attendees communicated the progress of the project in a totally different language and with information that was not deemed relevant.

Further analysing the communication methods used in the eLearning organisations uncovered a high emphasis on financial analysis around strategic activities and exploring new technologies. As noted by the Director of Informatics (Enterprise Ireland), when creating a vision for the organisation, they usually put it in “financial terms”. In addition, as noted by the CTO of Company C, communicating the benefit of exploring or investing in new technologies through financial models is often not possible. He stated, “imagine how hard that is to sell to the CEO”, when explaining a case when he needed to invest in research activities without being able to calculate the financial terms of the activity. In fact there is evidence that each of the CEO’s communicated strategies in financial terms. From the perspective of Company A, their strategy was to become a $100 million company. In order to fulfil this vision the CTO saw his role as making the technology as efficient as possible. As stated by the CTO (Company A), “the technology part of that (strategy) will not be too different from how we do it today…..do more of the same better faster”. Furthermore, the CTO clearly noted that “as long we have a business case to hang around it (technology investment) the CEO isn’t too concerned about what we do”. Company D pursued the strategy of monetising any effort across their clients through their services model. This was evident in a meeting with the CTO as he said he would only be able to participate in the project if he was able to demonstrate an “aggressive return”. In addition, Company B pursued a compound growth rate of 35% and similar to Company D would participate if substantial gains could be financially identified.

4.1.1 Summary of Diagnosis

Summarising the data gathered, it was found that the organisations had a communication breakdown that was impacting on their ability to create a shared understanding and shared vision. In an effort to overcome this obstacle a planned intervention to use the business model concept as defined by Osterwalder et al. (2005) was initiated. Consequently, it was predicted that with the business model as a communication platform the technology side would be better able to communicate the opportunities of the project with their business counterparts. In particular, it would allow the CTO’s better describe the potential of the new technologies being developed within the institute but also better demonstrate the potential impact of not developing the technologies. This would further enable the CEO’s better understand the business potential of the technologies and also get the CTO’s to explore aspects of the
new technologies beyond that of immediate technological obstacles. This would in-turn, increase participation in the project and allow the technology to explore new possibilities within the project.

4.2 Evaluation of Intervention

From an overall perspective the intervention and use of the business model was a success. The proceeding open day after the intervention showed a marked increase in participation by all the organisations. Furthermore each of the four firms committed to future participation. Company B wanted to get more involved in mobile devices and creating a future technology roadmap around that area. Company D brought one of their customers to the open day and noted very high interest in commercialising some of the technologies in the institute. Finally, Company A expressed keen interest in the area of new forms of learning assessment. The increased interest from the eLearning organisations was also noted by members within the institute. It was the first time that they felt the organisations had got actively involved in the project. From Enterprise Ireland’s perspective they were also satisfied with the results. the Programme Director reflected on the chances of some research being commercialised, and said “Not that I thought it was impossible but when we started off it was a ninety-one shot and now its still not passed the fifty-fifty mark but is more sixty-forty. But there was a chance that it would have never got off the ground and stayed at nine-to-one”. Also the Programme Director directly stated that we had “achieved our first objective…now there is more than just an awareness (by the e-Learning organisations) of whats going on”. However, he did also note that the eLearning organisations “were still in that firefighting mode … they are still driving their current business and they are all succeeding in that but that’s not giving them any bit of time to look that bit further down the road”.

Having noted the evidence of success it was important to attribute that success back to the intervention. During a meeting with the Programme Director (Enterprise Ireland) he discussed the reasons he felt were the cause for the increase in participation in the project. Firstly, the Programme Director felt that the business model intervention gained the buy-in from the upper level management of the eLearning organisations. As he stated “the priority was to open the eyes of the companies…. I think we definitely achieved that”. However, it was noted that this was not done by closing the divide between the technology and business functions but by directly communicating the business value of the technologies directly to them. Furthermore, there is evidence to demonstrate the intervention directly impacted on the participation of the eLearning organisations in the project. The Programme Director felt it “made it easier for the rest of the open days as they will be supported by their senior executives”. Secondly, he stated that after the intervention many of the “senior commercial guys” began to see opportunities in the eLearning market relating to the technologies being developed in the institute. Thirdly, he specifically noted that “from a CEO level they would have appreciated that you have helped them understand the need for a continuous method for their R (research) along with their D (development)”. Further endorsing the intervention and highlighting its effectiveness, the Director of Informatics (Enterprise Ireland) made the business model a compulsory component of future ILRP’s.

Success from the business model intervention was also quite evident within the companies themselves. While interviewing the CTO in Company A he said “you have the (CEO) all fired up about it (the project) and I don’t know what you did. It will make it a lot easier for us to have conversations” with the institute. He also noted that the CEO was “quite bullish” about the project and without the intervention the CEO “would have struggled to map that (the project potential) back into the realities of what we want to do over the next two years”. Directly after the intervention, Company D saw an innovative opportunity with one of their biggest customers (multinational – over 14000 employees) and brought their Chief Learning Officer from the US to the next open day. This highlighted the fact that senior managers began to see opportunities from the technology and began to take action in utilising the research. This was a very positive result for the project as there was a risk that the eLearning organisations would think the technologies had very little practical applicability. After the intervention, there is clear evidence to show that the CEO of Company C began to take notice of the new shift in eLearning. In a meeting he explicitly stated, “I am really intrigued in this formal/informal
learning and it scares the hell out of me a little bit”. The CEO also added that he saw “definite opportunities for companies like us” within the Web2.0 domain. This is a particular case in point, as it is coming from a company that saw very little value in the new technologies. They firmly believed that their customers would not be willing to pay for extra layers of functionality in their product.

As the evidence suggests there is a strong link between the intervention and the increased participation and innovative behaviour of the firms. However, further evaluation of the intervention shows that none of the eLearning organisations implemented the business model framework as a way of common communication platform for creating a shared vision after the intervention. Investigating why Company A didn’t use the business model framework, the R&D engineer found that it did give “some good use cases on how to apply the technologies” but for the most part was “boring”. He further stated that “it will be easier (to communicate the value) when we get to see pieces of the technology”. At that time the CTO was also having problems in communicating the value of the technologies. In an interview he explicitly stated having a problem in “getting a better handle on just what is it that you (research institute) are doing that might benefit us”. It was clear he was having a problem with this as he reiterated “it’s hard to tie all this together, unless you have the vision of where it ends up”. Interactive Services were having the same problem. Their CTO was also finding it difficult to tie all the new technologies together to form a business case or vision to communicate to the CEO and customers. He pushed the need to have “real problems to glue the technology together”. Furthermore, it was deemed that use cases were more appropriate for method for the CTO’s to communicate the value of the technologies. As a result, both ThirdForce and Intuition collaborated on generating use cases to aid the rest of the organisations to better see the potential of the technologies and understand how they could implement them in an industry specific scenario. Both companies dedicated a resource to the task and a number of use cases were developed that covered the pharma, financial and hi-tech industries. However, as noted by the Project Leader within the Institute, “they did not go far enough with the use cases” and failed to push the boundaries of what was possible. Nevertheless, after these use cases had been developed the organisations were “able to visualise” (CEO, Interactive Services) the technologies in an actual real life scenario.

4.2.1 Summary of Evaluation

Outlining the key findings Table 3 summarises the evaluation of the intervention. Above all the business model framework clearly communicated the potential value of the technologies being developed in the research institute, which in-turn gave increased freedom to the CTO’s and technology personnel to get more involved in the project. Evidence pointed to increased commitment to the project from the organisations, the cause of which can be linked back to the business model intervention. However, this all resulted from actions taken during the intervention. One of the objectives was to show the value of the business model framework so that the CTO’s would utilise it as a common communication platform. However, there was very little evidence that CTO’s utilised the business model framework. One of the main reasons was that the CTO’s were more interested in understanding how to apply the technologies to current real life use cases rather than exploring the future possibilities. This is consistent with the general modus operandi of the CTO’s as they were constantly dealing with day-to-day technological issues.

<table>
<thead>
<tr>
<th>Pre-Intervention</th>
<th>Intervention</th>
<th>Post-Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message</td>
<td>Potential of New Technologies</td>
<td>Potential of New Technologies</td>
</tr>
<tr>
<td>Medium</td>
<td>Financial Models</td>
<td>Business Model</td>
</tr>
<tr>
<td>Facilitator</td>
<td>CTO</td>
<td>External Facilitator</td>
</tr>
<tr>
<td>Result</td>
<td>Lack of Understanding by the CEO’s</td>
<td>Increased understanding from CEO’s</td>
</tr>
<tr>
<td></td>
<td>Lack of commitment from business executives</td>
<td>Increased Commitment</td>
</tr>
</tbody>
</table>

Table 3: Summary of Key Findings
5 DISCUSSION AND CONCLUSIONS

As stated by Reich and Benbasat (1996) one of the key influencing factors of social alignment is the type of communication used in the formulation of a shared understanding/vision. However, within the findings the use of financial analysis as a common language created a number of issues between the business and IT/IS domain. The primary issue was that while the CEO’s of the organisations used financial analysis to define strategies and objectives, the CTO’s primarily discussed technological issues relating to current products. This highlighted the increased effectiveness of the business model framework as a communication platform. However, increased scrutiny of the findings uncovered the impact of the researcher in the intervention as a factor in developing the shared understanding. Yet this was not the sole reason for the observed impact as the research had been involved with the organisations through the open days with little success. This would show that even though the business model framework was not adopted it did increase the impact of the external agent and vice versa.

The findings show that the business model framework clearly demonstrated value of the new/innovative technologies to the CEO’s of the eLearning organisations, which in turn increased their commitment to the initiative. One of the key reasons why the business model was more successful in communicating the value of new technologies than the financial analysis methods is that it did not exclude vital information in an effort to oversimplify the potential value. As highlighted by Christensen et al. (2008) using financial modelling can exclude the impact of not investing in innovative projects. In line with the research of Chesbrough and Rosenbloom (2002) this study demonstrated that the business model was able to map the financial and technological domains. Furthermore, the study shows that the business model was able to present possible opportunities that could be exploited by the eLearning organisations themselves or their competitors. In doing so the CEO’s were able to see the potential of the new/innovative technologies but also the risk in not assigning resources to exploring these technologies. In effect, the findings demonstrate the suitability of the business model as a communication tool between business and technology roles, particularly within an innovation setting. More specifically, the study supports research such as Amit and Zott (2001), which demonstrates the business model as a strong tool in presenting the value created within an eBusiness environment.

Within the findings, the success of the business model intervention in enabling the communication of the potential value of the new/innovative technologies is linked and substantiated through the marked increase in participation by the organisations. However, the key goal of the intervention was to create a common communication platform that would facilitate social alignment in the organisation and support the development of a shared understanding/vision. Nevertheless, the business model failed to become a common communication platform for the organisations. For the business model to become a common communication platform, it needed to be adopted by the CTO’s as they were the main connection between the research institute and the eLearning organisations. It has long been known that excellent business/technology relationships within an organisation require the technology representatives to “promote IT as an agent of business transformation” and to “contribute beyond the IT function role” (Feeny et al., 1992). Moreover, some have shown the inability to do this is at the onus of technology representative of the organisation (Lutfman et al., 1999). Yet, this study demonstrated that after the intervention little effort was made in developing a shared vision/understanding, utilising the business model or any other communication platform. Further analysing the success note during the intervention, its main benefit was that it provided a suitable tool that could be used for building a shared vision and understanding through a common communication platform. Moreover, the intervention demonstrated how the business model framework could be used and provided one shared vision for the eLearning industry. The key question is why did the organisations positively respond to the actual intervention but ultimately fail to implement the business model and further develop a shared vision after the intervention was completed?
One explanation can be found by discussing the main difference between the period during the intervention and period after the intervention. The two main differences between these two periods include: (i) the facilitating role of building the vision was not undertaken by either the CTO, and (ii) the industry vision was primarily built using an external viewpoint during the intervention. With regard to the first point, during the intervention the role of building the shared vision was undertaken by the researcher. The aim was that after the intervention this would be continued by the CTO. However, this did not happen. One reason for this is that the role of building a shared vision has been viewed more suitable for person outside of the technology domain. Literature has documented the need of an “organizational architect” to work with both technologists and strategists to grow the organisational and technological capabilities needed to build a vision (Sauer & Willcocks, 2002). Notwithstanding the fact that Sauer and Willcocks (2002) do not specify that the role should be performed by somebody external to the organisation. They do advocate the need for a participant that sits outside both the technology and business domains to build a mutual understanding between the two. This would explain the failure of technologists to facilitate the building of a shared vision after the intervention. Furthermore, the findings would suggest that placing the onus on the technologists (CTO’s) to build a shared vision (Luftman et al., 1999; Cohen & Toleman, 2006) or placing the onus on both technologists and business executives to understand the issues of their counterparts to build a share vision (Kearns & Lederer, 2003; Luftman & Kempaiah, 2007) is not an optimal solution within an eBusiness or innovation setting.

Furthermore, the second point which highlights the fact that the industry vision built during the intervention came from a viewpoint external to the eLearning organisations is also significant. In particular, external input is not discussed in Reich and Benbasat’s (2000) research on the factors of social alignment. Their research primarily focuses on internal communication between the technology and business roles of the organisation in order to build a shared cognitive map or shared vision. However, literature within the innovation domain and specifically around exploring opportunities has noted the need for cognitive diversity. In particular, Gilsing and Nooteboom (2006) highlight the fact that organisations that open up to a variety of insights are more successful than those who did not. Furthermore, they note that cognitive diversity or external input is especially important when an organisation is going from a stage of exploiting current technologies to exploring new technologies (as was the case with the eLearning organisations). In addition, in their model of absorptive capacity Zahra and George (2002) posit the need for social integration mechanisms to build a shared understanding but also the need for external input to feed into that understanding. Again this would indicate that an organisation that needs to build their absorptive capacity or ability to internalise new technologies need external input in building a shared vision.

In effect, this study has shown that the business model acted as a better tool for both codifying external knowledge and integrating it socially than financial languages. However, using the business model as a communication platform alone is not enough to increase the social alignment and build a shared understanding/vision between business and technology, as advocated by Osterwalder et al. (2005). The research highlighted that the external agent had an effect on the development of social alignment. This would infer that within an innovation setting, careful thought should be put into the person who is going to fulfil the role of implementing the business model as a communication platform. However, it must be reiterated that the business model framework did have a significant impact as the external agent was involved with the organisations before the intervention with no observed impacts.

References


Towards A Reference Model for Online Research Maps

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0592.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>knowledge transfer, Web Site Analysis, Conceptual modelling, Reference modelling</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
TOWARDS A REFERENCE MODEL
FOR ONLINE RESEARCH MAPS

Knackstedt, Ralf, University of Muenster, Leonardo-Campus 3, 48149 Münster, Germany,
ralf.knackstedt@ercis.uni-muenster.de
Lis, Łukasz, University of Muenster, Leonardo-Campus 3, 48149 Münster, Germany,
lukasz.lis@ercis.uni-muenster.de
Stein, Armin, University of Muenster, Leonardo-Campus 3, 48149 Münster, Germany,
armin.stein@ercis.uni-muenster.de
Barth, Ingo, University of Muenster, Leonardo-Campus 3, 48149 Münster, Germany,
bingo@uni-muenster.de
Becker, Jörg, University of Muenster, Leonardo-Campus 3, 48149 Münster, Germany,
joerg.becker@ercis.uni-muenster.de

Abstract

Online research maps are websites that present information about certain research activities in a
structured manner. Institutions like universities, states, or individual researchers use them as knowl-
edge base to identify and communicate “who knows what” and “where can the respective information
be found”. Furthermore, these items are used as a research marketing measure, stressing the impo-
tance of this kind of portal. However, research maps differ in their range of functionalities, their re-
spective naming, their target audience and so on. Thus, it is an exhausting task identifying and choos-
ing the right set of functionalities. Our goal is to offer a template for the creation of common practice
online research maps. For this purpose, we present a reference model and describe its development
process. As preliminary measure, 66 research maps have been carefully analysed with respect to the
formerly mentioned aspects. Derived from the results of our analysis, common practice was identified
and used as basis for developing a reference model for online research maps. For development pur-
poses, an existing language for describing internet portals was chosen and adapted to suite the re-
quirements of describing research maps. The reference model presented in this article was then evalu-
ated in a multi-methodical procedure.

Keywords: Web Site Analysis, Reference Modelling, Conceptual Modelling, Knowledge Transfer.
1 MOTIVATION

These days, main deficiencies in information supply are no longer caused by lacking availability of information but rather by information overflow and high complexity in information processing (Porter & Millar 1985). This problem is addressed by the research area of knowledge management, which elaborates on how to identify, gain, generate, disseminate, utilise, and retain knowledge (Probst & Raub & Romhardt 1999). Knowledge management is not only a success factor in business but it is also of highest importance for researchers and research networks. This paper addresses the above stated deficiencies by supporting the development of online research maps.

Online research maps are Internet-based knowledge management instruments, which present research activities through answering different questions, like (a) “who is conducting the research?”, (b) “what is being researched?”, (c) “what results have been achieved?”, and (d) “who is paying for the research?”. Research maps give a general overview of the involved parties, research topics, and achieved results, trying to emphasise existing mutual relationships. These relationships can be, for example, of geographical, organisational, financial, or causal nature. Online research maps can significantly reduce the effort put in the search for knowledge assets and the respective experts due to structured – often visual – representation (Eppler 2001). Thus, online research maps facilitate getting familiar with an unacquainted research area. Furthermore, the database of the research map can be quantitatively analysed to identify research gaps. By giving an insight into the research activities of an organisational unit, research maps are not only a knowledge management instrument but can also be used as a marketing instrument to make own research results available for public. For instance, the U.S. National Cancer Institute uses an online research map (http://www.cancermap.org) to “connect the dots between all researchers involved in pancreatic cancer research to speed the development of national strategies and to leverage resources for pancreatic cancer research”.

While using and hosting an online research map yields certain advantages, their development can be an exhausting undertaking when conducted manually in a try-and-error manner. As organisations conceive their individual solutions independently, much work is carried out redundantly. In many cases, research organisations’ budgets limit the effort that can be brought into the development of a research map. Especially in this case it is important for those organisations to get support designing the functionality and structure of their internet portal.

Reference modelling has been proven a successful means of knowledge transfer through capturing common-practise solutions for information systems development (for detailed information about the advantages of reference models see e.g. vom Brocke 2007; Frank & Strecker 2007; Fettke & Loos 2007). Our goal is to develop a reference model for online research maps, which can support constructors during the conceptual phase of a development project of an online research map.

Hence, there are two major research questions we address with our paper.

(1) What functionalities and structures should be included in the reference model for online research maps to capture common practise?

(2) How can this common practice be graphically represented by a modelling language?

Our development process of the reference model includes the collection and analysis of existing websites, the choice and adaptation of an appropriate modelling language, data-driven reference model construction, and finally the evaluation (cf. Figure 1). The remainder of this paper proceeds correspondingly. In Section 2, we present our explorative analysis of existing online research maps. It consists of an online-based search for websites, their analysis, and identification of implemented functionalities. Furthermore, as the naming of these functionalities varies from website to website, we establish a common vocabulary for the used functionalities. In Section 3, a modelling language appropriate for online research maps is chosen, adapted, and specified by a language-based meta-model. Section 4 covers the creation of our reference model, which is conducted in an iterative way. In the following
section the evaluation of our reference model is discussed. On the one hand, the model is validated against randomly selected websites. On the other hand, we evaluate the applicability of our model to a concrete real-life situation of research map construction. In Section 6, we conclude with an outlook for further application and evaluation activities as well as for important contributions of the reference model to other research questions.

![Diagram](image1.png)

**Figure 1: Development Process of the Reference Model**

## 2 EXPLORATIVE ANALYSIS OF RESEARCH MAPS

As a preliminary step, typical functionalities of research maps had to be identified. These functionalities should serve as basis for the selection of an appropriate modelling language. Thus, from July to August 2008, an analysis of existing research maps was conducted. For this, the search engine Google (http://www.google.com) was used to retrieve results for keywords like “research map” (and other corresponding English terms) and “Forschungslandkarte” (and other corresponding German terms). For each term, the first 250 results were analysed whether or not they belong to a research map. Besides many links leading to pages just providing information about research maps or solely providing links to them, 76 research maps were identified as being adequate for our analysis at the first moment. However, 9 of them were offered as print media and one of the web based research maps did not deliver the required login information in time. This led to a total of 66 research maps analysed, originating from eight different countries. Most of them (42) originate from Germany, another 10 from the United States. The remaining 14 maps originate from several European countries. Each of the 66 research maps was scanned for functionalities it provides. As the naming of the research maps’ functionalities differ, coding them with common terms was required. Once functionalities for a certain research map were identified, they were recorded in a contingency table. This enabled us to identify common expressions and aggregate them thematically (cf. Table 1). For this, every term found on a certain online research map and identified as being relevant was assigned to one of the 52 defined functionalities. Those again were aggregated to another set of 12 classes, characterising them in a more general manner. Furthermore, for common understanding we defined the meaning of each of the functionalities and counted their absolute occurrences (#Freq.) and their percentage share (%).

![Screenshot](image2.png)

**Figure 2: Screenshot of an Exemplary Research Map (Johns Hopkins Center for Global Health)**

---

*Proceedings ECIS 2009*
Table 1: Common Expressions, Their Relation to Functionality and Their Definitions

<table>
<thead>
<tr>
<th>Class</th>
<th>ID</th>
<th>Functionality</th>
<th>RM</th>
<th>Characterisation</th>
<th>#Freq</th>
<th>% in RM</th>
<th>% not in RM due to 5% rule</th>
<th>% not in RM due to specialisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge representation</td>
<td>01</td>
<td>Index of institutions</td>
<td>X</td>
<td>A list which offers information about organisations, institutions and enterprises.</td>
<td>42</td>
<td>63.6</td>
<td>not in RM due to 5% rule</td>
<td>not in RM due to specialisation</td>
</tr>
<tr>
<td></td>
<td>02</td>
<td>Index of professionals</td>
<td>X</td>
<td>A list which offers information about people.</td>
<td>19</td>
<td>28.8</td>
<td>not in RM due to 5% rule</td>
<td>not in RM due to specialisation</td>
</tr>
<tr>
<td></td>
<td>03</td>
<td>Knowledge Structure Map</td>
<td>X</td>
<td>A visual representation of a knowledge domains structure and its relations.</td>
<td>10</td>
<td>15.2</td>
<td>not in RM due to 5% rule</td>
<td>not in RM due to specialisation</td>
</tr>
<tr>
<td></td>
<td>04</td>
<td>Knowledge Asset Map</td>
<td>X</td>
<td>A map which visually qualifies the existing stock of knowledge of an individual, a unit or an organization.</td>
<td>49</td>
<td>74.2</td>
<td>Search: Project search; Search for text</td>
<td></td>
</tr>
<tr>
<td>Searching</td>
<td>05</td>
<td>Content search</td>
<td>X</td>
<td>A full-text search on the content of the website.</td>
<td>26</td>
<td>39.4</td>
<td>Search: Institution search</td>
<td></td>
</tr>
<tr>
<td></td>
<td>06</td>
<td>Institution search</td>
<td>X</td>
<td>A search limited on information about institutions, organisations and enterprises, which are part of the website.</td>
<td>15</td>
<td>19.7</td>
<td>Search: Contact name search; Employee Locator</td>
<td></td>
</tr>
<tr>
<td></td>
<td>07</td>
<td>Person search</td>
<td>X</td>
<td>A search limited on information about researchers or members of staff.</td>
<td>2</td>
<td>3.0</td>
<td>Cooperation search*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>08</td>
<td>Cooperation search</td>
<td>%</td>
<td>A search for enterprises requesting cooperations.</td>
<td>1</td>
<td>1.5</td>
<td>Special equipment search*</td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td>09</td>
<td>Area of research</td>
<td>X</td>
<td>An introduction to the area of research of the website.</td>
<td>27</td>
<td>40.9</td>
<td>Name of the area of research</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Research methods</td>
<td>X</td>
<td>An explanation of methods used to create the research map.</td>
<td>5</td>
<td>7.6</td>
<td>Name of the research method</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>Research project</td>
<td>X</td>
<td>The existence of a page containing research projects.</td>
<td>27</td>
<td>40.9</td>
<td>Projects; Activities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>Research results</td>
<td>X</td>
<td>The existence of a page containing research results.</td>
<td>35</td>
<td>53.0</td>
<td>Publications; &quot;Name of the research project&quot; ; Reports</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Publications</td>
<td>X</td>
<td>A list of articles, proceedings and reports published.</td>
<td>18</td>
<td>27.3</td>
<td>Publications; Papers; Proceedings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>References</td>
<td>X</td>
<td>The existence of references to pages of the website to bring evidence to statements and reports.</td>
<td>11</td>
<td>16.7</td>
<td>References; Papers; Key documents</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>Literature database</td>
<td>X</td>
<td>The existence of a database containing a list of textbooks or scientific journals related to the map.</td>
<td>4</td>
<td>6.1</td>
<td>Library; Journals</td>
<td></td>
</tr>
<tr>
<td>Research funding</td>
<td>16</td>
<td>Funding</td>
<td>X</td>
<td>Information about scientific grants and funding opportunities.</td>
<td>8</td>
<td>12.1</td>
<td>Funding; Funding Opportunities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>Awards</td>
<td>%</td>
<td>Information about science awards gained by the portal operator.</td>
<td>7</td>
<td>10.3</td>
<td>Awards</td>
<td></td>
</tr>
<tr>
<td>Analysis</td>
<td>18</td>
<td>Link to external Websites</td>
<td>X</td>
<td>A link to an external Website which contains statistical data and analysis of an area of research.</td>
<td>9</td>
<td>13.6</td>
<td>Statistics*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>Tabular reports</td>
<td>X</td>
<td>Analyses as tabular reports.</td>
<td>10</td>
<td>15.2</td>
<td>Overview of...; Quick view of...; Analysis*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>Chart reports</td>
<td>X</td>
<td>Analyses as chart reports.</td>
<td>12</td>
<td>18.2</td>
<td>Apply to join; Join the researchmap</td>
<td></td>
</tr>
<tr>
<td>Participation</td>
<td>21</td>
<td>Intranet login</td>
<td>X</td>
<td>Enters to a password protected area of the website.</td>
<td>19</td>
<td>28.6</td>
<td>Login</td>
<td></td>
</tr>
<tr>
<td>Legal information</td>
<td>22</td>
<td>Imprint</td>
<td>X</td>
<td>General and legal information about the company, responsible for the webpage.</td>
<td>43</td>
<td>65.2</td>
<td>Imprint; Terms of use</td>
<td></td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>Privacy Policy</td>
<td>X</td>
<td>Legal document stating with the information related to customers and merchants private profiles.</td>
<td>14</td>
<td>21.2</td>
<td>Privacy Policy; Privacy and Cookies Policy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>Disclaimer</td>
<td>X</td>
<td>Statement intended to define or delimit the scope of rights.</td>
<td>5</td>
<td>7.6</td>
<td>Disclaimer</td>
<td></td>
</tr>
<tr>
<td>Common components</td>
<td>25</td>
<td>Site map</td>
<td>X</td>
<td>A representation of the structure of a webpage.</td>
<td>24</td>
<td>36.4</td>
<td>Site map</td>
<td></td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>Language</td>
<td>X</td>
<td>Opportunity to change a websites used language.</td>
<td>28</td>
<td>42.4</td>
<td>&quot;Text&quot;; &quot;Icon&quot;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>Contact</td>
<td>X</td>
<td>A page containing the portal operations contact details.</td>
<td>57</td>
<td>86.4</td>
<td>Contact; Contact us: Main Contacts</td>
<td></td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>FAQ Help</td>
<td>X</td>
<td>A list of frequently asked questions.</td>
<td>9</td>
<td>13.6</td>
<td>FAQ</td>
<td></td>
</tr>
<tr>
<td>Portal components</td>
<td>29</td>
<td>Latest news</td>
<td>X</td>
<td>A page containing information about the latest research results and news of the area of research or the portal.</td>
<td>19</td>
<td>28.6</td>
<td>News; What’s new?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>Portal operator</td>
<td>X</td>
<td>Information about the portal operator like the history, company organisation and directions.</td>
<td>37</td>
<td>56.1</td>
<td>&quot;Name of the operator&quot;; About...; About us</td>
<td></td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>Events calendar</td>
<td>X</td>
<td>A list of events related to the research map in chronological order.</td>
<td>25</td>
<td>37.9</td>
<td>Events; Events Calendar; Conferences</td>
<td></td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>Objectives</td>
<td>X</td>
<td>Information about the ambition and purpose of the research portal.</td>
<td>15</td>
<td>22.7</td>
<td>&quot;Welcome page&quot;; Mission; Purpose; Vision; Objectives</td>
<td></td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>Sponsors</td>
<td>X</td>
<td>Pages or links containing companies’ names or logos supporting the website financially.</td>
<td>12</td>
<td>18.2</td>
<td>&quot;Banner ad&quot;; Funding Partners; Funders</td>
<td></td>
</tr>
<tr>
<td></td>
<td>34</td>
<td>Links</td>
<td>X</td>
<td>A list of links related to the website.</td>
<td>25</td>
<td>37.9</td>
<td>Links; Related links; Useful websites</td>
<td></td>
</tr>
<tr>
<td>Media</td>
<td>35</td>
<td>Press release</td>
<td>X</td>
<td>A page listing press releases chronologically.</td>
<td>37</td>
<td>56.1</td>
<td>News; News Center; Press release*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>Newsletter</td>
<td>X</td>
<td>Gives an user the opportunity to subscribe to newsletters.</td>
<td>13</td>
<td>19.7</td>
<td>Newsletter; Mailing Service; Subscription Service</td>
<td></td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>Podcast</td>
<td>X</td>
<td>Gives an user the opportunity to watch or subscribe to podcasts.</td>
<td>4</td>
<td>6.1</td>
<td>Podcasts; Video Gallery</td>
<td></td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>RSS Feed</td>
<td>%</td>
<td>Gives an user the opportunity to subscribe to RSS Feed.</td>
<td>2</td>
<td>3.0</td>
<td>RSS Feed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>Vacancy</td>
<td>X</td>
<td>A page containing job advertisement.</td>
<td>9</td>
<td>13.6</td>
<td>Jobs; Vacancy; Job exchange*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>Further education</td>
<td>X</td>
<td>A page containing information about advanced training opportunities.</td>
<td>7</td>
<td>10.6</td>
<td>Further education*; Advanced training*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>41</td>
<td>Internship</td>
<td>X</td>
<td>A page containing trainees and internship offerings.</td>
<td>2</td>
<td>3.0</td>
<td>Internship*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>Professional training</td>
<td>X</td>
<td>A page containing information about professional training opportunities.</td>
<td>3</td>
<td>4.5</td>
<td>Professional training*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>43</td>
<td>Degree dissertation</td>
<td>X</td>
<td>A page containing diploma theses.</td>
<td>5</td>
<td>7.6</td>
<td>Diploma Thesis; Diploma*; Dissertation*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>Academic studies</td>
<td>X</td>
<td>A page containing information about academic studies related to the area of research.</td>
<td>9</td>
<td>13.6</td>
<td>Education; Academic studies*</td>
<td></td>
</tr>
<tr>
<td>Career</td>
<td>45</td>
<td>Recommend this page</td>
<td>X</td>
<td>Opportunity to email the URL of a page to a friend.</td>
<td>5</td>
<td>7.6</td>
<td>Email to a friend; Share an article; Send it to a friend</td>
<td></td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>Feedback</td>
<td>X</td>
<td>Opportunity to submit ideas and corrections to the portal operator.</td>
<td>4</td>
<td>6.1</td>
<td>Feedback; Submit ideas; Corrections</td>
<td></td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>47</td>
<td>Shopping cart for papers</td>
<td>X</td>
<td>Opportunity to put papers and brochures into a shopping cart before ordering.</td>
<td>11</td>
<td>16.7</td>
<td>Shopping cart</td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>Visitor counter</td>
<td>%</td>
<td>A function which counts the amount of visitor of a page.</td>
<td>2</td>
<td>3.0</td>
<td>Hits; You are the x visitor to the site</td>
<td></td>
</tr>
<tr>
<td></td>
<td>49</td>
<td>Keyword index</td>
<td>X</td>
<td>A list containing domain related key words.</td>
<td>1</td>
<td>1.5</td>
<td>Index</td>
<td>Make a gift</td>
</tr>
</tbody>
</table>
For this article, common German expressions were translated to English (marked with *). Terms in quotation marks serve as placeholders (e.g. “Text” stands for “English” or “German” (ID 28), “Icon” stands for the respective flag symbols). After having built the overview, we created a screenshot of each research map and tagged it with the respective term (cf. Figure 2). Thus, documentation of the assignment results is available. As an example for the assignment, the research map of the “Johns Hopkins Center for Global Health” (http://research.hopkinsglobalhealth.org/researchmap.cfm) provides the functionality Project Map, which was coded with the respective term Knowledge Asset Map (ID 04). As another example, About the Center was coded with the term Portal Operator (ID 32).

3 ADAPTATION OF A MODELLING LANGUAGE

A necessary task during the reference modelling process is the selection and – if necessary – adaptation of an adequate modelling language (Becker et al. 2008). For websites, various approaches towards modelling requirements definition are proposed, which can be differentiated into three branches of development (cf. Figure 3). One group of approaches originates from data modelling and is based on Entity-Relationship Models (Chen 1976). Other approaches primarily derive from Hypertext Modelling (Garzotto & Paolini & Schwabe 1991; Halasz & Schwartz 1994). A further branch of development comprises object-oriented approaches (Rumbaugh et al. 1991; Rumbaugh & Jacobson & Booch 1998).

![Image of Figure 3: Modelling Approaches for Web Applications (Brelage 2006)]

The extended World Wide Web Design Technique (eW3DT) by Scharl (1998) was chosen for the construction of the reference model for research maps. The crucial advantage of eW3DT was the fact that reference models for internet portals have already been constructed using this modelling language (Scharl 1997). Thus, this modelling language is intended to be partly reused in the context of this paper. Furthermore, this modelling language incorporates all required elementary model constructs provided by the other modelling languages. eW3DT distinguishes the five website element types page, interaction, index, menu and file, which can be either static or dynamic. In contrast to static web elements, the content of dynamic web elements may vary depending on the time being accessed. A website structure is modelled by establishing relations between website elements using different connection types. The contents of single website elements can be refined in sub-models. The analysis as presented in Section 2 revealed requirements that made it necessary to adapt eW3DT. The meta-model in Figure 4, for which we used an extended version of the ERM data modelling language (Chen 1979), illustrates the project-specific extensions to eW3DT. Derived from the functionalities of the class analysis (cf. Table 1, IDs 19-21), the reference model is required to be able to describe the content-related structure of reports that should be available in the research portal. Thus, the new element type Web-OLAP represents a quantitative report that is available on the research map and may also be manipulated through the OLAP operations: rotate, slice, dice and drill-down respectively roll-up.
Concerning the extension of the meta-model for the specification of analysis reports, the model constructs that are commonly used in modelling languages for the functional specification of OLAP systems have been adopted (Bulos 1998; Golfarelli & Maio & Rizzi 1998; Holten 2003). The content-related specification of OLAP reports is defined by values of columns and rows which determine a two-dimensional projection of a navigation space. Navigation spaces are made up of dimensions and dimension scopes (e.g. date defined by year, institution defined by country and research theme defined by topic); and systems of metrics (e.g. number of research results). Thereby, the navigation space clarifies which metrics should be analysable according to which structure of dimensions. The design relying on rows and columns determines the tabular structure of the report. If the element type Web-OLAP-Report is being marked as dynamic, the report is generated from the database on demand and thus can contain different metric values at different access times. If the report is static, however, its contents remain the same regardless the time the web portal is accessed. By specifying additional attributes of the Web-OLAP element, the available OLAP operations can be chosen. Further information functionalities were identified empirically, some of which pointed out that indices are frequently used in research portals to give an overview of various entities, like institutions, people, and topics. Within the extension of the meta-model, the element type index is intended to be assigned to different
dimensions – respectively dimension scopes – which can also be used within the context of report-based analysis. For example, by using this meta-model extension it is possible to express that the index directory of institutions is presented in different ways, e.g. ordered alphabetically/geographically or by Institution Type/Area of Research (cf. Figure 6). Besides the hierarchically structured index navigation, it should be possible to search throughout thematically separated areas, e.g. via full-text search. For this purpose, the additional element type search has been included. By defining a navigation space, the content-related restriction of a supported search can be specified.

Another extension to the meta-model was motivated by the fact that in eW3DT database symbols are assigned to the element type interaction in order to specify that data is collected or provided during this interaction. As decisions are to be made concerning the naming and structure of databases, we used data clusters to specify this data in our adapted eW3DT variant. The data clusters are excerpts of an Entity-Relationshipship Model (ERM). The ERM excerpts describe the data addressed in the interaction from the requirements definition point of view. Thereby, a consistent application of reference modelling in requirements definition is guaranteed. However, this abstracts from the way the data is physically structured in the database and from the database itself.

4 A REFERENCE MODEL FOR RESEARCH MAPS

With the functionalities identified in Section 2 and the modelling language as prerequisite for the creation of the reference model at hand, we chose an iterative process for building its structure. During the first iteration step, a set of five research maps, which provided a considerable amount of functionalities, was analysed regarding their structure (cf. Table 2). For each functionality the respective web element type (Page, Menu, Index, Search, Web-OLAP, File and Interaction) as Static or Dynamic, and the links between them (Dynamic, Static, Including, or Representing) were chosen. If necessary, additional functionalities were included and placed at an appropriate position.

<table>
<thead>
<tr>
<th>Step</th>
<th>Research Map</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation</td>
<td>Forschungslandkarte zur hybriden Wertschöpfung (Ec); Kompetenznetzwerk Stammzellforschung NRW (Me); Forschungsportal Sachsen-Anhalt (Ec); Forschungsportal der Universität für Bodenkultur Wien (Env); FIDES der Wirtschaftsuniversität Wien (Sc)</td>
</tr>
<tr>
<td>Iteration 1</td>
<td>Forschungslandkarte Asklepios Kliniken (Me); Forschungslandkarte zur Excellence Initiative des BMBF (Ed); Forschungslandkarte zu Bundeseinrichtungen des BMBF (PA); Fachportal Pädagogik (Ed); Forschungslandkarte Windenergie (Ene); Duke University (Me); Knowledge Base Social Science Eastern Europe der GESIS (CS); Kunststoffland NRW e.V. (Ch); Landesinitiative Projekt Zukunft Berlin (Sc); Alaska Science Portal des USGS (Env)</td>
</tr>
<tr>
<td>Iteration 2</td>
<td>Forschungslandkarte zur Helmholtz-Gemeinschaft des BMBF (RC); Security Research Map des BMBF (SE); Kompetenznetzwerk Katalyse von DECHEMA (Ch); Forschungsportal der Deutschen Rentenversicherung (IF); eigenfactor.org (Sc); Research Map des John Hopkins Center for Global Health (Me); Kompetenznetzwerk Neuro NRW (Me); Sandia National Laboratories (Sc); Kompetenzcluster des VDI (Ec); Forschungslandschaft Sachsen der VEMAS (In)</td>
</tr>
<tr>
<td>Iteration 3</td>
<td>Forschungslandkarte der Hochschulen des BMBF (Ed); Kompetenznetzwerk dezentrale Energietechnologien e.V. (Ene); Informations- und Wissensplattform Chem.de (Ch); Holzcluster Steiermark (Ti); Forschungskarte der Deutschen Massivumformung e.V. (In); Cancermap des National Cancer Institute (Me); Bundesumweltamt (Env); Forschungsatlas QPT der Universität Zürich (Ps); Forschungslandkarte des Deutschen Instituts für Erwachsenenbildung (Ed); Science Map der University of Saskatchewan (Sc)</td>
</tr>
</tbody>
</table>

Legend of research areas:
Ch = Chemistry; CS = Culture und Social Affairs; Ec = Economics; Ed = Education; Ene = Energy; Env = Environment; IF = Insurance and Finance; In = Industries; Me = Medicine; PA = Public Administration; Ps = Psychology; RC = Research Centre; Sc = Science (multi disciplinary); SE = Safety Engineering; Ti = Timber Industries;

Table 2: Iteration Steps for Building the Reference Model

The first iteration step consisted of ten research maps being compared to the model originating from the initiation step. Ten functionalities had to be additionally included and placed in an appropriate position in the model. Others had to be repositioned. During the second iteration, only three more elements had to be included.

During the third iteration, no more additional elements were required and no repositioning had to take place. Thus, we defined the result as our reference models first version (cf. Figure 5). The greyed out elements in Table 1 express, which functionalities of the initial list were left out in the reference model either because not more than five percent of the analysed research maps included them, or because
they were very domain-specific and not of any value for research maps belonging to different domains.

Figure 5: Reference Model Proposal after Three Iteration Steps

The reference model all in all consists of 148 elements and is divided into 14 sub-models, describing subordinate contents. If not accessing the research map via a deep link, e.g. by using a link provided by a search engine, the user enters the research map via the Homepage (level 1.0). There, elements of level 1.1 are included and immediately presented. For maximum impact, on this level the user should be enabled to choose his preferred Language. Furthermore, the Objectives of the map should be presented to give a quick overview about what the site is about. The interactions Intranet Login and Feedback communicate the possibility to get in contact with the operators by sending annotations or proposals for corrections. By entering a secured area, the (active, thus not visiting) user gets the possibility to contribute to the research map, a necessity e.g. for research networks or communities. A file Privacy Policy generates trust and informs about how the operator deals with members data and their profiles. Content Search enables the user to search the content of the research map by keywords. To send the URL of one of the page to interested third parties, the interaction Recommend this Page can be used. The Site Map gives an overview of the structure of the web portal whereas the Disclaimer is a serves as legal insurance.

Web site elements on level 2.0 can be accessed from the homepages navigation menu. Figure 6 illustrates this for the menu Institution, which consists of a dynamic element Directory of Institutions and an element Institution Search.
Institutions

Figure 6: Sub-Model of the “Institutions” Menu

Similar to the menu Institutions, 14 more sub-models were created. A condensed description of all available sub-models is listed in Table 3:

<table>
<thead>
<tr>
<th>Sub-model</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis</td>
<td>This sub-model contains four dynamic Web-OLAP-Applications. Each application is configured to its special scope. The Web-OLAP-Applications Project Analysis, Area of Research Analysis, Institution Analysis, and Publication Analysis will generate significant reports.</td>
</tr>
<tr>
<td>Career Center</td>
<td>Four dynamic indices in the sub model Career Center exist. Each index contains domain-specific information, external links for further information, e.g., Vacancy, Further Education, Academic Studies, and Professional Training.</td>
</tr>
<tr>
<td>Events Calendar</td>
<td>The dynamic index Events Calendar announces events, exhibitions, workshops, and advanced trainings related to the area of research. An external link to an organisers website lets users sign up for an event.</td>
</tr>
<tr>
<td>Institutions</td>
<td>This sub-model lists research institutions alphabetically, geographically or by type of institute. One institute is characterised by its Contact Data, Profile, Main Research Area, involved Members of Staff, Publications, Projects, Patents, Awards, Cooperations, Special Equipment and handled Methods. For large numbers of registered institutions, an Institution Search should be implemented.</td>
</tr>
<tr>
<td>Intranet login</td>
<td>This sub-model contains a menu Intranet linking to two interactions: One interaction manages the administration of user data, the other interaction manages the administration of content data like Projects, Research Results, Publications, and Events Calendar.</td>
</tr>
<tr>
<td>Links</td>
<td>The static or dynamic index Links is a link collection to external websites related to the research portal.</td>
</tr>
<tr>
<td>Media</td>
<td>The dynamic menu Media contains three dynamic indices and one static page. The page Press Contact keeps contact data for press only. The index Press Release lists press reports chronologically. Older reports will be stored in the index Press Archive. The third index Podcasts enables users to play audio and video documents related to the portal. Additionally, there is a possibility to subscribe to the Newsletter.</td>
</tr>
<tr>
<td>Portal Operator</td>
<td>This sub-model consists of five static pages. A Profile describes the portal operator with his core competences, the service offering and philosophy. The History of a firm will be presented as well as its Organisation. Pages with Contact data and Directions are in this sub-model, too.</td>
</tr>
<tr>
<td>Professionals</td>
<td>The sub-model Professionals consists of a Person Search and an Index of Professionals. One professional is characterised by his Contact Data, membership to an Institution, Profile, Main Area of Research, Publications, Projects, Patents, and Awards.</td>
</tr>
<tr>
<td>Publications</td>
<td>A publication is characterised by its Publication Status, Type of Publication, Title, Location, Publisher, Year, Authors, Project, Research Area, Institution and URL. The index Publications lists data sets sorted by time, type or author. Categorial filters in a Publication Search could be used to specify database requests.</td>
</tr>
<tr>
<td>Research Funding</td>
<td>This sub-model includes an index with information about Funding opportunities.</td>
</tr>
<tr>
<td>Research Maps</td>
<td>Maps facilitate the search process in an area of research. Knowledge Asset Maps and Knowledge Structure Maps are modelled as Web-OLAP-Reports. Web-OLAP-Reports are being real-time-generated from the database, so they will be up-to-date permanently.</td>
</tr>
<tr>
<td>Research Projects</td>
<td>This sub-model contains the index of Research Projects which will be described by a Title, a realising Institution, Participants, a Description, its Duration, Keywords, used Methods, Research Results, Cooperations, Financial Supporters and its Project Status. A powerful Project Search helps to control the project data sets.</td>
</tr>
<tr>
<td>Research Results</td>
<td>One or more research results could originate from one research project, so research results are outsourced in its own sub-model. The index of Research Results lists all data sets alphabetically. Using Search Research Results, the output could be limited and sorted by categories. One result will be described by its Title, its publishing Institution, Contact Person, Description, Classification, Keywords, URL its mapping to a Research Project, its situation of Patent Registration, Financial Supporters, its Status, and Publications.</td>
</tr>
</tbody>
</table>

Table 3: Descriptions of the Sub-Models’ Contents
5 EVALUATION OF THE REFERENCE MODEL

Having built the reference model, we chose a multi-methodical approach for our evaluation, initially starting off with testing the abilities of the reference model against a random sample of ten of the remaining research maps. During this, we found out that no more website elements had to be added to the reference model to completely describe the sample. This can be interpreted as an indicator for the reference model completeness. Subsequently, we used the reference model as accompanying measure for the creation of a research map from scratch. By doing so, we wanted to evaluate its usefulness in day-to-day business. The map that was to be set up was on order of an international research network, settled in the domain of information systems. It can be accessed via the URL http://www.forschungslandkarte-hybridewertschoepfung.de. The research map generally required the ability to present information about research artefacts created by the members of the network. It was also required to describe the artefacts with different dimensions, list the respective contact person and additional downloads, if available. For reporting purposes, the dimensions should be analysable by an OLAP system.

The reference model successfully served as a guide for the discussion on and the choice of functionalities implemented for the portal. The artefacts of the research maps are analysable with the OLAP system IBM Cognos by their dimensions. During the next step, we called the users of the research map for participation by filling in a paper-based questionnaire about the applicability and usefulness of the elements provided. No negative comments regarding the content of the portal were given, however the naming of the website was questioned, which, however, was not subject to the application of the reference model anyway.

Furthermore, we conducted semi-structured interviews with seven users. All in all, they too found the research portal to be complete in regards to the requirements. Besides successfully describing the required content, the reference model turned out to be helpful during the creation process of the research map. For a detailed description of how this process can be evaluated, see Frank (2007). However, up to this point, all personnel except third party users of the research map developed are known to the authors and therefore – although leading to very good results – their judgement has to be questioned. Thus, in the next step, the reference model will be used during the development process of a research map for clients personally not known to the authors.

6 OUTLOOK

The reference model in its as-is state seems to be able to completely describe most of the research maps, which are available online. However, as motivated in Section 5, the fourth step of the evaluation phase still has to be taken, which will be done in a larger scale project with participants that are not acquainted or connected to the authors in any way. For this, the reference model will be handed over to the research map development team. The authors will accompany the whole process and observe the reference models applicability for both the choice of elements and the process of developing the research map. The result of this analysis should provide the means for a further adaptation of the modelling language and an even more suitable definition of functionalities. By doing this, another iterative process will be started, improving the reference model’s significance (cf. Figure 1). The reference model serves as basis for a larger scale empirical analysis, encompassing more online research maps. As prerequisite for such an empirical analysis, it is important to define adequate criteria for coding the objects of investigation. Here, the reference model for research maps comes into play by offering criteria in terms of functionalities and common expressions. Once collected, further results should be expected from this data pool:

(1) By being able to work with a larger number of results, we should be able to identify different types of research maps. Concluding from the sample analysed in Section 2, it seems that there are different purposes for which research maps are set up. Firstly, single researchers use the maps as online documentation of their own work. These kinds of maps are very limited in their number of functionalities,
the number of research domains covered and the number of people accessing the site and contributing to it. Secondly, organisations like universities or sponsors set up research maps to either present their own research results or those of the supported projects. The amount of users accessing and contributing to this kind of map is naturally considerably larger. Furthermore, the research map might cover more than one research domain. Due to the larger amount of stakeholders, the demand for functionalities, too, will be higher. Thirdly, research networks use research maps for sharing their results among each other similar to organisations. In opposite to them, there is no need for a central institution managing the research map. Once empirically identified, these different types of research maps might lead to clusters of functionalities connected to them. If this should succeed, the reference model could be enriched with meta-data, specifying for which type of research map the respective functionalities are relevant. This would greatly simplify the choice of functionalities for a certain kind of map, as the adequate set could be derived from the selection of a certain configuration parameter (Becker & Delfmann & Knackstedt 2007).

(2) Another application scenario of the empirical data collected is the construction of a maturity model (Nolan 1979; de Bruin et al. 2005) for research maps. A similar approach addressing portals in general was proposed by Baroni de Carvalho & Ferreira & Choo (2005). A quantitative analysis can show which functionalities are typical for different degrees of maturity. Assuming that the maturity model consists of five – as widely used – stages, it is our expectation that our reference model corresponds with level three, implying a medium level of maturity. While research maps can be considered as knowledge management instruments, such a maturity model is a specific kind of a knowledge management maturity model as e.g. proposed by Freeze & Kulkarni (2005).

References


### VALUATION OF ONLINE SOCIAL NETWORKS – AN ECONOMIC MODEL AND ITS APPLICATION USING THE CASE OF XING.COM

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0483.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Social Networking (e.g. Facebook, second life), Business Case, Customer relationship management (CRM), Web 2.0</td>
</tr>
</tbody>
</table>
VALUATION OF ONLINE SOCIAL NETWORKS – AN ECONOMIC MODEL AND ITS APPLICATION USING THE CASE OF XING.COM

Gneiser, Martin, Department of Information Systems Engineering & Financial Management, University of Augsburg, Universitaetsstrasse 16, D-86159 Augsburg, Germany, martin.gneiser@wiwi.uni-augsburg.de

Heidemann, Julia, Department of Information Systems Engineering & Financial Management, University of Augsburg, Universitaetsstrasse 16, D-86159 Augsburg, Germany, julia.heidemann@wiwi.uni-augsburg.de

Klier, Mathias, Department of Information Systems, University of Innsbruck, Universitaetsstrasse 15, A-6020 Innsbruck, Austria, mathias.klier@uibk.ac.at

Weiß, Christian, Department of Information Systems Engineering & Financial Management, University of Augsburg, Universitaetsstrasse 16, D-86159 Augsburg, Germany, christian.weiss@eliteakademie.de

Abstract

Online social networks are gaining increasing economic importance in light of the rising number of members. The numerous recent acquisitions priced at enormous amounts illustrate this development. Therefore, the growing relevance of online social networks in science as well as in practise revealed the need for adequate valuation models, which take into account the networks’ specific characteristics. Thus, this article develops an economic model for valuation of online social networks. The model allows the evaluation of whether the purchase prices on the market, which recently amounted to millions, are justifiable. Finally, the practical application of the model is illustrated by an example of the major European online social network XING.com.

Keywords: social networking, customer relationship management, business case.
1 INTRODUCTION

One of the most important current changes with regard to the use of the Internet is the transformation of passive information users into active actors, which increasingly create the content of the World Wide Web (WWW) themselves. Along with these changes, the economic impact of established media declines and experts predict heavy socio-economic and political implications (Bernoff et al. 2008). A main driver for this development is the active use of online social networks, where people are connecting and communicating more and more online with one other (Kazienko et al. 2006, Gross et al. 2005). Networking sites such as Facebook.com or XING.com not only provide a technical platform to establish and maintain relationships between users, but also enable users to present themselves to a wide public and to make visible their own social networks (Boyd et al. 2007). This emergent technical and social phenomenon generates an increasingly important economic impact and has spurred enormous attention among researchers and practitioners.

Thus, media and IT companies have been acquiring recently online social networks for considerable amounts to adapt their business models to the new environmental conditions and to reorganize their companies for the future. In 2005, for example, the media company News Corp. acquired the online social network MySpace.com for US $ 580 m. Two years later Microsoft paid US $ 240 m for a 1.6% minority interest in the online social network Facebook.com. The extrapolated value of this company thus amounts to staggering US $ 15 bn. This trend can also be observed in Germany: following a bidding war with the publisher Springer, the German publishing company Holtzbrinck acquired the online student network StudiVZ.de for approximately € 85 m (Sievers et al. 2008). However, the enormous purchase prices for online social networks are also considered critical and experts compare the situation with the dotcom bubble over the turn of the millennium: Martin Sorrell for instance, CEO of the WWP Group – the world’s largest communications services group (and one of the six largest advertising holding companies) –, is cited in the Financial Times Deutschland seriously questioning the valuation of Facebook.com at US $ 15 bn (Lambrecht 2008).

This makes clear, that the important question of how online social networks can be valued using well-founded valuation methods has not yet been answered. Therefore, the objective of this paper is first to develop an economic model for the valuation of online social networks, which takes into account the specific characteristics of these companies, and second to illustrate the practical application of the model to the online social network XING.com using only publicly avaiable data. The paper is structured as follows: In section 2 we define and describe online social networks as a current phenomenon. In section 3 we briefly review the existing valuation approaches to online social networks, before we develop our own quantitative approach in section 4. The practical use of the new model is extensively illustrated by an example of the major European online social network XING.com in section 5. The last section summarizes the results and suggests areas for further research.

2 ONLINE SOCIAL NETWORKS: A CURRENT PHENOMENON

Although Facebook.com was only established in 2004, today more than 90 m people get together in the digital friendship network (Agarwal et al. 2008). This is only one example of how online social networks – aroused by the web 2.0 boom – have evolved into a new, mostly free of cost, mass medium where users\(^1\) present themselves to a wide public and voluntarily reveal parts of their privacy. Beside the exponential growth of online social networks there is a growing realization that online social networks are not simply forums in which individuals congregate. Rather, “these networks create

---

\(^1\) The terms customer, member and user are used synonymously.
substantial value for the individuals who participate in them, the organizations that sponsor them, and the larger society in multiple ways” (Agarwal et al. 2008). The community idea itself, which was long known and extensively researched especially in the field of social sciences (see Bagozzi et al. 2006) and in social network analysis in general (Milgram 1967, Granovetter 1973, Watts 2003), took on new dimensions with the development of the Internet and the emergence of online social networks. In this context, this article focuses mainly on the users’ integration in the online social network (e.g. number of contacts etc.) and the consequences in regard to the economic valuation.

We generally perceive an online social network as a set of actors, which are represented by nodes, and a set of edges (ties) linking pairs of nodes (Adamic et al. 2003, Kazienko et al. 2006, Bampo et al. 2008). The edges represent connections between actors and describe social interactions or relationships. The nodes and edges are usually presented by a graph (Hanneman et al. 2005), as shown in Figure 1. This visualisation especially highlights so-called hubs (Bampo et al. 2008), i.e. actors who have a particularly large number of connections to other actors.

![Elements of an online social network](image)

**Figure 1. Elements of an online social network**

In the following we define – according to Boyd et al. (2007) – an online social network in particular as a web-based service that enables ”individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system” (Boyd et al. 2007). Thereby, the aspect of networking, i.e. establishing and maintaining relationships between users, is prevailing. However, relationships are not as tangible as those from the real world (Kazienko et al. 2006). Currently, there are a lot of online social networks both for business (e.g. Doostang.com, LinkedIn.com or Xing.com) and for private purposes (e.g. Facebook.com, MySpace.com or StudiVZ.de), focusing on different target groups. Moreover, they differ in size and who can see your profile and how much of it is visible as well (Howard 2008). While most of the key technological features are fairly consistent, the culture that emerges around online social networks varies (Boyd et al. 2007). Furthermore and in addition to the fostering of individual contacts, the community idea is actively lived over forum and group functions.

At the moment many online social networks are basically funded through advertising proceeds. An extension of the business model that includes user fees, is therefore a great challenge for coming years (Pauwels et al. 2008). Critical both regarding the introduction of user fees and in particular the economic valuation of online social networks is the fact that the individual benefit of members considerably depends on the number of members within the online social network. For instance, if a part of the members leaves the online social networks, the individual benefit of the remaining members consequently decreases. On the other hand, every new additional contact of a member raises his or her barrier to leave the network. Such characteristic effects have to be considered when attempting a valuation of online social networks.
3 RELATED WORK

Researchers and practitioners have written a plethora of articles and books on the valuation of firms in general (see Koller et al. 2005, Brealey et al. 2008, Damodaran 2002). However, according to the predominant view in literature standard business valuation approaches are very restricted in their ability to value young, fast growing companies in a dynamic environment, such as Internet companies (see e.g. Gollotto et al. 2003). Reasons are, for instance, the backward orientation using traditional financial balance sheet figures (e.g. liquidation value, substantial value, (adjusted) book value), a lack of acceptance and application in business matters (e.g. real option approach), a lack of academic foundation (e.g. venture capital approach), and the limited history to draw on for future cash flow projections and the handling of just negative cash flows (e.g. discounted cash flow approach).

What makes the economic valuation of online social networks even more difficult is the fact that customers, relationships between customers and the resulting network effects – and therefore intangible assets (especially social capital e.g. Kazienko et al. 2006) – represent a major part of the firm value than assets being currently reflected on the balance sheet. Hence, the value of each customer and the integration of the customer in the online social network as well as the growth of the network have to be considered explicitly to get a reasonable estimate of the firm value. Established standard business valuation models do not sufficiently consider these aspects yet. However, in recent years new approaches (e.g. for the services sector) have been developed, which take into account the value of customers as the most important factor for a firm’s valuation (cf. e.g. Gupta et al. 2004, Bauer et al. 2005, Krafft et al. 2005). Although these models are still based on the discounted cash flow approach, the focus has shifted from the projection of cash flows on a company level to the projection of cash flows obtained from the existing and future customer relationships.

The basic idea behind these valuation methods is measuring the value of the customer base by summing up all discounted cash flows (in and out cash flows) arising from all existing und future customer relationships. The obtained value of the customer base represents the entire value of the discounted operating cash flows of a company. Finally, the value of the customer base “and all cash flows generated from non-operating assets yield the overall value of the company” (Bauer et al. 2005). This change of perspective is quite beneficial for the valuation of online social networks. Although several methods of customer-based valuation have been developed which take into account important aspects, we are not aware of any approach so far that is applicable for the valuation of online social networks. A significant aspect which has to be taken into account when evaluating the customer base of online social networks is considering the number of individual contacts. This is crucial, since the loyalty of a customer strongly depends on the integration into the online social network, as every additional contact of a customer raises the barrier to leave the network (see Algesheimer et al. 2006). The online social network XING.com, for instance, reports that well-connected users have (due to network effects) a higher retention rate (i.e. they are less inclined to leave the network) and lead to a higher activity among users (XING, 2006).

Based on these premises, we develop a model for the economic valuation of online social networks considering the findings from previous research in customer-based valuation and network theory.

4 DESIGN OF THE ECONOMIC MODEL

The long-term value of online social networks is largely determined by the value of the network’s customer relationships, since tangible assets usually play a tangential role. Hence, the online social network’s existing and future customers provide its most reliable source of future revenues. Thereby the value of all existing and future customer relationships is denoted as the customer equity (CE) (Blattberg et al. 1996, Rust et al. 2004). To determine the value of a single customer the widely accepted customer lifetime value (CLV) approach is used, which is similar to the discounted cash flow
approach in firm valuation (see Koller et al. 2005, Damodaran 2002). CLV is defined as the present value of all existing and future cash flows generated by a certain customer (Berger et al. 1998).

Incorporating the CLV approach for determining the value of the online social network, we first partition all existing and future customers into different cohorts $c$ (with $c=0, 1, \ldots$), where $c$ denotes the period in which the customer joined or will join the online social network. Then customers are referred to as $i=1, \ldots, N_c$ for each cohort $c$, whereas all existing customers at the instant of valuation are assigned to cohort $c=0$. With this notation, an online social network’s CE can be expressed as the sum of discounted CLVs of all existing (cohort 0) and future (cohorts 1, 2, \ldots) customers:

$$CE = \sum_{c=0}^{\infty} \frac{\sum_{i=1}^{N_c} CLV_{c,i}}{1 + d^{f}},$$

where $CE$ denotes the total value of all existing and future customer relationships, $CLV_{c,i}$ the CLV of customer $i$ of cohort $c$, $N_c$ the number of customers in cohort $c$ (with $N_c \in \mathbb{N}$) and $d$ the periodical discount rate (with $d \in \mathbb{R}^+)$.

In order to determine the CLV of customer $i$ of cohort $c$ ($CLV_{c,i}$), we obtain the present value at the beginning of period $c$ of all cash flows $CF_{c,i,t} \in \mathbb{R}$ that the online social network expects to receive from the customer over the entire relationship (Berger et al. 1998). Assuming $T_{c,i} \in \mathbb{N}$ as the duration of the customer’s relationship (for existing customers: remaining duration) and index $t$ as the period of the customer relationship (for existing customers: period since the instant of valuation), $CLV_{c,i}$ can be expressed as follows:

$$CLV_{c,i} = \sum_{t=0}^{T_{c,i}} \frac{CF_{c,i,t}}{1 + d^{f}},$$

where $CF_{c,i,t}$ denotes the cash flow in period $t$ of the customer relationship for customer $i$ of cohort $c$ and $T_{c,i}$ the duration of the customer relationship for customer $i$ of cohort $c$.

However, the implementation of Equation (2) is not easy, as it requires detailed information regarding both the future cash flows $CF_{c,i,t}$ and the duration of the customer relationship $T_{c,i}$ for every single (future) customer. Therefore, we use a common approach to bypass the estimation of the concrete duration of the customer relationship $T_{c,i}$ and consider retention rates $r_{c,i,t}$ (cf. for example Berger et al. 1998, Gupta et al. 2004). The retention rate $r_{c,i,t}$ of a customer $i$ of cohort $c$ for a period $t$ (with $t \geq 1$) is defined as the (conditional) probability that the customer remains in the online social network in period $t$, given that the customer has been a member in the previous period ($t-1$). Thus, an undifferentiated approach calculating average retention rates for the whole customer base is often used. To avoid this, we compute individual retention rates for each customer, considering his or her individual degree of interconnectedness in the online social network. Assuming that the online social network is modelled as an undirected graph (see Figure 1), where members are represented by a set of nodes and communication relationships (also known as contacts) by a set of edges linking pairs of nodes (Bampo et al. 2008), the number of incident edges of a node $i$ represents the number of

2 Strictly speaking all determined values are expected values. For simplification we avoid to state all determined values as expected values.

3 An edge respectively a contact between two members exists technical if and only if one of the members has confirmed the contact request of the other.
members customer \(i\) has a connection to or keeps in touch with. This can be expressed in terms of the period \(t\) through the variable \(e_{c,i,t} \in \mathbb{IN}\). Regarding the estimation of the individual retention rate \(r_{c,i,t}\) for customer \(i\) the following requirements have to be fulfilled\(^4\):

R.1 For a customer \(i\) with a larger number of contacts the individual retention rate should be ceteris paribus, higher than for a customer \(j\) with less contacts (lock-in effect). This results in a strict monotone increasing retention rate function of the number of contacts (i.e. \(e_{c,i,t} > e_{c,j,t-1}\) implies \(r_{c,i}(e_{c,i,t}) > r_{c,j}(e_{c,j,t-1})\)).

R.2 An additional contact – i.e. an increase in the number of contacts by one – leads to a ceteris paribus less marginal change in the individual retention rate of customer \(i\) with a larger number of contacts than in the individual retention rate of a customer \(j\) with fewer contacts. This results (in combination with R.1) in a decreasing marginal utility of the number of contacts in regard to the retention rate (i.e. \(e_{c,i,t} > e_{c,j,t-1}\) implies \(r_{c,i}(e_{c,i,t}+1) - r_{c,i}(e_{c,i,t}) < r_{c,j}(e_{c,j,t-1}+1) - r_{c,j}(e_{c,j,t})\)).

These requirements as a starting point, we intensively searched for appropriate functions. The arctangent based formula (3) fulfills both requirements R.1 and R.2 for all numbers of contacts \(e_{c,i,t}\) and can therefore be used for our purpose. We compress the arctangent function (\(\arctan\)) to restrict the obtained values for \(r_{c,i}(e_{c,i,t})\) to the interval \([0; 1]\). Then the individual retention rate for a customer \(i\) of cohort \(c\) in period \(t\) can be defined as a function of the number of contacts as follows:

\[
r_{c,i,t}(e_{c,i,t-1}) = \frac{\arctan(\alpha_{c,t} \cdot e_{c,i,t-1})}{\pi/2},
\]

where \(r_{c,i,t}\) denotes the retention rate for customer \(i\) of cohort \(c\) in period \(t\), \(e_{c,i,t-1}\) the number of contacts of customer \(i\) of cohort \(c\) in period \(t-1\) and \(\alpha_{c,t}\) the calibration factor for the number of contacts in period \(t-1\).

Note that the parameter \(\alpha_{c,t} \in \mathbb{IR}^+\) is used to calibrate the model in regard to the empirical observed average retention rate of the particular period \(t\) of the customer relationship (the empirical observed average retention rate can be interpreted as the fraction of customers that had been members for \(t-1\) periods and remained in the online social network in period \(t\)). Figure 2 illustrates the function \(r_{c,i,t}\) of the number of contacts \(e_{c,i,t-1}\) for some selected values of the calibration factor.

![Figure 2. Retention rate as a function of number of contacts](image)

Taking into account the customers’ individual retention rates \(r_{c,i}(e_{c,i,t})\) we can derive Equation (4) for the CE of an online social network\(^5\). Since the future numbers of contacts of a customer \(i\) are

\[\text{CE} = \sum_{t=0}^{\infty} \frac{r_{c,i,t}}{1+r_{c,i,t}} \cdot (1-e_{c,i,t}).\]

\(^4\) Cf. e.g. Varian (2003), where a detailed literature overview of network effects is given.

\(^5\)
unknown, his or her recent number of contacts has to be used for a forecast. We will demonstrate a corresponding procedure as well as how to determine all other parameters of the model in detail in the following section using the case of XING.com.

\[
CE = \sum_{c=0}^{\infty} \sum_{j=1}^{N_c} \frac{CLV_{c,j}}{(1+d)^j} = \sum_{i=0}^{\infty} \sum_{l=1}^{N} \frac{CF_{c,i,l} \prod_{l=1}^{i} r_{c,i,l} (e_{c,i,l})}{(1+d)^j}
\]

Finally, assigning the approach of Bauer et al. (2005), we have to add up the value of the non-operating assets and to subtract the value of all non customer-specific costs as well as the market value of dept to obtain the corporate value of an online social network. However, according to empirical research, for some companies, the CE is “a useful proxy for firm value” (Gupta et al. 2004). In order to demonstrate the valuation, the following section illustrates the practical application of the model to XING.com, one of the largest and well-known online social networks in Europe.

5 APPLICATION OF THE ECONOMIC MODEL

In this section, we illustrate the application of the model designed in the preceding section and determine the corporate value of the online social network www.xing.com (referred to as XING) on January 1st 2008. As XING is a publicly listed corporation (IPO in 2006), we can resort to data published in the annual reports from 2006 and 2007 for our valuation. This ensures a better transparency and traceability. To avoid a blanket valuation of XING based on average values and the disregard of essential information such as the customers’ individual degree of interconnectedness, we drew a sample of 1,000 customers (Premium Members) on December 31st 2007. By choosing the members randomly (using the search for “random members” provided by XING), it is assured that the sample is really characteristic for the whole network. Based on this data, we determined each customer’s individual CLV considering the individual number of contacts and the initial year of registration in XING. In a final step, we derive the corporate value from the CE.

5.1 The online social network XING

The online social network XING was founded in August 2003 under the name OPEN Business Club and is one of the leading online social networks within the realms of professional online networking platforms in Europe. At the end of 2007, XING counted over 5.7 m members worldwide. These customers use XING to find useful business contacts, new business opportunities, employees, jobs and ideas by posting a profile on the Internet platform. In addition to the free of cost Basic Membership, XING offers a Premium Membership for a monthly fee of € 5.95 which is the backbone of the business model and booked by 362,000 members (December 31st 2007). Besides these membership fees we disregard additional revenue generating sources like banner-ads and e-commerce in a first step as so far these sources of revenue are not crucial to the XING business model.

---

5 As it is not possible to draw a conclusion of the customer’s individual retention rate directly after his or her initial registration to the online social network (t=0) an average value for the rate \( r_{c,i} \) is used. For \( t>1 \) see (3).

6 We assume that all Premium Members joined the network on January 1st within their year of registration. Note that the year of registration is publicly available for each member (cf. www.xing.com).

7 In 2007 94% of XING’s revenues were generated by Premium Memberships (Xing 2007).
5.2 Determination of the parameters of the model

Determination of the number of members

Starting from the IPO at the end of 2006, XING reports a compound annual growth rate (CAGR) of XING’s Premium Members of 64% (Xing 2007). As corporate cash flows are almost exclusively generated by Premium Members, we only consider Premium Members’ cash flows in our model. Nevertheless, Basic Members contribute indirectly to the value of the online social network: On the one hand they are “potential contacts” for Premium Members and therefore increase the attractiveness of the network. On the other hand Basic Members are also “potential Premium Members” in future periods. However, a projection of a compound annual growth rate for Premium Members of 64% seems to be not reasonable. For instance, mature Internet companies like Amazon, Ameritrade, Capital One, eBay, and E*Trade usually show compound annual growth rates in the range of 15% to 25% (Gupta et al. 2004) and a survey of the Global Industry Analysts Group (Xing 2006) projects an annual growth rate of 21.1% for chargeable Internet services in the next years. Moreover, from 2007 to 2008 the number of online social network users increased by 25% to 580 m users worldwide (ComScore 2008). Thus, we adjust the annual growth rate for XING to 25% for the years 2008 to 2010 (cf. Table 1). For the subsequent time period up to 2017, we project a more conservative growth rate of 10%. Beyond the year of 2018 we do not assume any network growth for XING, i.e. numbers of new members and numbers of members leaving the online social network are the same.

<table>
<thead>
<tr>
<th>Year</th>
<th>Premium Members</th>
<th>2006</th>
<th>2007</th>
<th>2008e</th>
<th>2009e</th>
<th>2010e</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>221,000</td>
<td>362,000</td>
<td>452,500</td>
<td>565,625</td>
<td>707,031</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Number of Premium Members of XING 2006 to 2010 (cf. Xing 2007)

Determination of the individual retention rates

As described in section 4 retention rates for \( t \geq 1 \) represent the probability that a Premium Member generating cash flows up to period \( t-1 \) will still be a Premium Member in period \( t \). First of all, we determine average retention rates \( r_{0,t-1} \) for the Premium Members derived from the published fraction of members still remaining \( t \) years (or periods) after their year of registration (cf. Table 2). As the Premium Membership fees for XING are payable in advance, we assume that all customer cash flows are generated at the beginning of a period. Considering this, we derive an average retention rate of 100% \( (r_{0,1}=100\%) \) for the first year of membership (=first period), as all new customers generate cash flows in their first year. For the second year we consequently consider only those customers, that are still Premium Members of the online social network after one year \( (r_{0,2}=82\%) \). Furthermore, the average retention rate for the third year \( r_{0,2,3} \) is determined as 93% \( =76%/82\% \), as 82% of Premium Members remain after their first year of membership in the online social network (paying members in \( t=2 \)) and 76% of Premium Members after their second year of membership (paying members in \( t=3 \)). Starting from year 4 onwards, we assume \( r_{0,t} \) being constantly 99% \( =75%/76\% \).

<table>
<thead>
<tr>
<th>Period ( t ) (year of membership)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4 etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction of remaining Premium Members after period ( t )</td>
<td>82%</td>
<td>76%</td>
<td>75%</td>
<td></td>
</tr>
<tr>
<td>( \emptyset ) retention rate for period ( t ) ( (r_{0,t}) )</td>
<td>100%</td>
<td>82%</td>
<td>93%</td>
<td>99%</td>
</tr>
</tbody>
</table>

Table 2. Rate of remaining Premium Members and average retention rates

In a second step, to account for the individual degree of interconnectedness of each Premium Member, we determine individual retention rates based on the actual number of contacts \( e_{g,t-1} \) of a Premium Member and on the calibration factor for a specific period \( a_{t-1} \) (cf. Equation (3)). For the determination of the calibration factor, we use the average retention rate. In detail, we choose \( a_{t-1} \) so that the overall
average of the individual retention rates for period \(t\) (i.e. the average of all \(r_{0,i,t}(e_{0,i,t-1})\)) corresponds to the observed average retention rate for this year of membership \(r_{0,2,t}\) (cf. Table 2: e.g. 82% for the second year of membership). The results of this calibration for the periods 1 to 3 are illustrated in Table 3. For further periods we do not need this calibration factor, as starting from period 4 onwards we assume constant individual retention rates.

<table>
<thead>
<tr>
<th>Period (t) (year of membership)</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration factor for period (t) ((\alpha_t))</td>
<td>0.0643</td>
<td>0.1560</td>
<td>0.4170</td>
</tr>
</tbody>
</table>

Table 3. Calibration factor for the calculation of the individual retention rates

To calculate the individual retention rates for the existing customers (using Equation (3)) not only for the next period \(t\) (based on \(e_{0,i,t-1}\)) but also for further periods \((t+1, t+2, \ldots)\) we have to forecast the individual number of contacts \((e_{0,i,t}, e_{0,i,t+1}, \ldots)\). For this projection, we calculate in a first step the average number of contacts depending on their individual period of membership \(t\) (e.g. 126 for the second period after registration). Thereon, we derive average growth rates respectively. For example we obtain an average growth rate of 29.9% \((=126/97-1)\) from \(t=1\) to \(t=2\). The rates are presented in Table 4, whereas these are only relevant for the periods 1 to 3.

<table>
<thead>
<tr>
<th>Period (t) (year of membership)</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\emptyset) number of contacts for period (t)</td>
<td>97</td>
<td>126</td>
<td>230</td>
</tr>
<tr>
<td>(\emptyset) growth of number of contacts from period (t-1) to (t)</td>
<td>29.9%</td>
<td>82.5%</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Average growth of the number of contacts

Finally, we can determine the individual retention rates using the calibration factors \(\alpha_t\) (cf. Table 3), the individual information about the year of registration (to determine period \(t\)), the current number of contacts and the average growth rates of contacts (cf. Table 4). The latter is essential for the forecast of the individual number of contacts in the following periods. Table 5 illustrates individual retention rates exemplarily for Premium Members A1 and A2.

<table>
<thead>
<tr>
<th></th>
<th>Year of membership in 2007</th>
<th>Number of contacts 2007</th>
<th>Retention rate 2008e</th>
<th>Number of contacts 2008e</th>
<th>Retention rate 2009e</th>
<th>Number of contacts 2009e</th>
<th>Retention rate 2010e etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>50</td>
<td>80.8%</td>
<td>65</td>
<td>93.7%</td>
<td>119</td>
<td>98.7%</td>
</tr>
<tr>
<td>A2</td>
<td>1</td>
<td>150</td>
<td>93.4%</td>
<td>195</td>
<td>97.9%</td>
<td>356</td>
<td>99.6%</td>
</tr>
</tbody>
</table>

Table 5. Example for the calculation of individual retention rates

For better understanding, we show the calculation following the example of A2: Customer A2 has \(e_{0,A2,1}\)=150 confirmed contacts at the end of his first year of membership. Using the calibration factor \(\alpha_1=0.0643\) and Equation (3) we determine the individual retention rate \(r_{0,A2,1}(e_{0,A2,1})=arctan(0.0643\cdot150)/(\pi/2)=93.4\%\). Hence, the probability that A2 still remains Premium Member in the next period (i.e. in 2008) is 93.4%. For the calculation of the individual retention rate of A2 for 2009, we project the number of contacts by the end of 2008 as follows: \(e_{0,A2,2}=150\cdot(1+29.9\%)\approx195\). This leads to a individual retention rate of \(r_{0,A2,2}(e_{0,A2,2})=arctan(0.1560\cdot195)/(\pi/2)=97.9\%\). The individual retention rates of further years (e.g. for 2010) can be calculated analogically. For the determination of retention rates of future customers, we have to use average numbers of contacts (cf. Table 4) as their individual numbers of contacts are unknown.
Determination of the cash flows

The revenue generated per Premium Member is € 5.95 per month, which accounts to € 71.40 per year. In order to project future cash flows, we determine in a first step the EBITDAM-margin (Earnings Before Interest, Tax, Depreciation, Amortisation, and Marketing) based on figures published in the annual report 2007 (Xing 2007) in the amount of (€ 6.894 m+€ 1.651 m)/€ 19.609 m≈43.6%. Due to the negative margin of the previous year and the long-term rather truncating growth we use a more conservative margin which is extrapolated to a constant figure of 35%. With regard to the amount of marketing spending we have to rely on an assumption, as we could not find precise information in the annual reports about the allocation to existing and to new customers. Therefore we follow the often used rule-of-thumb (cf. Greenberg 2001) and assume that it is five times more expensive to win new customers than to keep existing ones. Taking into account the customer distribution of existing and new customers in our sample of 2007 (55% of the sample are existing customers and 45% are new customers), we allocate marketing-spending of € 8.14/year for new customers and € 1.63/year for existing customers. Following these, we determine the cash flow per Premium Member amounting to $ CF_{c,i,1} = \text{€ 71.40} \times 35\% - \text{€ 8.14} = \text{€ 16.85}$ for the first year of membership and to $ CF_{c,i,t} = \text{€ 71.40} \times 35\% - \text{€ 1.63} = \text{€ 23.36}$ for the following years ($t>1$).

Determination of the discount rate

Due to the dominating equity financing of XING, we assume in a simplified model that the weighted average cost of capital (WACC) is solely based on equity. The cost of equity capital is derived by applying the after-tax CAPM using the average yield of a 10-year European government bond of 4.4% for the base rate (European Central Bank 2007). Applying a common used income tax rate of 35% the tax adjusted risk free rate accounts to 2.86%. Furthermore we assume an expected risk premium of the stock market after taxes of 5.5% (Stehle 2004). Taking into account that online social networks bear more risk than traditional software companies and the fact that XING is relatively small, we increase the published beta-factor of 1.27 for software-companies (Drukarczyk et al. 2007) to 1.48. In summary after applying the after-tax CAPM, we derive a discount rate of 11% (=2.86%+(5.5%\times1.48)).

5.3 Key findings of the application

Applying the economic model to XING, we obtain a CE of € 219.14 m. The value of the existing members sums up to € 63.89 m. In contrast, the value of the future members consists of the discounted cohort values of all acquired members up to the year 2026 (amounting to € 151.77 m) and of the discounted terminal value (amounting to € 3.48 m). Table 6 gives an overview of the key findings. With our results we help investors to make well-founded investment decisions on the basis of public available data. If we take into account that further residuals such as the value of the non customer-specific cash flows, fixed costs that are not attributable to the individual customer and the value of the non-operating assets are negligible, the corporate value equals the CE. Comparing this value with the market capitalization in the amount of € 229.89 m on January 1st 2008, we can state only a slight difference of 4.7% from our findings. This difference can be explained on the one hand by general volatility of the stock market and divergent estimation of valuation parameters by the stock market. On the other hand we neglected additional revenue sources such as advertisements, e-commerce or merchandising products as these sources of revenue are (so far) not crucial to the XING business model. Therefore a stock price of € 44.21 at the instant of valuation seems to be reasonable.

---

8 This extrapolation is consistent with the projected EBITDA-margin according to the XING guidance.
9 As cash flows $ CF_{c,i,t}$ are generated at the beginning of each period, we discount the values contrary to Equation (4) by $t-1$ periods and assign acquisition payments to period $t=1$. In period $t=0$ there are no cash flows.
10 XING shows equity of € 41.5 m and long-term debt of € 0.85 m in 2007 (Xing 2007).
11 From the year 2018 on we assume a net growth of zero relating to the number of members. Therefore the cohort values are almost constant from the year 2027 on, so that we can take a terminal value based on the perpetuity.
<table>
<thead>
<tr>
<th>Year of registration / Cohort</th>
<th>≤2007 / 0</th>
<th>2008 / 1</th>
<th>2009 / 2</th>
<th>2010 / 3</th>
<th>…</th>
<th>2026 / 19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing Premium Members</td>
<td>362,000</td>
<td>452,500</td>
<td>565,625</td>
<td>707,031</td>
<td>…</td>
<td>1,377,804</td>
</tr>
<tr>
<td>New Premium Members</td>
<td>164,710</td>
<td>129,180</td>
<td>147,714</td>
<td>178,530</td>
<td>…</td>
<td>17,651</td>
</tr>
<tr>
<td>Discounted value of cohort [m€]</td>
<td>63.89</td>
<td>20.32</td>
<td>20.93</td>
<td>22.79</td>
<td>…</td>
<td>0.42</td>
</tr>
<tr>
<td>Discounted terminal value [m€]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.48</td>
<td></td>
</tr>
<tr>
<td>Customer equity (CE) [m€]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>219.14</td>
<td></td>
</tr>
</tbody>
</table>

Table 6. Key findings of the application to XING

6 SUMMARY

The increasing economic relevance of online social networks and the numerous recent acquisitions priced at enormous amounts revealed the need for adequate valuation models. However, traditional valuation approaches are restricted in their ability to value young, high-growth online social networks in a dynamic environment. Thus, we developed an economic model for the valuation of online social networks taking into account their specific characteristics. The model allows the evaluation of whether the purchase prices on the market, which recently amounted to millions, are justifiable. The practical application of the model was illustrated by an example of the major European online social network XING. The results show that the model fits quite well, as the results of the model were in the range of the market capitalization of XING at the instant of valuation. For the practical use of the model, we illustrated that, although some assumptions within the application to XING were necessary, public available data in connection with specific market data are sufficient to get reasonable results. These results help investors to make well-founded investment decisions. However, future research has to focus on the application of this approach to other business models of online social networks, as only membership fees are the core basis of XING’s revenue model were currently considered in a first step. For example, for online social networks without membership fees, it is possible that people do not unsubscribe but simply do not use it (passive users with low or no value for the company). In this case, it could be a good idea to check the number of contacts added or the number of accesses in a certain period of time instead of the customers’ number of contacts as an enhancement of our model.

Furthermore, we assumed average retention rates for future customers so far. This assumption could be released by accepting more computational complexity for the determination of the customers’ individual retention rates. This can be achieved through network simulations of the development of the individual number of contacts. We are currently working on taking into account these aspects.

References

References

# Revealing Knowledge Networks from Computer Mediated Communication in Organizations

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0633.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Knowledge Management Systems, Social Network Analysis, Learning, Computer-mediated communication (CMC)</td>
</tr>
</tbody>
</table>
REVEALING KNOWLEDGE NETWORKS FROM COMPUTER MEDIATED COMMUNICATION IN ORGANIZATIONS

Abstract

In today’s knowledge driven economy, knowledge is considered to be the key factor in defining the success of an organization. We have learned that knowledge is residing in the informal network of the organization. Hence, to improve performance, it is the informal knowledge network that should be examined and developed. For this purpose, social network analysis is increasingly applied in business contexts. This is, however, a new domain, which is still in development. This paper aims to aid in this development by researching how representative knowledge networks can be revealed in organizations. While surveying is a common first option to capture an organizational network, this technique may not always be suitable. Communication sources (e.g. e-mail) may provide an alternative, however, we do not know to what extent these sources can represent the actual knowledge network. This paper examines a Dutch IT services organization. Here, a web-survey among the employees baselines the knowledge network, which is compared to 3 communication networks from the same organization, captured by means of e-mail, telephone and SMS (Short Message Service) communication (also known as text messaging or texting). A comparison is made by means of correlating the network matrices and by comparing essential network properties. Findings show that only the e-mail network is significantly representative for the baselined knowledge network. This exercise is exploratory in nature as only one organization is examined, but comprehensive with regard to the richness of data that is available for examination. From our findings we gain insight in the extent to which networks, captured from e-mail, telephone and SMS archives can represent an organizational knowledge network.

Keywords: Informal network, Learning network, Advice network, Social Network Analysis, Mail analysis
1 INTRODUCTION

Knowledge management is widespread in organizations and involves the management of knowledge processes concerning generating, sharing and using knowledge (Davenport & Prusak, 1998). Typically, there are two approaches toward knowledge management: codification and personalization (Hansen, Nohria & Tierney, 1999). The first approach assumes that knowledge within the organization can be made explicit and stored within information systems such as knowledge repositories. Hence, knowledge should be retained so that it is available for re-use. The second approach considers organizations as transactive knowledge systems where knowledge is exchanged between individuals (Wegner, 1987). In this approach, knowledge management focuses on stimulating the interaction between individuals so that knowledge diffusion takes place throughout the organization.

The personalization approach has gained more attention of researchers in recent years. Part of this research focuses on the pattern of knowledge interactions that we refer to as knowledge networks (Hanssen, 2002). A visual representation of such a knowledge network consists of nodes that represent the employees of an organization while the links between these nodes represent the knowledge exchanges between these employees. This contextual specification is the main difference between a social network and a knowledge network: a knowledge network is a social network regarded from a knowledge perspective. A useful method for studying these networks is Social Network Analysis (SNA) as it supports both qualitative and quantitative network research (Wasserman & Faust, 1994; Cross & Parker, 2004). Network data that is required for this research is typically collected using surveys or in some instances using interviews. In a survey, a respondent is asked about his/her knowledge exchange relations with his/her colleagues within the organization. Aspects of the relation that can be of interest include the type of knowledge that is exchanged, how it is exchanged, and how frequently it is exchanged. A disadvantage of surveys is that, especially for large surveys, it takes considerable time from the respondent to fill out the survey and that high response rates are required to do qualitative analysis in case studies (Helms, 2007; Teigland, 2002). Consequently, there is hardly any longitudinal analysis of knowledge networks as it requires that a survey is sent out several times during a certain period.

To overcome these issues, some researchers analyze e-mail traffic within organizations as a source for collecting network data (e.g. Tyler, et al., 2005). The advantage here is that because respondents do not need to fill out a survey, the response rate is 100% and longitudinal analysis becomes possible by examining consecutive time frames. A limitation, however, is that the assumption with analyzing e-mail traffic is that e-mail communication is representative for communication that concerns knowledge exchange. Principally, this idea is not wrong as knowledge is an act of personal interaction by means of communication. But this communication does not only take place through e-mail messages. Face-to-face communication is also an important means for getting knowledge across (Lock Lee & Neff, 2004). Furthermore, e-mail messages are not only used for knowledge exchange but for instance also for coordinating tasks and personal messages (Whittaker & Sidner, 1996). Next to e-mail traffic, other communication sources may be applied to overcome the issues of capturing an informal network by means of surveying. One alternative is by means of tracking telephone conversations within an organization. Also, tracking text messages (texting), further referred to as SMS messages (Short Message Service), may be a valuable source for capturing the informal network of the organization.

This paper intends to examine the extent to which networks that are created from sources such as e-mail, telephone and SMS communication are representative for the knowledge network of an organization. We define our research question as follows:

“To what extent are networks, captured from computer mediated communication sources representative for the knowledge network of an organization?”
The need for this research initiative stems from the fact that the analysis of knowledge networks is becoming a more popular instrument in both science and management. Data capturing by means of surveying knows several limitations that data capturing from computer mediated communication sources does not have. The latter form of data capturing always yields a 100% participation, covering the full network, and it supports longitudinal analysis opportunities. If this research can indicate if and to what extent computer mediated communication sources are suitable to represent knowledge networks, it enables new possibilities for analyzing networks for both science and business.

The remainder of this paper is structured as follows. In section 2 this research is embedded in its theoretical background. Section 3 elaborates on the research method and approach. A case study were networks are captured and compared is introduced and a comparison approach of the networks is elaborated upon. In section 4 the results of the case study are provided, based on the proposed measuring and comparison approach. In section 5 our findings are discussed and conclusions and limitations are provided.

2 THEORETICAL BACKGROUND

A knowledge network concerns the knowledge management activities (e.g. knowledge creation or sharing) that take place among people in an organization. When focusing on a particular knowledge management activity, several types of knowledge networks can be distinguished. Two types of knowledge networks that are often distinguished in literature include learning networks (cf. Skerlavaj, Dimovski, Mrvar & Pahor, 2008) and advice networks (cf. Borgatti & Cross, 2003). Both types of knowledge network are included in this research and described in more detail below. Furthermore, the role of communication in knowledge exchange is briefly explored, which finally leads to the idea to use communication data stored by technology supported media as a source for collecting data on knowledge networks.

The traditional, cognitive view on learning assumes that learning is a mental process within the heads of individuals (Hustad, 2007). In this view, learning is separated from doing in practice and therefore neglects what can be learnt through experience and collaboration. More recent views consider learning as a social process of mutual engagement and relate learning explicitly to practice. Two major accounts in this field are the work by Lave & Wenger (1991) and Brown & Duguid (1991). Based on their studies, they introduced the concept of Situated Learning and Learning in Working respectively. The concept of situated learning is based on master-apprenticeship relationships where apprentices learn and master a practice through legitimate peripheral participation. Peripheral participation means that the apprentices start participating by performing relatively easy tasks and with low risk to the practice, i.e. in the periphery of the practice. During the execution of his tasks, the apprentice engages with and is coached by more experienced practitioners in the practice. If an apprentice is successful, i.e. he is learning, he gets more responsibilities until he finally becomes a master himself.

In this research, we refer to the learning relations between employees in an organization as knowledge networks (Skerlavaj et al., 2008; Palazzolo et al., 2006). Knowledge that is exchanged in learning relations typically involves tacit knowledge and consists of skills, experience and attitudes. This makes knowledge exchange difficult and therefore not every type of knowledge exchange is as effective as another. More active learning approaches, e.g. guided problem solving or guided observation, are preferred as they result in richer knowledge transfers from the master to the apprentice (Leonard & Swap, 2005; Davenport & Prusak, 1998). Richer knowledge transfers results in a higher level of knowledge exchange and therefore contribute to a deeper understanding of the practice. The goal of learning networks is to increase and preserve the knowledge and competence level of employees in the organization, which should ultimately result in a better performance of the organization (Skerlavaj et al., 2008).
Advice networks refer to the idea that employees are dependent on the knowledge of others to execute their job (Cross & Parker, 2004). In that case, it is important to know who knows what and to have access to these people so that they can be asked for advice when needed (Borgatti & Cross, 2003). Cross, Borgatti & Parker (2001) defined five different types of advice that can be sought: Solutions, Meta-knowledge, Problem reformulation, Validation, and Legitimation. These types indicate the intention for seeking advice from a colleague. Providing advice takes place by exchanging knowledge, however, this knowledge exchange is considered to be different from knowledge exchange in a learning network. In the case of an advice network, the goal is to transfer just enough knowledge so that the advice seeker can solve his problem. Hence, more passive learning approaches for knowledge exchange can be used such as directives, rules of thumb or pointers to information sources (Leonard & Swap, 2005). Besides knowing who knows who and having access, it is also important that you can tap into this knowledge quickly (Cross & Parker, 2004). The speed of knowledge exchange is referred to as velocity (Davenport & Prusak, 1998). It is defined as the time between contacting a colleague and finally receiving the requested knowledge from this colleague, either directly from him or via him from another colleague. The higher the velocity of the knowledge exchanges in the advice network, the better it is for the job performance (Hansen, 2002).

Communication is a vital mechanism when it comes to exchanging knowledge in learning as well as advice networks (Palazzolo et al., 2006). It is through communication, consisting of for example discussions and asking questions, that an apprentice can learn from a master. Furthermore, communication is also the basis for asking advice from colleagues. Several media are available for communication and hence for knowledge exchange in learning and advice networks. Basically, two types of media can be distinguished: Face-to-face contact and technology supported media. According to media richness theory (Daft & Lengel, 1986), the selection of the right medium depends on the richness of the knowledge that needs to be exchanged. Consequently, it might be wise to use different media for learning and asking for advice as the richness of the knowledge transfer in both cases is different. Research shows that people have a strong preference for face-to-face contact when it comes to exchanging knowledge (Cross et al., 2001; Smith & McKeen, 2003; Lock Lee & Neff, 2004). All studies reported that computer supported media such as e-mail, telephone or intranet have a lower preference. It is not just that face-to-face contact is the best medium per se but also that other factors are involved why people prefer face-to-face contact. These factors include trust and social cues for example (Wenger, McDermott & Snyder, 2002; Hooff, De Ridder & Aukema, 2004). But it is not just a matter of selecting the best medium, but using several media at the same time as they reinforce and support each other (Davenport & Prusak, 1998). Hence, it is likely that people will use several computer supported media to support face-to-face contact. Based on this assumption, it might be worthwhile to explore to what extent data about knowledge network relations can be retrieved from the computer supported media that store data about communication.

E-mail is a computer supported medium that is more frequently used to capture the social network of an organization (e.g. Tyler, Wilkinson & Huberman, 2005) and a variety of tools exist that support capturing a social network from e-mail data (e.g.: Gloor (2004), Edwards (2005), Mika (2005) and Viegas, Golder & Donath (2006)). E-mail is argued to be a plausible source for capturing a social network. Farnham, Portnoy & Turski (2004) for example reveal that the people we mail most are also the people we tend to work with most. This does, however, not argue for the fact whether e-mail data may also be applied to capture a knowledge network. Hence, it is relevant to investigate to what extent e-mail data, along with other computer mediated communication sources, such as telephone communication and SMS messages, are plausible as sources to capture the knowledge network of an organization.
3 RESEARCH METHOD

3.1 Approach & Sample

In order to examine the research questions, a comparison was made between an advice and a learning network on the one hand and 3 communication networks on the other hand, derived from e-mail traffic, mobile phone calls and SMS messages. The choice for these networks stems from the availability of data in our sample. A case study was conducted in a Dutch consultancy firm in the IT services sector. The firm consists of 68 employees, of which 41 were included in our sample due to their participation in the survey. Of these 41 employees, 10% is female. Moreover, the sample is divided over 6 functional areas (e.g. management, consultancy) and 3 levels of seniority (junior, medior, senior). The case study was conducted in June 2008, a non-holiday period, boosting the representativeness of the networks while no colleagues were unavailable. The scope of the data collection was 1 month. Most employees in our sample are geographically separated in their daily work. Therefore, these employees heavily rely on computer mediated communication media such as e-mail, telephone and SMS to communicate with each other.

3.2 Data collection

The data for the advice and the learning network was collected by means of an online web-survey that was specifically created for this research. In the survey, employees were asked to indicate to what extent they give advice to each other and to what extent they learn from others. Both questions could score on a 5-point scale, ranging from “sporadically” to “intensively”. From the survey results, two adjacency matrices were created: an advice network and a learning network. The scores that each respondent provided for 1..n colleagues were placed in the cells of an adjacency matrix where each respondent is represented by a row and each colleague by a column. The cells were the row and column match the same person were not taken into account. While no advice or learning (score: 0) was also denoted in the adjacency matrix, a matrix is always squared. Eventually, before the calculations were performed, the 2 matrices were recoded. The 5 points in the scale were reduced to 4 points, cancelling out the 1 on the scale, leaving a scale of 2-5. This was done after an examination of the scores in the web-survey. About 50% of the respondents indicated a connection to all other respondents of at least 1 on the scale, meaning that these respondents indicated that they at least sporadically contacted all other respondents. The other 50% of the respondents did not indicate this connection of at least 1 on the scale to all other respondents. This distortion was solved by removing the 1 on the scale. The communication networks (e-mail, telephone and SMS) were not recoded.

For the respondents that participated in the survey, we also collected data about their e-mail, telephone and SMS communication. For this purpose, the Exchange server was accessed to extract one month of e-mail communication data and mobile phone bills were accessed to extract telephone and SMS communication data. Both the data from the Exchange server and the mobile phone bills were applied in a data mining application that was specifically designed and developed for this research, called ESNE: E-mail Social Network Extraction. The application uses the header data of e-mail messages (the “from” and “to” headers). The content of the messages (body data) was discarded due to privacy regulations in our organization of study. Also, the application supports filtering of the e-mail messages. Filtering was applied in various ways. First of all, the data was filtered for a specific time frame of 1 month. Moreover, an upper bound was defined for mass e-mails. Messages that were sent to over 20 recipients were filtered out from our research. It is argued that these e-mails do not represent personal communication as is the case in an advice and a learning network. No lower bound was defined. Different from other tools, our application also filtered out automatically generated messages (e.g. an out-of-office reply) and all senders and recipients external to the company, leaving
only the communication between employees, as again is the case in an advice and a learning network. Similar to other tools (e.g. Holzer, Malin & Sweeney (2005) and Bird et al. (2006)), the application also supports finding aliases. The application scanned all employees on their first and last name and provided possible aliases for each employee (i.e. different mail addresses). All possible aliases were manually linked to an employee or discarded by a manager from the organization. The application itself does not support the visualization of a social network from the data but it is capable of creating adjacency matrices from the e-mail data in the format of the social network analysis software used in this research (.ntf file to import in Cyram Netminer 3.30a). The .ntf file contains both the adjacency matrix and an attribute table containing actor attributes that were derived from employee data managed in the application. This way, an adjacency matrix was created from the filtered e-mail data (including aliases) from June 2008. The weight in the adjacency matrix is expressed by the amount of e-mail messages exchanged between two actors.

Likewise, two adjacency matrices were created from the telephone and SMS data. The application filtered the telephone data to exclude all but mobile phone calls from and to the employees of the focal organization. Also, calls that lasted less than 10 seconds were filtered out. Concerning the SMS data, only the messages that were sent from and to employees were included. The weight in both matrices is again expressed by the amount of calls or messages exchanged between two actors. All matrices i.e. the advice and learning matrices from the web-survey and the e-mail, telephone and SMS matrices from the analysis tool, were manually combined to one file containing the 5 matrices, identical in size and actors, and one attribute table to be used in our social network analysis software.

3.3 Measures

In order to investigate the differences between the networks, a means to compare the networks is required. In the comparison process, the networks that were captured by means of the web-survey were treated as a baseline. It is argued that these networks represent the actual advice and learning network of the organization, as these are the only networks where the employees indicated their relations themselves. In comparison: the e-mail, telephone and SMS networks are a result of communication, moreover than an indication of relations by the employees themselves.

For the comparison between the networks, a two-way comparison format is proposed. First, the networks are compared by means of correlating the matrices. Correlation coefficients can indicate the extent to which the matrices of the networks are similar to each other. The correlation coefficients can be calculated over either the non-weighed (binary) data and the weighed data. The first calculation indicates similarity when only observing whether actors are tied or not. The second calculation indicates the similarity when also taking into account the weight of links between actors. The correlation coefficients for the non-weighed matrices are calculated using the Jaccard coefficient. This coefficient is based on only the existing relations in a matrix and neglects the non-existing relations (i.e. it matches the 1 and neglects the 0). For the weighed matrices, the Pearson’s correlation coefficient is applied. For all correlation measures, a QAP permutation test (Borgatti, Everett & Freeman, 1992) is applied with 2,500 iterations in order to support the significance of the correlation scores.

In the second way of comparing the networks from our comparison format, the networks are compared by means of calculating and comparing network properties that tell something about the structure of the networks. By comparing the networks by means of their properties, further insight can be gained in the equivalence of the networks. For comparing the network properties, we use a number of well-known network measures (derived from Hanneman & Riddle (2005)) that cover different aspects of network structure. An overview of the the network properties that are examined is provided in table 1. For a more elaborate description of the network properties, the reader is referred to Hanneman & Riddle (2005)).
Table 1. Network properties used to compare the networks

<table>
<thead>
<tr>
<th>Basic Demographics</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Links</td>
<td>Total amount of links that exist between the actors in the network (e.g., 50)</td>
</tr>
<tr>
<td>Average Degree</td>
<td>Average amount of links per actor in the network (e.g., 5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Connection</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Distance</td>
<td>Network average of geographic path distance between 2 actors in the network (e.g., 2,2)</td>
</tr>
<tr>
<td>Diameter</td>
<td>Network maximum of geographic path distance between 2 actors in the network (e.g., 5)</td>
</tr>
<tr>
<td>Density</td>
<td>Ratio of existing links relative to the amount of possible links (e.g., 35%)</td>
</tr>
<tr>
<td>Connectedness</td>
<td>Extent to which the network is a single component (e.g., 35%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Embedding</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reciprocity (Dyad method)</td>
<td>Extent to which relations are dyadic (e.g., 35%)</td>
</tr>
<tr>
<td>Transitivity</td>
<td>Extent to which relations are triangular (e.g., 35%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Centrality</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Degree Centralization</td>
<td>Degree of variance in in-degree centrality opposed to a perfect star network (e.g., 35%)</td>
</tr>
<tr>
<td>Out-Degree Centralization</td>
<td>Degree of variance in out-degree centrality opposed to a perfect star network (e.g., 35%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clustering</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering Coefficient</td>
<td>Extent to which connections overlap per actor (my link is also your link) (e.g., 35%)</td>
</tr>
</tbody>
</table>

4 RESULTS

4.1 Matrix correlations

In table 2 and 3 below, an overview is provided of the correlation coefficients between all five networks. For the QAP permutation test, in all cases the $p >= observed$ values were .000, indicating that all results are significant. For the Pearson correlations of the weighed matrices, the significance is expressed by either * or **.

As can be derived from table 2 and 3, the correlation coefficient between the advice and learning network indicates that both networks have no specifically high overlap. From this fact we may conclude that it is plausible to interpret the correlations of the e-mail, telephone and SMS networks to the advice network separate from the correlations of those communication networks to the learning network. A clear decreasing trend is visible for the similarity between the survey networks (advice and learning) and the communication networks, decreasing from mail, to telephone, to SMS. This trend applies for both the non-weighed and weighed data and for both the advice and learning network. Another result from the calculations is the fact that the weighed communication network matrices show higher scores than the non-weighed communication network matrices, meaning that e.g. the e-mail network is more similar to the advice and learning network if the weight of the links is taken into account. The e-mail, telephone and SMS networks tend to be more similar to the advice network than to the learning network. However, these differences are only minor.
Table 2: Jaccard coefficient calculations between the non-weighed networks

<table>
<thead>
<tr>
<th></th>
<th>Advice</th>
<th>Learning</th>
<th>Mail</th>
<th>Phone</th>
<th>SMS</th>
<th>Extended</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advice</td>
<td>.376</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning</td>
<td>.409</td>
<td>.397</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mail</td>
<td>.212</td>
<td>.212</td>
<td>.312</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phone</td>
<td>.074</td>
<td>.079</td>
<td>.085</td>
<td>.202</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMS</td>
<td>.389</td>
<td>.386</td>
<td>.502</td>
<td>.184</td>
<td>.054</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Pearson correlation coefficient calculations between the weighed networks

<table>
<thead>
<tr>
<th></th>
<th>Advice</th>
<th>Learning</th>
<th>Mail</th>
<th>Phone</th>
<th>SMS</th>
<th>Extended</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advice</td>
<td>.458**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning</td>
<td>.419*</td>
<td>.383*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mail</td>
<td>.287</td>
<td>.230</td>
<td>.487**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phone</td>
<td>.132</td>
<td>.098</td>
<td>.199</td>
<td>.413*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMS</td>
<td>.499**</td>
<td>.456**</td>
<td>.830**</td>
<td>.438*</td>
<td>.214</td>
<td></td>
</tr>
</tbody>
</table>

4.2 Extended mail network

From our findings from the matrix correlations, we learned that of all communication networks, the e-mail network is the most similar to the advice and learning network. Therefore we decided to extend the research concerning the similarity between the advice and learning network on the one hand and the e-mail network on the other hand. It is argued that the timeframe of one month that is applied in this research may yield that the e-mail network is less similar to the advice and learning networks than it could be if the timeframe is extended. Therefore, the timeframe of the e-mail network was extended to 6 months, from January to June 2008. From this data, a new network, the extended e-mail network, was constructed using the data mining application (ESNE). The extended e-mail network was then correlated against the advice and learning network. Again, the $p >= observed$ values were .000 in all cases, indicating that all results are significant. The results from these calculations are incorporated in table 2 and 3 above.

The results show a different outcome for the non-weighed than for the weighed data. The non-weighed data shows that the extended e-mail network is less similar to the original e-mail network that consists of only one month of data. However, the difference between the scores from the non-weighed and the weighed data are minimal. On the other hand, the weighed data does indeed show that the extended e-mail network is more similar to both the advice and learning networks than the original e-mail data of one month.

4.3 Network properties

In table 4 an overview is provided of the scores from 11 network properties that are commonly used in social network analysis. The scores are calculated for all basic weighed networks (i.e. the extended e-mail network is excluded from our scope here).
The network properties displayed in table 4 show different scores for the different networks. For most network properties, the same trend can be visualized as with the matrix correlations. Of all communication networks, it is again the e-mail network that most resembles the advice and learning networks. The telephone network is less similar to the advice and learning networks and the SMS network even less. This trend does, however, not apply for the reciprocity network property. An explanation for the deviating reciprocity value in the e-mail network may be that of all forms of communication studied in this research, e-mail is typically a medium of bilateral communication, i.e. an e-mail is typically replied to with another e-mail. This way of communicating is not found in media such as telephony or in an advice or learning network, where one instance of communication already involves both sender and receiver. The matrix correlations showed that the e-mail network, which is the communication network that most resembles the surveyed networks, is more similar to the advice network than to the learning network. In the case of the network properties, however, the e-mail network is more similar to the learning network than to the advice network. Again, these differences are only minor. This does, however, not apply for the centralization properties, where the in-degree centralization index of the e-mail network is far more similar to the advice network and the out-degree centralization index of the e-mail network is far more similar to the learning network.

<table>
<thead>
<tr>
<th>Network Properties</th>
<th>Advice</th>
<th>Learning</th>
<th>Mail</th>
<th>Phone</th>
<th>SMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Links</td>
<td>464</td>
<td>465</td>
<td>493</td>
<td>193</td>
<td>57</td>
</tr>
<tr>
<td>Average Degree</td>
<td>11</td>
<td>11</td>
<td>12</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Mean Distance</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td>2.1</td>
<td>2.4</td>
</tr>
<tr>
<td>Diameter</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Density</td>
<td>28%</td>
<td>28%</td>
<td>30%</td>
<td>12%</td>
<td>4%</td>
</tr>
<tr>
<td>Connectedness</td>
<td>81%</td>
<td>90%</td>
<td>100%</td>
<td>37%</td>
<td>6%</td>
</tr>
<tr>
<td>Reciprocity (Dyad method)</td>
<td>36%</td>
<td>36%</td>
<td>76%</td>
<td>30%</td>
<td>43%</td>
</tr>
<tr>
<td>Transitivity</td>
<td>54%</td>
<td>54%</td>
<td>53%</td>
<td>35%</td>
<td>26%</td>
</tr>
<tr>
<td>In-Degree Centralization</td>
<td>38%</td>
<td>71%</td>
<td>44%</td>
<td>39%</td>
<td>14%</td>
</tr>
<tr>
<td>Out-Degree Centralization</td>
<td>71%</td>
<td>48%</td>
<td>46%</td>
<td>37%</td>
<td>12%</td>
</tr>
<tr>
<td>Clustering Coefficient</td>
<td>72%</td>
<td>71%</td>
<td>69%</td>
<td>59%</td>
<td>21%</td>
</tr>
</tbody>
</table>

Table 4: Network properties and scores

5 DISCUSSION & CONCLUSIONS

5.1 Discussion & conclusions

This paper intended to examine to what extent computer mediated communication in organizations may be representative sources for capturing and analyzing organizational knowledge networks. As the analysis of organizational networks is becoming an increasingly more important management instrument, there is a need to continuously improve the analysis method. While capturing network data by means of surveying knows several limitations, the analysis of computer mediated communication...
sources may provide a promising alternative. In the Dutch IT services organization were we conducted our case study, 6 networks were captured: the advice and learning network by means of surveying and 4 communication networks from e-mail traffic (1 month and 6 months in length), telephone conversations and SMS (text) messages. Consequently, a comparison between the networks was conducted by means of two comparison approaches: correlating the network matrices and comparing basic network properties, applied in social network analysis. Results from the matrix correlations show that only the e-mail network (both the original and the extended network) is significantly representative for both the advice and learning network. The telephone and SMS network are only slightly representative to the advice and learning network. The results from the network property calculations support the findings from the matrix correlations to a large extent. A secondary finding is the fact that the extended e-mail network has a mixed effect on the quality of representativeness of the original e-mail network. While the weighed extended e-mail network is more representative to the learning and advice network as its weights are more representative, the non-weighed extended e-mail network is less representative for the advice and learning network. This is probably caused by a distortion due to introducing new relations in the extended e-mail network that were not present in the original e-mail network and also not in the advice and learning network. This distortion may be removed by defining a minimum amount of messages sent between two actors before including the relation in the network (recoding the data). Although this examination was only conducted in one organization, the richness of the data proved worth to investigate how the computer mediated communication networks resemble the advice and learning network that provided a baseline for the knowledge networks in the organization of our case study. We may conclude from our research that a network based on e-mail traffic proves to be a valid representation for a knowledge network when conducting knowledge network analysis if surveying the focal organization is not an option.

5.2 Limitations

Probably the most important limitation on our research is the fact that this research was conducted in only one organization. Therefore, we cannot conclude on the repeatability of our findings in other organizations. It is, however, imaginable that it is quite a challenge to find multiple organizations that are willing to participate in a survey and also provide access to their Exchange server (e-mail) and to provide mobile phone bills (telephone and SMS). This is the reason that this research was addressed as an exploratory case study research. Another limitation is that the scales used in the data capturing process for the 6 matrices were not normalized. The advice and learning networks consist of a 5-point scale, whereas the e-mail, telephone and SMS networks were based on the amount of messages or calls. However, the analysis of the matrix correlations lowers the effect of this limitation as the coefficients calculated over both the weighed and non-weighed networks are similar to a large extent. A final limitation is the fact that the telephone and SMS networks could not be captured for 100% of the data. This is due to the fact that the telephone bills were partially concealed. This was done by the provider of the telephone network and cannot be altered. In the telephone network 20% of the data was removed and in the SMS network 9% of the data was removed due to concealed data (anonymous calls and SMS messages). Because the blocked entries on the telephone bills are spread over all of the actors, it is false to conclude that the measured indicators deviate by 20% and 9%. The percentages only decrease the amount of calls and messages, they do not remove 20% and 9% of the links. From the existence of concealed entries it is, however, evident and fully in line with the findings of this research that telephone and SMS networks should not be favored as a surrogate for the advice and learning network of an organization.
5.3 Acknowledgements

We would sincerely like to thank Professor Robert A. Hanneman (University of California, Riverside) for his help on structural equivalence measures for matrices and on the interpretation of the significance of the results of our paper.

References


Proceedings ECIS 2009


## Analyzing Community Contributions to the Development of Community Wireless Networks

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0709.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Social Network Analysis, Social computing, Socio-technical theory, Social inclusion</td>
</tr>
</tbody>
</table>
ANALYZING COMMUNITY CONTRIBUTIONS TO THE DEVELOPMENT OF COMMUNITY WIRELESS NETWORKS

Abdelaal, Abdelnasser and Ali, Hesham, College of Information Science and Technology
University of Nebraska at Omaha, Omaha, NE 68182, {aabdelaal, hali}@mail.unomaha.edu

Abstract

Community wireless networks (CWNs) have emerged as collective actions achieved by many communities worldwide to access the information highway. Developing autonomous CWNs depends, in large part, on community contributions that may include time, money, efforts, expertise and computer resources. However, there is a lack of instruments for measuring such contributions, as well as the outcomes of these networks. This study uses the social network analysis analytical approach to model, measure and analyze community contributions in the development of their wireless infrastructures. In particular, we model community contributions as a two-mode (or bipartite) graph composed of two sets of nodes: the first represents a set of community contributors and the other represents a set of wireless networks. The edges between these two sets stand for the inputs of contributors. Their contributions include volunteering time and manpower, sharing their wireless nodes with community members, donating money, donating hardware, providing technical support, and developing open source software for the network. The model is used to analyze these tangible and intangible forms of contributions. We hope this study provides a better understanding and sounder measurement of the role of communities in developing these emerging common wireless infrastructures and similar digital collective actions.

Keywords: community wireless networks, social network analysis, collective actions, common systems

1. INTRODUCTION

The phenomenal growth of community wireless networks (CWNs) has captured the attention of a broad range of academicians and professionals in several disciplines and practical fields. CWNs are socio-technical crucibles where community resources are shared, mobilized and reproduced to build a common telecommunication infrastructure. Numerous communities have built autonomous wireless infrastructures using their own resources to join the information society. The majority of these networks provide free, or affordable, Internet to community members. In other words, CWNs have been built by the community, as an outgrowth of private WLANs, for the community. We argue that CWNs could serve as a third solution, in addition to the private and public solutions, for achieving digital inclusion (Abdelaal and Ali 2009a). Consequently, scholars recognize them as wireless commons achieved by collective actions (Abdelaal and Ali 2009b; Bina and Giaglis 2006; Damsgaard, Parikh, and Rao 2006; and Sicker, Grunwald, Anderson, Doerr, Munsinger and Sheth 2006 and Negroponte 2002). The term collective actions refers to the voluntary collaboration of group members to construct a common project (Olson 1971:pp7). Literally, a collective action requires a collectivity, or a group of people, which collaborate to achieve an action (Sakurai 2002). Along this notion, developing collective wireless networks, require a motivated and committed group of agents, sometimes organizations, with shared objectives who contribute to these networks for the well-being of the community at large (Abdelaal and Ali 2007; Quinn 2006; Bina and Giaglis 2006 and Camponovo and Cerutti 2005). These agents may include local residents, open source software (OSS) developers, students, technology vendors, municipalities, civil rights activists, developing agencies and researchers.

Agents of a social structure are those who participant in converting both tangible and intangible resources into negotiable offerings (Allee 2008). The agents of CWNs could be classified into three categories: beneficiaries, contributors, and isolated actors (Abdelaal and Ali 2008). Beneficiaries are individuals who gain from these collective networks. Their benefits may include obtaining free Internet access, donated old PCs, technical
Expertise, exposure, and/or other social gains. Contributors are those who donate to the network. Their contributions may include providing manpower or technical skills, donating money or used equipment, sharing their wireless nodes with others or boosting the publicity of the project (Quinn 2006 and Drunen, Koolhaas, Schuurmans and Vijn 2003). Scientists and practitioners emphasize that engaging community members and employing their resources are essential to building and sustaining these community-driven ventures (Siochru and Girard 2005; and Sandvig, Young, and Meinrath 2004 and Bina and Giaglis 2006). Isolated actors, within the proximity of the network, are those who neither benefit from nor contribute to the network. These isolated actors include the “have-nots” or those who are digitally disconnected. The subject matter of this paper is the contributions of the agents of CWNs to the development and operation of these collective actions. CWNs provide community members the opportunity to contribute (e.g. money, time, expertise, computer resources) and develop innovative applications, software, and hardware for their communities (Powell 2006). In this study, we use the term contributions to refer to the tangible and intangible resources or assets provided by community members to build and maintain CWNs.

Despite the exponential growth of these networks, closely linked to the deregulation of the 2.4GHz spectrum, there is a lack of empirical studies and quantitative methods that investigate their resources and outcomes. For instance, previous attempts, mostly case studies, have conceptual and instrumental limitations. Our objective is to complement previous literature and develop a quantitative model, supported by empirical evidence, for community contributions to CWNs. Hopefully this model would guide future research and help policy makers and community developers to promote this innovation. Such collective projects, we believe, require new perspectives to treat their resources and outcomes for the following reasons:

1) CWNs are built and operated by flows of tangible (e.g. money, hardware) and intangible (e.g. knowledge, efforts, and software) contributions from community members. Since these contributions are provided for free, it is difficult to measure and aggregate them using conventional measures. Therefore, it is hard to comprehend the contributions of these resources to the overall cost of the project;
2) Most of these networks are not official organizations, but rather loosely affiliated groups or collectivities;
3) As with other collectivities and collaboration groups, it is not easy to track the contributions and benefits of participants. This problem is exacerbated if the diversity of stakeholders of CWNs is considered; and
4) Community contributions, and the benefits of CWNs, are usually provided outside the market mechanism. Therefore, it would be useful to convert them into negotiable market values in order for researchers to control such factors and for practitioners to manage their networks.

Therefore, paying special attention to this emerging form of collective actions is warranted and developing quantitative models or metrics for their variables is necessary to advance this intellectual stream and accumulate related knowledge. The developed models, or artifacts, should be efficient and general enough to encompass different types of involved actors and acknowledge their contributions and benefits. The key question we address is:

How can we measure the role of communities in developing their wireless networks?

Social network analysis (SNA), depends on graph theory, is used to answer this question. Answering this research question is important for understanding how community resources are mobilized and reproduced to achieve collective actions, taking advantage of digital innovations. This would be a first step towards advancing CWNs as a third solution to integrate the society in the information age. This analytical approach has been used to study a wide-range of similar networks such as computer, social, biological, chemical and transportation systems (Hanneman and Riddle 2005 and Peter, Scott, and Wasserman 2004). Sufficient modeling and analytical details are provided so that researchers and practitioners can apply the adopted research approach in their own investigations. This study is part of a larger research project whose objective is to investigate the tension between these emerging socio-technical networks and their societies. Investigating the tension between "structure and agency" or "macro and micro" is one of the key intellectual themes in sociological inquiry (Hanneman and Riddle 2005). It is important to note that this paper is an extension to a previous work that proposes a conceptual framework of community contributions to CWNs (Abdelaal, Ali and Khazanchi 2009b). In this previous work, we
treat community contributions as collective actions mobilized by the social capital in the community. In another work, we explore the role of these networks in achieving digital inclusion for their communities (Abdelaal and Ali 2009a). In particular, we also discuss their size, capacity, service charging, and finances. We also provide a thorough review of literature related to their outcomes.

The rest of this paper is organized as follows: section 2 discusses the literature of the role of communities in building their wireless networks and briefly reviews similar work that uses SNA to study complex networks. In section 3, we present our research methodology. In section 4, we propose our model. In the fifth section, we use the proposed model to compare and classify CWNs. Section 6 discusses the significance of research and suggest ideas for future work.

2. LITERATURE REVIEW

The notion of community and information commons have attracted the attention of a large number of Internet scientists (Lesser, Fontaine and Slusher 2000). Although there is an extensive research on collaboration in the domain of the Internet, very little empirical work has focused on the creation of telecommunication infrastructures as outcomes of such collaboration. In other words, a small body of research, mostly qualitative, addresses how CWNs are built. For instance, Damsgaard, Parikh, and Rao (2006) describe CWNs as wireless commons. They define wireless commons as a group which share their private WLANs to create a common resource and open it for others. The authors point out that the group members might have conflicts of interest or they may overuse such a common resource. Misusing or overusing common resources is a classical problem in social science known of the tragedy of commons. They discuss the causes and the preventions of this tragedy in the domain of the wireless commons or CWNs. Sicker, Grunwald, Anderson, Doerr, Munsinger and Sheth (2006) use simulation tools to study the relationship between the network capacity and the number of users to examine if the wireless commons are misused. They use the density, usage patterns, environmental conditions, and application demand as assessment parameters. Battiti, Cigno, Sabel, Orava, and Pehrson (2005) call these networks open access networks. According to the authors, the main advantage of CWNs is fostering win-win partnerships between actors, increasing freedom of choice for users and providers, decreasing costs and expanding service coverage. Along this line of research, we also consider CWNs as a form of collective or common projects (Abdelaal and Ali 2009a and Abdelaal, Ali and Khazanchi 2009b).

Best and Maclay (2002) identify six factors that must be considered for designing a self-sustainable CWN: costs, revenue, networks, business models, policy, and capacity. Camponovo and Cerutti (2005) classify the actors of CWNs into four categories: (1) those who share their access points with others; (2) the beneficiaries of these networks who obtain free Internet access; (3) the Internet service providers; and (4) the regulatory authority that sets the regulations to govern the use of the spectrum. Bina and Giaglis (2006) use the concept of collective actions to explore the motivations of participants of CWNs. Mandviwalla, Jain, Fesenmaier, Smith, Weinberg, and Meyers (2006) identify the following stakeholders: underserved individuals, municipalities, schools, small businesses, nonprofit organizations, community groups, utility companies, healthcare providers, and state and federal governments. This study focuses only on the inputs of such stakeholder to CWNs. We discuss the outcomes of these networks in Abdelaal and Ali (2009).

Scholars point out that such inputs may include volunteerism, money and hardware donation, providing technical support and developing OSS for the network.

For instance, Quinn (2006) discusses the role of community engagement, volunteerism, OSS, and donated computer hardware in the development of three CWNs in the Chicago area (United States). The author also
proposes a guide to help practitioners in this regard. Similarly, Drunen, Koolhaas, Schuurmans and Vijn (2003) point out the importance of community contributions in building and maintaining the Wireless Leiden network in the Netherlands. According to the authors, such contributions include low-cost network interfaces, OSS, home-built antennas, and voluntary manpower and technical skills. We model and analyze these types of community resources used to build the investigated CWNs. Using a case study, Shin and Venkatesh (2008) suggest that community participation should continue through all the developing stages is essential for the sustainability of CWNs. They also identify four groups of stakeholders emerged around this network: the Kutztown municipal authority, technology vendors, the project team and the community. They argue that inputs of citizens, as lay designers, are important for the success of CWNs. Powell (2006) discusses contributions of volunteers to British Columbia Wireless. Their contributions include site surveys, hardware hacking, software and content development, technical support and legal and regulatory research.

Reviewing previous literature, we highlight the following drawbacks:
1) The types of community contributions, to CWNs, adduced in previous literature are plausible and supported by evidence from case studies. However, they are not categorized and attributed in terms of quantitative variables that could be controlled. This is important for guiding future research, allowing for comparison, and accumulating knowledge; and
2) There are no conventional instruments or common approaches to measure the tangible or intangible resources and outcomes of CWNs.

Similar to our analytical approach, Jackson (2003) discusses several examples of economic applications of graph theory. These applications include obtaining information about jobs from social contacts, exchanging goods between market actors, and contracting trade agreements. Gale and Kariv (2007) propose a graph model for financial networks where nodes represent traders and weighted edges represent the probabilities of trade between them. Souma, Fujiwara and Aoyama (2005) model the Japanese shareholders network using a directed graph where nodes represent companies and edges represent activities, ownership and governance. Spulber and Yoo (2005) use graph concepts to develop a pricing policy for telecommunication services. Their proposed pricing policy takes into account the impact of changes in one node on the entire system, particularly on the economies of scale. Tesfatsion and Pingle (2003) use graph concepts to examine the effects of a non-employment payoff on network formation and behaviors of workers. The authors model the interactions between networks of workers in a form of a directed graph. The vertices represent the workers and employers. Edges denote to the work flows between workers and employers. In our pilot study, we use data from the Omaha Wireless Group to describe CWNs as socio-technical networks where each network is composed of two graphs: wireless network and a social network served by this wireless network (Abdelaal and Ali 2009b). We argue that the interactions between these two graphs impact different aspects of CWNs. Evidently, graph theoretic concepts and SNA is a sounder analytical approach for studying CWNs as value networks.

2. RESEARCH APPROACH

As with similar emerging phenomena, there are no standard research approaches or widely-recognized empirical studies on CWNs. This is mainly because the CWNs movement is large and diverse, geographically dispersed, and informally structured. In addition, CWNs are not well-defined and practitioners and researchers usually mix between them and similar networks such as municipal wireless networks and commercial Wi-Fi hotspots (Abdelaal and Ali 2009b). It is, therefore, difficult to collect high quality data about the transactions of all actors. To avoid these problems, we collected 2-mode network data using a survey instrument during the annual International Summit for Community Wireless Networks (ISC4WN). The collected data represents a network of “affiliations” between different categories of community contributors, by their inputs to develop the network, and their CWNs. In other words, this data describes which type of contributor (e.g. volunteer, money donor, OSS developer, etc.) is affiliated with which CWN and how much he/she contributes. Another quality of the data is

---

1 The Omaha Wireless Group at http://omahawireless.unomaha.edu/index.html
2 The International Summit for Community Wireless Networks was held in Washington, DC, U.S., May 28th to the 30th, 2008. Its objective was to explore the opportunities and challenges of CWNs.
that it represents 16 networks from different parts of the world. The data describes ties between two sets of nodes at two different levels of analysis: a set of CWNs on the micro level and the community members on the macro level. The third quality of the data lies in reflecting diverse opinions of CWNs developers and advocates who attended this annual summit. We posted the survey online and sent its link to those who could not fill it during the summit. We received 41 responses and eliminated the incomplete ones. The survey questions are designed to ask respondents to choose the types of contributions their networks receive from their communities. We categorized six distinctive facets of contributions (or variables): volunteering time, money donation, hardware donation, technical support, sharing access points, and developing software for the system, as shown in Table 1. These variables are used as proxies for community resources, or contributions. In order to study collective actions in a quantitative manner, we need to identify a collective action to examine and construct variables that represent this action (Sakurai 2002). These variables have been identified based on extensive review of literature, three years of experience working with the Omaha Wireless Group and discussions with leaders of CWNs. Seven individuals examined the clarity and relevance of the survey questions to the measured constructs. The literature and collected data shows that there are other types of contributions such as obtaining the necessary political support and promoting the awareness of the benefits of the project. We excluded such contributions for simplicity. Again, we use graph theory concepts to study how communities, by their contributions, create CWNs. A graph is a mathematical model consisting of two sets V and E. V is a set of nodes called vertices connected by a set of links (or E) called edges. Graphs have been used to represent many similar complex networks and solve related problems (Peter, Scott, and Wasserman 2004). The SNA and the UCINET software is used to visualize, measure and analyze the data. We adopt this approach for its efficiency in representing network data in a compact and systematic manner, suitability for computation processing, capability of using theories and concepts of graph theory and ability to infer patterns of relations between actors in a mathematical manner (Hanneman and Riddle 2005). For more details about concepts, measures and suitable data for 2-mode networks (or bipartite graphs), we refer readers to Hanneman and Riddle (2005) and Borgatti (2008).

4. MODELING THE CWNs-BY-CCONTRIBUTIONS NETWORK

<table>
<thead>
<tr>
<th></th>
<th>Time (h)</th>
<th>Money</th>
<th>Tech. supp.</th>
<th>access points</th>
<th>Hard. Dona.</th>
<th>OSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Court Housing Co-op</td>
<td>5</td>
<td>200</td>
<td>10</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>SeatleWireless</td>
<td>100</td>
<td>5000</td>
<td>50</td>
<td>1000</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>Champain-Urbana</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>AirStream</td>
<td>250</td>
<td>10000</td>
<td>100</td>
<td>2000</td>
<td>50000</td>
</tr>
<tr>
<td>5</td>
<td>PTAWUG</td>
<td>1250</td>
<td>7500</td>
<td>850</td>
<td>250</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>Keur Sedaro</td>
<td>15</td>
<td>30000</td>
<td>20</td>
<td>250000</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>Pretoria Wireless</td>
<td>0</td>
<td>3000</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>Cstle Square WiFi</td>
<td>300</td>
<td>60000</td>
<td>1000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>WUG</td>
<td>500</td>
<td>6000</td>
<td>240</td>
<td>300</td>
<td>25</td>
</tr>
<tr>
<td>10</td>
<td>Cape Town Wireless</td>
<td>20</td>
<td>3895</td>
<td>400</td>
<td>300</td>
<td>500</td>
</tr>
<tr>
<td>11</td>
<td>Nepal Wireless</td>
<td>50</td>
<td>30000</td>
<td>100</td>
<td>1250</td>
<td>1250</td>
</tr>
<tr>
<td>12</td>
<td>Red Libre De Ometepe</td>
<td>25</td>
<td>30000</td>
<td>20</td>
<td>50</td>
<td>750</td>
</tr>
<tr>
<td>13</td>
<td>Jawug</td>
<td>10</td>
<td>389</td>
<td>20</td>
<td>1500</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>NYCwireless</td>
<td>10</td>
<td>0</td>
<td>20</td>
<td>200</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>ZGwireless</td>
<td>100</td>
<td>0</td>
<td>200</td>
<td>5000</td>
<td>50</td>
</tr>
<tr>
<td>16</td>
<td>Digital el Paso</td>
<td>200</td>
<td>50000</td>
<td>400</td>
<td>1500</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1. A matrix of the contributions of CWNs actors

The investigated observations are transformed into a data matrix in which each observation, or a CWN, occupies one row and each variable, or type of contribution, occupies one column. Table 1 summarizes the data collected about these variables. Three kinds of intangible resources, or community contributions, have been distinguished in the realm of CWNs: the cooperation of volunteers, the technical expertise of computer specialists and the
bandwidth shared by those who share their nodes with their neighbors. One of the challenging questions related to intangibles (e.g. human knowledge, software, reputation, political support, and collaboration) is how to convert them into negotiable forms of value (Allee 2008). Since community contributions are provided for free, we estimate their value using their opportunity costs. We assume that the opportunity cost of volunteered time is $5 per hour, the donated hardware (e.g. used computers and access points) is $25 per unit, the technical support is $10 per hour and the shared access point is $50 per unit. This process is called value conversion through which we convert intangible inputs or assets into financial values (Allee 2008). The third column in this table presents the dollar value of volunteered hours per week for the investigated projects. For instance, Court Housing obtains voluntarily time of $5(1x5) and SeattleWireless obtains voluntarily time of $100 (20 hours x 5 dollars). Column four represents the money donations received so far by these networks. Column five represents the dollar value of the weekly technical support provided to these networks by community members. For instance, SeattleWireless obtains 5 hours of technical support whose money value is $50. Similarly, Court Housing has two community members who share their access points, with estimated money value of $100 and received two units of hardware donation of $50 value. Column six represents the money value, as estimated by respondents, of the OSS developed to the network.

Powell (2006) calls networks that depend on the contributions of volunteers a network of aid. However, we view CWNs as value networks composed of social and technical resources contributed by the community. They also generate social, economic and technical outputs for the community. In general, a value network is defined as a group of people who work together via relationships to create public goods or economic value (Allee 2008). We modeled community contributions in a bipartite graph: $G = (U, E)$, where $U = (V U R)$ where $V$ and $R$ are disjoint sets of vertices (nodes). In other words, there is no link either between any pair of nodes with the set $V$ or the set $R$. The set $V$ represents the investigated CWNs and the set $R$ depicts the types of contributors represented by their contributions. The set $E$ depicts the ties between $V$ and $R$. A tie is an ordered pair of nodes $(v_i, r_j)$. These ties represent different forms of community contributions. Each edge has a nonnegative capacity or weight $W$ for all edges $(v_i, r_j) \in E$, as shown in Table 1 and Figure 1. If there is no edge between two nodes $v_i$ and $r_j$ or the weight $w(v_i, r_j) = 0$, then the group of contributors $r_j$ do not support the network $v_i$. Figure 1 depicts the two sets described in our model: a set of CWNs (the red nodes) and a set of community contributions (the green nodes). The size $|R|$ of the set of contributions is 6. The set $R$ represents the inputs, or the dollar value, of community members, discussed earlier, which include money donation, time volunteered, sharing access points, developed software for the system, and provided technical support. The size $|V|$ of the CWNs set is 16 networks. In this model, the set $V$ represents the sinks, or CWNs, and the set $R$ represents the sources of values or set of contributors. Clearly, the rows, in Table 1, represent the sinks and the columns represent the sources of contributions. To illustrate more, Figure 2 demonstrates NYCwireless as a graph of a super sink and three sources of value.

One of the problems with respect to CWNs is finding efficient ways to summarize and visualize data about the transactions between their actors. Figure 1 depicts the investigated CWNs-by-contributions network. Visualizing such collective projects may uncover some of the hidden information about the interactions between their components. For instance, Figure 1 shows the “goingtogetherness” or “correspondence” of types of contributions, which represent contributors, and associated CWNs. This visualization depicts bundles of CWNs-contributions as clusters in the joint space. For instance, the SeattleWireless, Jawug, NYCwireless, Keur Sedaro, Digital el Paso, Cstle Square WiFi, Champaign-Urbana, and Pretoria Wireless are located close to each other. This togetherness indicates the similarity between these two groups of networks. We also visualized the tangible and intangible resources of these networks in forms of negotiable values. Figure 1 shows that hardware donation (HD) is located far away from the other types of contributions. The other five facets of contributions are located close to each other. We will provide more quantitative analysis to such togetherness in the following sections.
Figure 2 shows the NYCwireless as a super sink with three sources of values which are shared access points, technical support, and time. It receives weekly volunteered hours of $10 and $20 dollars of technical support and it has 200 dollars worth of shared access points between community members. It does not receive hardware donations, money donations, or free OSS form the community.

Although we provided monetary values for these facets of contributions in this study, this is not meant to measure the real value of the infrastructure. Instead, we used it only for modeling and analytical purposes in order to help researchers and practitioners to apply our method in their investigations. In other words, the main purpose is to provide a new approach to quantify, map and measure the assets of these networks or the community contributions used to build them. It is important to measure these resources so that network managers can manage them. We can also visualize how the wireless nodes are geographically distributed and to what extent these networks are embedded in their communities. Using this approach would be more useful for an in-depth case study that focuses on one, or few, network to have a clearer understanding and a thorough estimation of the resources and assets of this specific network.
5. EMPIRICAL ANALYSIS

The proposed graph theoretic model represents a valuable quantitative tool to address key issues related to the resources of CWNs. Following is an empirical analysis of community contributions with the help of this model.

5.1 A comparative study of CWNs

In this experiment we used the developed model to compare different CWNs based on the type and value of community contributions. In particular, we used the similarity function of the UCINET software to compare these networks using the sum of the weights, or the capacity, of their incoming edges or \( \text{in} (E) \). The similarity function measures to what extent two CWNs are similar in this regard. Table 2 shows the similarity between the investigated networks. A similarity of 1 means the two networks are identical (we compare the networks to themselves), and -1 means they are totally opposite to each other in terms of types of involved actors and their contributions. The similarity between the Court Housing and SeattleWireless networks, 0.94, means they are very similar to each other based on the types and value of community contributions. A high dissimilarity between two networks may indicate a good opportunity of collaboration between them. They could exchange knowledge, expertise or OSS. For instance, there are collaboration opportunities between the Champian-Urbana and Cape Town networks, as the similarity between them is -.92. We can also compare the set of contributions to each other using the same approach, and this one of our future work.

Another way to compare the similarity between CWNs is using the sum of the weights of the outgoing edges (or \( \text{out} (E) \)). The set of \( \text{out} (E) \) could be used to depict the number of nodes the network has, the number of users it serves or the opportunity cost of its services or value offerings. These potential measures will be addressed thoroughly when we study the outcomes (e.g., accruing social capital, generating human capital, and improving the business environment, achieving digital inclusion). In other words, we will address the dual version of this phenomenon which is considering CWNs as super sources of a pool of added values which feed, or benefit, multiple sinks (e.g., local community, developing organizations, municipalities, technology vendors), or subgroups of stakeholders of CWNs.

We can also use graph concepts to identify the maximum size of the CWNs-by-contributions network. This is a classical problem in graph theory known of the maximal complete bipartite graph or the maximal bicliques (Li, Sim, Liu, and Wong 2008). In our case, this graph should include all types of contributions or inputs.

<table>
<thead>
<tr>
<th></th>
<th>Court</th>
<th>Seattle</th>
<th>Champ</th>
<th>AirStr</th>
<th>PTAW</th>
<th>Keur</th>
<th>Preto</th>
<th>Cstle</th>
<th>WUG</th>
<th>Cape</th>
<th>Nepal</th>
<th>Red L</th>
<th>Jawur</th>
<th>NYCw</th>
<th>ZGWi</th>
<th>el Pa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Court</td>
<td>1.00</td>
<td>0.94</td>
<td>-0.55</td>
<td>-0.14</td>
<td>0.81</td>
<td>0.30</td>
<td>0.88</td>
<td>0.88</td>
<td>0.87</td>
<td>0.89</td>
<td>0.90</td>
<td>0.88</td>
<td>0.42</td>
<td>0.12</td>
<td>0.16</td>
<td>0.89</td>
</tr>
<tr>
<td>Seattle</td>
<td>-0.55</td>
<td>-0.26</td>
<td>1.00</td>
<td>-0.24</td>
<td>0.96</td>
<td>0.06</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
<td>0.98</td>
<td>0.31</td>
<td>0.25</td>
<td>-0.11</td>
<td>0.08</td>
<td>0.99</td>
</tr>
<tr>
<td>Champ</td>
<td>-0.14</td>
<td>-0.24</td>
<td>-0.93</td>
<td>1.00</td>
<td>-0.25</td>
<td>0.29</td>
<td>-0.23</td>
<td>-0.24</td>
<td>-0.28</td>
<td>0.16</td>
<td>0.21</td>
<td>-0.27</td>
<td>-0.26</td>
<td>-0.26</td>
<td>0.20</td>
<td>-0.26</td>
</tr>
<tr>
<td>AirStr</td>
<td>0.91</td>
<td>0.96</td>
<td>0.79</td>
<td>-0.25</td>
<td>1.00</td>
<td>-0.19</td>
<td>0.99</td>
<td>0.95</td>
<td>0.99</td>
<td>0.97</td>
<td>0.98</td>
<td>0.29</td>
<td>0.02</td>
<td>-0.35</td>
<td>0.33</td>
<td>0.99</td>
</tr>
<tr>
<td>PTAW</td>
<td>0.30</td>
<td>0.06</td>
<td>-0.33</td>
<td>-0.23</td>
<td>0.19</td>
<td>0.13</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.13</td>
<td>-0.13</td>
<td>-0.11</td>
<td>-0.14</td>
<td>0.95</td>
<td>0.98</td>
<td>0.99</td>
<td>-0.11</td>
</tr>
<tr>
<td>Keur</td>
<td>0.28</td>
<td>0.98</td>
<td>0.00</td>
<td>-0.23</td>
<td>0.99</td>
<td>-0.13</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.30</td>
<td>-0.27</td>
<td>1.00</td>
</tr>
<tr>
<td>Preto</td>
<td>0.98</td>
<td>0.98</td>
<td>-0.04</td>
<td>0.24</td>
<td>0.99</td>
<td>-0.14</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-0.28</td>
<td>1.00</td>
</tr>
<tr>
<td>Cstle</td>
<td>0.97</td>
<td>0.98</td>
<td>0.80</td>
<td>-0.23</td>
<td>0.99</td>
<td>-0.13</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.98</td>
<td>0.95</td>
<td>1.00</td>
<td>0.01</td>
<td>-0.29</td>
<td>-0.26</td>
<td>1.00</td>
</tr>
<tr>
<td>WUG</td>
<td>0.89</td>
<td>0.97</td>
<td>-0.92</td>
<td>-0.16</td>
<td>0.97</td>
<td>-0.13</td>
<td>0.99</td>
<td>0.99</td>
<td>0.98</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.30</td>
<td>-0.27</td>
<td>0.99</td>
</tr>
<tr>
<td>Cape</td>
<td>0.99</td>
<td>0.98</td>
<td>-0.60</td>
<td>-0.21</td>
<td>0.99</td>
<td>-0.11</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.03</td>
<td>0.27</td>
<td>0.25</td>
</tr>
<tr>
<td>Nepal</td>
<td>0.88</td>
<td>0.31</td>
<td>-0.35</td>
<td>-0.27</td>
<td>0.29</td>
<td>-0.14</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.22</td>
<td>-0.31</td>
<td>0.28</td>
</tr>
<tr>
<td>Red L</td>
<td>0.42</td>
<td>0.25</td>
<td>-0.33</td>
<td>-0.25</td>
<td>0.99</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.03</td>
<td>-0.22</td>
<td>1.00</td>
<td>0.95</td>
<td>0.96</td>
<td>0.03</td>
</tr>
<tr>
<td>Jawur</td>
<td>0.12</td>
<td>-0.11</td>
<td>-0.33</td>
<td>-0.30</td>
<td>-0.35</td>
<td>0.98</td>
<td>-0.30</td>
<td>-0.30</td>
<td>-0.29</td>
<td>-0.30</td>
<td>-0.27</td>
<td>-0.31</td>
<td>0.95</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.27</td>
</tr>
<tr>
<td>NYCw</td>
<td>0.16</td>
<td>0.00</td>
<td>-0.31</td>
<td>-0.23</td>
<td>-0.32</td>
<td>0.99</td>
<td>-0.27</td>
<td>-0.28</td>
<td>-0.28</td>
<td>-0.27</td>
<td>-0.25</td>
<td>-0.28</td>
<td>0.96</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.25</td>
</tr>
<tr>
<td>ZGWi</td>
<td>0.89</td>
<td>0.95</td>
<td>-0.32</td>
<td>-0.25</td>
<td>0.99</td>
<td>-0.31</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.03</td>
<td>-0.27</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Table 2. The similarity between
Figure 2 shows that the maximum complete bipartite graph, for the investigated networks, is of size \( k_{2,6} \). It includes two networks: Airstream and Red Libre De Ometepe. Both have affiliations with all the investigated set of contributors. The question arises here is: how can we increase the size of this graph? In other words, how can we enable other networks to affiliate with all possible segments of contributors and engage the community at large?

5.2. Classifying CWNs

The purpose of this experiment is to partition, or group, CWNs-by-contributions into classes based on the density of community contributions. Specifically, we identified CWNs which have high-density of contributions and presented them as close as possible to each other in one block called “the core.” We also obtained another set of CWNs which have very low-density of contributions in one group, called “periphery.” This approach is called “core-periphery” analysis (Hanneman and Riddle 2005). Others call it “Blockmodel,” through which the cells of the data matrix are sorted such that rows and columns that belong to the same class are organized close to each other (Borgatti 2008). We applied this type of analysis to the collected data to obtain an ideal image of high-density and low-density groups along the main diagonal. This approach uses the concept of graph density to classify CWNs-by-contributions. In particular, we partitioned the columns (types of contributions) and the rows (CWNs) into four groups based on the density of contribution for each group. In a directed graph with weighted edges, the term graph density refers to the average contribution of each group or partition. Table 3 shows that the bottom right block has the highest density, 26400.5 dollars, of contributions and the upper right block has the lowest density of contribution which is 283.6 dollars.

![Figure 3. a complete bipartite graph of size \( k_{2,6} \)](image)

Table 3. Partitioning CWNs into groups

<table>
<thead>
<tr>
<th>Starting fitness: 0.164</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final fitness: 0.179</td>
</tr>
<tr>
<td>Correlation to ideal: 0.179</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Blocked Adjacency Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>13</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>15</td>
</tr>
</tbody>
</table>

Density matrix

```
1   2
1 20143809 203654
2 1510263 26400626
```
A deep look at these blocks shows that the networks in the upper half (e.g. WUG, Keur Sedaro, etc.) have more shared access points, money donations and OSS. They also obtain less volunteerism of time, hardware donations, and free technical support from their communities. On the contrary, the networks in the lower half (e.g. SeattleWireless, Champaign-Urbana, etc.) have less sharing access points, money donations and OSS. They also receive more volunteerism of time, hardware donations, and technical support. The question is: why these two groups are different in terms of types and amounts of contributions they receive. We hypothesize that these different patterns between these groups could be due to the differences (e.g. social values, average income, education, etc.) between the contributors across these networks. We can also compare the density of each group to the density of the whole graph which is 12719.6 dollars. This density could be used to measure the level of community engagement in the investigated networks. Yet, we would like to emphasize that these facets of contributions are dynamic in terms of the number of contributors and the type of their inputs. This is because of the loose structure and nonprofit status of these networks, similar to other collectivities.

6. SIGNIFICANCE OF RESEARCH

Through the Internet, users share data, exchange massages and collaborate on research, and develop OSS (Lesser, Fontaine and Slusher 2000). In this study, we show how they build and share telecommunication infrastructures. In particular, we use graph theoretic approach to propose fresh perspective for understanding the role of communities in the development of their wireless infrastructures. In particular, we modeled, measured and analyzed the contributions of community members in a quantitative and compact manner. To measure community contributions in these collective projects, we construct the variables of OSS, money donations, technical expertise, voluntary work, node-sharing, and hardware donations. Identifying, mapping, measuring and evaluating these resources are important for a successful CWN endeavor. We also distinguish contributions of technical processes (e.g. developing OSS, providing technical support) from unskilled manpower. This would be useful if we study the role of CWNs in building technical expertise or human capital. Furthermore, we convert them into marketable values, using their opportunity cost, to help us aggregate them and measure the market value of these collective projects.

This study provides policy makers with important insights into CWNs as a potential solution for the digital inequality problem. The used analytical approach could be used to assess the socioeconomic benefits of such collective projects. Measuring such benefits may help researchers to obtain new insights into how to engage more participants. In particular, it may help community leaders to convince governments, OSS groups, students and developing agencies to contribute to this movement. The participation of government entities, for instance, may take different forms such as providing funds, deregulating the necessary spectrum, and/or providing the legal rights of communities and nonprofit organizations to collectively build, own, and run such telecommunication infrastructures. We also hope that the proposed model would expand the problem-solving abilities of practitioners when they use it to visualize, map, measure and manage the resources and outcomes of CWNs.

For academicians, our study provides an elegant analytical approach and a rich set of conceptual insights to guide current and future research on CWNs and similar collectivities such OSS groups, MySpace, Wikipedia, Facebook and other online forums and collaboration groups. In particular, our approach in treating CWNs as value networks is necessary to capture the complexity of relationships and transactions between their actors. Proposing such an approach, at this embryonic stage, is important in order to advance this intellectual stream. Another scientific contribution of this paper is providing an example of how to treat 2-mode network data and how to assess intangibles using their opportunity costs.

We believe that this perspective of treating the resources and assets of CWNs is illuminating but incomplete. Its main limitation is excluding important facets of community contributions such as boosting the publicity and awareness of the project. Another limitation is assuming that both CWNs and contributors are disjoint sets. In other words, we assumed that there are no relationships between individual CWNs and that groups of contributors do not intermix; however, this is not the case in reality. For example, volunteers may donate money, provide technical support, and/or develop software for the system. We isolate these inputs in order to distinguish financial contributions from nonfinancial contributions and cognitive contributions (e.g., technological processes) from
non-cognitive contributions (e.g. manpower). In addition, CWNs collaborate with each other in OSS development (e.g. their annual summit). We assume that these two sets are disjoint only for analytical purposes. Another important observation is that the agents of CWNs are spatially distributed and their cooperation may occur virtually or over time. Furthermore, the used data, Table 1, is correct only at the time of data collection, which is June of 2008. These networks are dynamic and their lack of commitments of participants. However, the adopted analytical approach should be useful in helping researchers resolve a number of critical issues related to CWNs. There is also methodological issue as the used SNA approach is limited only to small networks where we have confidence in the reliability of our observations, according to Hanneman and Riddle (2005).

It would be interesting if future work relates contributions of actors to their attributes such as their size, social values, income, age, gender, diversity, social cohesion and education. This may help answer questions such as why there is so much voluntary participation or money donation in specific networks compared to others. Another potential topic is focusing on one or few CWNs and study the one-mode, or micro, relationships (e.g. benefits, contributions, reciprocation, influence, solidarity, trust) between their actors. For instance, focusing on one network would help researchers to visualize and classify the contributors, beneficiaries, and isolates in a specific community. Our future work will focus on the dual version of the relationship between CWNs and their actors. In particular, we will investigate how CWNs, as macro-structures, create values for participants (e.g., volunteers, OSS developers, underserved communities) as micro-structures. We will use the adopted analytical approach to treat CWNs as sources of values. In particular, we will provide a careful analysis of the role of these networks in creating social capital and human capital.

7. CONCLUSIONS

This study presents a new direction in understanding the role of communities in developing their CWNs. In particular, In particular, it provides new insights into the role of communities, as macrostructures, in developing CWNs, as microstructures. We used graph concepts to model community contributions in the development of their wireless infrastructures. The proposed model is used to visualize, classify, and compare CWNs. The originality of our work lies in providing a generic approach to help practitioners and academicians to treat CWNs as value networks and systems of transactions. We considered both the tangible and intangible resources shared, mobilized or reproduced to build these networks. This study is an important step towards advancing this area as a research stream. It fills a gap between theory and practice for community activists, developing agencies, and researchers as it describes CWNs as systems of exchange. The phenomenal growth of this new form of collective actions is expected to completely revolutionize the telecommunication landscape and provide momentum to the digital paradigm. However, they need the attention of governments, developing and outreach organizations, technology vendors, students and researchers.

References


PRESENTING DATA FOR TEAM-BASED DECISION-MAKING IN AGILE INFORMATION SYSTEMS PROJECTS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0540.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Agile computing, Decision making, Decision support systems, Research in progress</td>
</tr>
</tbody>
</table>
PRESENTING DATA FOR TEAM-BASED DECISION-MAKING IN AGILE INFORMATION SYSTEMS PROJECTS

Conboy, Kieran, National University of Ireland, Galway, University Road, Galway, Ireland, kieran.conboy@nuigalway.ie

Acton, Thomas, National University of Ireland, Galway, University Road, Galway, Ireland, thomas.acton@nuigalway.ie

Halonen, Raija, National University of Ireland, Galway, University Road, Galway, Ireland, raija.halonen@nuigalway.ie

Abstract

Although there are a number of purported benefits of agile systems development methods over traditional forms, many in the Information Systems Development (ISD) community highlight the difficulties that the dynamic, turbulent nature of agile development environments may present. One prominent difference compared to traditional ISD projects is the emphasised presence of the customer in decision-making. Specifically, project management and decision-making can be significantly more challenging than in projects involving traditional development approaches. This paper describes a research-in-progress, that focuses on decision-making in agile projects, and aims to develop a better understanding of how agile teams present dynamic data to decision makers who may act “on the fly”, and in so doing lead to effective, high quality decision-making in agile environments. The goal of the paper with its three case studies is to develop a set of best practices for representing decision-making data in agile ISD projects.

Keywords: Agile IS project, Decision-making, Data presentation.
1 INTRODUCTION

This research captures one of the state-of-the-art development methods of today. The last 10 years or so has seen the emergence of a number of information systems development (ISD) methods, which have collectively been labelled as agile. Some of the most popular include eXtreme Programming (XP) (Beck 1999), the Dynamic Systems Development Method (DSDM) (Stapleton 1997), Scrum (Schwaber & Beedle 2002), Crystal (Cockburn 2001), Agile modelling (Ambler 2002), Feature Driven Design (FDD) (Coad et al. 1999), and Lean Software Development (LSD) (Poppendieck 2001), along with variants of each e.g. XP-Lite (Aveling 2004). These methods have been well received by those in ISD and there is strong anecdotal evidence to suggest that awareness and indeed use of these methods is highly prevalent across the community. However, use of these methods has a significant impact on the project’s decision-making processes.

Therefore, this research-in-progress paper aims to investigate and develop appropriate methods of data presentation to support high quality decision-making in agile information system projects with the means of qualitative multiple-case design (Benbasat et al. 1987).

In an agile development environment, the project manager’s role as a decision-maker is greatly reduced, and is more akin to that of a facilitator or coordinator (Nerur et al. 2005; Alleman 2002; Lindstrom & Jeffries 2004). Further, the development team makes most of the decisions, creating a “pluralist decision-making environment” (Nerur et al. 2005) due to the diverse backgrounds, attitudes, goals, and cognitive dispositions of the team members (Highsmith 2004; Chin 2004; Cockburn & Highsmith 2001). In agile projects, the organisation or team structure is “organic and flexible”, as opposed to traditional structures which are “mechanistic, bureaucratic and formalized” (Nerur et al. 2005); the project is completed through a series of iterations, each often as short as a few working days (Fowler & Highsmith 2001; Fitzgerald et al. 2006), resulting in more frequent, short-term decision-making; and software is valued over documentation (Fowler & Highsmith 2001) which was traditionally used as a vital decision aid. Significantly, the customer plays a more continuous and embedded role, and thus is intrinsically involved in most decisions (Griffin 2001; Farell et al. 2002; Beck & Andres 2004; Beck 2000). Moreover, developers are not confined to a specific specialised role and are encouraged to self-organise, interchanging and blending roles (Nerur et al. 2005) and involved in decisions that may fall outside their traditional skill areas.

One could assume that decision-making in an agile environment should be ad hoc, unstructured, and without discipline. Indeed, there is a common misconception that agile methods are centred around improvisation, and care-free deviation from rules and regulations. However, Beck dismisses this stating that agile “is not an excuse for unilateral behaviour” and he views agility versus discipline as a “false dichotomy” (Beck & Boehm 2003). In fact he argues that agility “is only possible through greater discipline on the part of everyone involved”. The need for discipline has been stressed by many key texts across a broad range of the agile methods (e.g. Schwaber 1996; Cockburn 2001; Schwaber and Beedle 2002; Beck and Andres 2004).

Next, we present the research objectives of our study. After that, the theoretical background for decision-making in agile IS projects is presented. Then we describe the research approach and the empirical cases of our study.

2 RESEARCH OBJECTIVES

As is often the case with new and emerging phenomena in ISD, agile method practice has led research, with the creation, promotion and dissemination of these methods almost completely due to the efforts of practitioners and consultants. Agile method research has gained momentum in more recent years,
but has been argued that the current body of agile method knowledge suffers from a number of conceptual problems such as a lack of clarity, parsimony, cumulative tradition, and theoretical ‘glue’.

One key weakness of agile method research is a lack of focus on how best to present data for decision-making in such an environment, particularly when many contribute to and are involved in the decision process. Indeed decision-making in this context is much different to that associated with traditional systems development, with many additional complexities, uncertainties and hurdles to overcome. There is a need to capture and represent metrics on the progress of the project, for example, estimates and bounds for stage delivery, dependencies and required resources for iterative development, timeline information on next stages, personnel information such as planned vacations, and external factors such as customer availability for development across iterations. Typically these are dynamic data, and are best represented either in tabulated formats with columns and column attributes, or in other aggregate representations such as appropriate graphs and charts. Indeed the latter has been shown to be particularly valuable in decision-making instances and scenarios involving key performance indicators. It follows that the format of presented data can be influential in decision-making scenarios. In agile development the customer plays a more continuous and embedded role, and thus is intrinsically involved in most decisions (Griffin 2001; Farell et al. 2002; Beck & Andres 2004; Beck 2000). This in contrast to more traditional approaches where customers do not get involved in day-to-day operational development; rather their involvement is limited to intermittent events such as prototyping sessions and release meetings.

It follows that data presented to decision makers in agile scenarios needs to be accessible to many people with varying roles and foci: considering the frequency of decisions in these environments together with the team-based approaches to decision-making endemic in agile projects and a need for timely decisions, it would seem sensible that data are presented to all actors in a manner that enables team interactions and facilitates the decision making process. One such manifestation of this may involve a single presentation technique to all involved, perhaps displayed on a computer screen or projection, that is easy to understand, displays interactions between variables, and allows data manipulation during the decision-making process. Such a presentation mechanism would not be solely output-based insofar as offering only data display, but would be pliable in terms of enabling decision-maker interaction with the data. In so doing the presentation mechanism could support the decision-making process.

In this study we seek to investigate decision-making efficacy in agile projects by exploring how best to present dynamic data to decision makers in team-based decision scenarios. The primary focus is in understanding the decision-making strategies in agile project management, supporting high quality decision-making through data presentation. The study will examine the decision strategies involved by project managers and teams in ensuring the success of agile projects in a context with both customer and IS teams consisting of diverse professions as the decision-making can not be studied outside its natural setting (see Benbasat et al. 1987). Furthermore, the study will investigate the appropriateness and value of various data presentation techniques to best support quality decision-making. In so doing, as output the study aims to develop a data presentation model for quality decision-making in agile projects involving the composition, structuration and presentation of relevant data.

The framework (Fig. 1) for the study is as follows:
Commensurate with the IS literature in this regard, we assume that decision makers will employ an effort-accuracy tradeoff in the decision-making process, and seek to reach an optimal decision with least effort. We also assume that the format of the data used in the decision-making process can support particular decision-making strategies: it remains to be seen whether such presentation formats can encourage quality decision-making by promoting superior strategies, but in so doing reduce the effort required for their use, and therefore provide a mechanism whereby decision makers are likely to employ such strategies.

Through an examination of the impact of data presentation on decision strategies used in the decision-making process, the study will provide a best-fit model of data presentation formats best suited to quality decision making in agile project management. Our research objective is thus: To investigate and develop appropriate and effective methods of data presentation to support high quality decision-making in agile information systems project management.

To realise the objective we will explore the decision-making strategies in agile project management, and investigate the effects of various data presentation formats on these strategies. Considering that there is substantial evidence in the literature that decision processes and employed strategies can directly impact decision quality, we strive to investigate decision making in agile development projects from a process perspective, with the aim of that identifying optimal decision strategies and the presentation formats underpinning them, so that decision quality can be maximised.

3 THEORETICAL BACKGROUND

Decision making is the performance of a task, that of making a particular decision (Bahl & Hunt 1984). Project managers and agile teams are faced with decision tasks on a daily and more frequent basis, with the occurrence of such tasks increased in agile development projects as a consequence of rapidly changing requirements, expectations, and their underlying data. Typically these tasks involve choice, with the decision makers presented with large amounts of data, and categories of data describable by various attributes that may influence the value of those data for decision-making. Decision makers in dynamic contexts need to assess various courses of action, various potentially good avenues of resource, or investigate the relativity of one potential decision over others, and make a choice. Such activities are multifaceted, and the process of quality decision-making includes aspects such as decision maker behaviours and strategies, the inherent timeliness of decisions particularly impactful in agile development scenarios, as well as data-related aspects such as data representation.
A decision strategy can be considered as a method (sequence) of operations for searching through a problem space (Payne et al. 1988; Todd & Benbasat 1999). Some strategies are better than others, insofar as leading to higher quality decisions. Benbasat and Todd (1996) argue that decision makers have “a series of pre-existing strategies which have been learned and stored in memory”, and to formulate a strategy, they make a selection “from a repertoire of pre-existing strategies which may be applied or adapted to the problem at hand.” They argue that the determination and application of a particular decision strategy is contingent upon the capabilities of decision makers as well as the characteristics of the problem (that is, the decision-making task) and the availability of decision aids as tools to support particular strategies. However, in choice-based decision-making scenarios, with decision makers choosing from alternatives, the use of particular decision strategies is dynamic, and decision makers are highly adaptive in strategy selection (Payne 1976; Kuo et al. 2004; Häubl & Trifts 2000). Decision making is the accomplishment of a decision task, where decision influences are assessed and where a decision choice is made. The application of strategies and decision-making approaches can be task-centred, with decision-making behaviours influenced by the nature of the task as well as other factors.

The choice and use of a particular decision strategy, or combination of strategies, is dependent upon many factors. The cognitive ability of the decision maker in conjunction with behavioural characteristics and decision styles can influence strategy selection and can impact decision making. Further, the nature of the decision-making task can lend itself to the employment of particular strategies: the use of accurate but more effortful strategies can arise in tasks where there is an importance attached to the decision outcome, and where decision-making time may be secondary to decision outcome. Also, data presentation formats may influence strategy selection, with particular display formats commensurate more so with higher quality strategies.

The cognitive ability of decision makers in conjunction with behavioural characteristics and decision styles can influence strategy selection and can impact decision making. Indeed many other person-centric and behavioural factors can be influential in the decision-making process, including self-efficacy, decision and management styles, team skills and others: a more thorough discussion is presented by Moldafsky and Kwon (1994). Decision makers perceive only that information relevant in order to achieve a specific goal within the problem space under consideration (Yadav & Khazanchi 1992). Zmud (1979) argues that whilst information requirements of decision makers are task-dependent as well as influenced by situational and individual characteristics, decision makers too often do not understand their information requirements, and demand too little or too much information. Further, he argues that decision makers are typically unable to make full use of the information provided or available in reaching a decision. Tasks best requiring the implementation of particular decision strategies for their resolution may only be optimally addressed if those strategies are recognisable as appropriate or useful by the decision maker. If decision makers possess the skills, intellect, and confidence to match particular behavioural strategies to particular task scenarios, and have the cognitive abilities required to choose and implement such strategies, then more optimal strategies may be implemented (Kuo et al. 2004; Eierman et al. 1995). Further, the nature of the decision-making task can lend itself to the employment of particular strategies: the use of accurate but more effortful strategies can arise in tasks where there is an importance attached to the decision outcome, and where decision-making time may be secondary to decision outcome. Also, data presentation formats may influence strategy selection, with particular display formats commensurate more so with higher quality strategies.

Studies have shown that modes of presentation of information to decision makers can be extremely influential on task performance (Garrity et al. 2005; Benbasat & Dexter 1986; 1985; Gershon et al. 1998; Jarvenpaa 1989; Hu et al. 1999; Kester et al. 2001; Lamberti & Wallace 1990; Marsden et al. 2002; Montazemi et al. 1996; Noyes & Garland 2003; Tan & Benbasat 1990). Vessey and Galletta (1991) found that individuals’ mental representations of the task space and the processes required to complete the (problem-solving) task were influential in determining the strategies employed in task completion. Literature from behavioural decision theory as well as consumer behaviour support
problem solvers’ use of processes and strategies that match the representation of the task (McClintock & Becker 1967; Barron 1974; Edwards 1961; Eierman et al. 1995; Gray & Wert-Gray 1999; Häubl & Trifts 2000; Payne 1982; 1976; Payne et al. 1988; Slovic et al. 1977). Further, there is evidence that individuals employ different decision-making and problem-solving processes and strategies for different types of task. Vessey and Galletta (1991) argue that matching the problem representation directly to the task positively impacts task performance, and that individuals’ skills have the greatest impact on performance when there is a close match between task representation and the task itself. Indeed they found that the former was largely unimportant in isolation and that the latter was substantially more important: where task representation alone was matched with the individual skills necessary for task resolution, there were no performance effects. Where individual skills supported the task itself, or both the task and its representation, there was significantly positive performance effects. As such, skill affected task solution rather than the acquisition of information alone. However, in this study the overall performance impact of individual skill-based differences were minimal compared to the effect of a fit between the task representation and the task problem. This study (involving the comparison of alternatives across a number of attributes) found that individuals prefer tabulated information to other representative forms.

It is clear that modes of presentation of information to decision makers can be extremely influential on decision making. Vessey and Galletta’s (1991) study involving the comparison of alternatives across a number of attributes found that individuals prefer tabulated information to other representative forms. Indeed Speier and Morris (2003) found that informational display format was influential in affecting decision performance. Further, Jarvenpaa (1989) found that display formats used in a multiattribute choice task were central in influencing decision strategies when the format was matched to the task, and that strategies which minimised the cognitive effort were preferred.

In this study we aim to examine the most appropriate data presentation formats to best support quality decision-making. In examining the usage, employment and selection of particular decision-making strategies, we aim to support decision making in agile projects by mapping presentation format to decision strategies, so as to promote the use of high quality strategies in the decision-making process, and thus positively impact decision quality. The study will examine a number of cases exploring how data are presented; investigate the decision strategies in use; map presentation format to decision process and strategies used; and produce a model for data presentation formats most likely to support and lead to high quality strategies in agile situations, and therefore superior decision-making.

4 RESEARCH METHODOLOGY & CURRENT PROJECT STATUS

To reach understanding of how teams act in agile IS projects; we apply experiential research methodology i.e. an integrated academic-practitioner team approach (Grant et al. 2001). With the methodology we aim to understand the "process behind the actual decisions being taken in the context of the inherent characteristics of entrepreneurs" (p. 66) in the teams of diverse professionals. The study will involve three case studies. Case studies are considered to be a suitable research approach for this study since it is exploratory in nature, with the intention of investigating decision-making in a real-life context (Stake 2000; Yin 2003) and they explore a phenomenon in its natural setting, applies several methods of data collection to gather information from one or a few entities (Benbasat et al. 1987). Empirical data will be collected over a 6 month period from December 2008 to May 2009.
<table>
<thead>
<tr>
<th></th>
<th>Project A</th>
<th>Project B</th>
<th>Project C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team size</td>
<td>8 Ireland, 5 U.S.</td>
<td>24 people (sub-group of 5 using agile) Ireland</td>
<td>7 Ireland, 6 U.S.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ireland</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 U.S.</td>
<td></td>
</tr>
<tr>
<td>Team composition</td>
<td>1 SVP, 1 VP, 1 director, 1 principal engineer, 1 tester</td>
<td>1 SVP, 1 VP, 1 director, 2 principal engineers</td>
<td>1 VP, 1 project manager, 1 principal engineer, 1 tester, 3 engineers</td>
</tr>
<tr>
<td></td>
<td>1 directors, 1 project manager, 1 principal engineer, 5 engineers</td>
<td>2 directors, 2 project managers, 2 architects, 15 principal engineers, 4 testers</td>
<td>1 project manager, 1 principal engineer, 1 tester, 3 engineers</td>
</tr>
<tr>
<td>Project Duration</td>
<td>1.5 years</td>
<td>7 years (2 years agile sub-group)</td>
<td>1 year</td>
</tr>
<tr>
<td>Type of system</td>
<td>Exploring new strategic opportunities of cutting edge Web 2.0 technologies</td>
<td>Security system for provision of enterprise wide access control</td>
<td>Suite of tools to monitor SLA adherence across Pennysoft’s core application service providers (response times etc) Application for monitoring/tracking system issue resolution</td>
</tr>
<tr>
<td>developed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customer</td>
<td>Internal</td>
<td>Internal</td>
<td>Internal</td>
</tr>
<tr>
<td>End Users</td>
<td>General population</td>
<td>Pennysoft staff</td>
<td>Pennysoft staff</td>
</tr>
</tbody>
</table>

Table 1. The Profiles of the Three Cases.

In this research, we will observe the meetings and collect literary material written by the project members. Further, we will collect emails and other notes that are written related to the project. To analyse the collected material, we will use discourse analysis (Potter, 1999) that supports naturally occurring talk and discussions, keeping in mind its context.

In addition, data collection will involve personal face-to-face interviews, a technique well suited to case study data collection, and particularly for exploratory research such as this because it allows expansive discussions which illuminate factors of importance (Yin 2003; Oppenheim 1992). The information gathered is likely to be more accurate than information collected by other methods since the interviewer can avoid inaccurate or incomplete answers by explaining the questions to the interviewee (Oppenheim 1992). The questions will be largely open-ended, allowing respondents freedom to convey their experiences and views, and expression of the socially complex contexts that underpin ISD (Yin 2003; Oppenheim 1992). The interviews will be conducted in a responsive (Rubin & Rubin 2005; Wengraf 2001), or reflexive (Trauth & O'Connor 1991) manner, allowing the researcher to follow up on insights uncovered mid-interview, and adjust the content and schedule of the interview accordingly. In order to aid analysis of the data after the interviews, all will be recorded with each interviewee’s consent, and subsequently transcribed, proof-read and annotated by the researcher. In any cases of ambiguity, clarification will be sought from the corresponding interviewee, either via telephone or e-mail. In relation to the interviews, supplementary documentation will also be collected, including project management plans, budgets and budget reports, meeting minutes and relevant e-mail communications.
Alternatively, we will carry out an interview done by questionnaires that will include both closed, semi-structured and open questions. With this choice we will ensure that we will reach also those respondents whose voice might otherwise remain unheard (Stapleton 2008).

Data analysis will use Strauss & Corbin’s (1998) open coding and axial coding techniques. Open coding is “the process of breaking down, examining, comparing, conceptualizing, and categorizing data” (Strauss & Corbin 1998). Glaser (1992) argues that codes and categories should emerge from the data, while with Strauss & Corbin’s approach (1998) these are selected prior to analysis. The approach adopted in this study is more akin to the latter, where the interview questions and subsequent analysis will be based on Payne’s (Payne et al. 1988) decision-making model. This will provide a list of “intellectual bins” or “seed categories” (Miles & Huberman 1999) to structure the data collection and the open coding stage of data analysis.

The second phase of analysis will use axial coding. Axial coding is defined by Strauss and Corbin (1998) as a set of procedures whereby data are put back together in new ways after open coding; whereas open coding fractures the data into categories, axial coding puts the data back together by making connections between the categories and sub-categories. As the data is coded, theoretical questions, hypotheses and code summaries will arise. These will be documented in analytic memos (Miles & Huberman 1999) to aid understanding of the concepts being studied and to refine further data collection. Miles and Huberman (1999 p. 72-74) offer advice on effective analytic memos, and these practices will be followed where possible.

As categories emerge, follow-up interviews will be arranged to elicit further, richer, more focused information. This will be done to confirm, extend, and sharpen the evolving list of categories. As categories become integrated, further data collection will not tend to cause any additional categories to emerge, but rather reinforce those already in existence. At this point, the categories will be deemed to be “theoretically saturated” (Strauss & Corbin 1998), and data collection ended.

Once the decision-making process is documented and analysed, the focus will shift to constructing decision-making scenarios involving underlying data presented in various formats to decision-makers. Revisiting the cases using a think aloud protocol, the study will investigate the impact of data presentation format on the decision-making process, in terms of the decision strategies used to reach decision finality. Using decision strategies as indicators of decision-making quality from a process perspective, the study will construct a model illustrating the relationship between data presentation format and decision-making processes in agile project management.

In all, the case study database will fulfil recommendations by Yin (2003) with its components of case study notes, case study documents, tabular material and a case study narrative.

The Cases

The primary operation of Pennysoft, a large, privately owned U.S. company, (employing 40000) involves the provision of financial services and investment resources. The company has been developing software at its site in Ireland since 1995, and currently employs around 300 people at this Irish site. The software products developed are supplied mainly to internal customers in the U.S. Many projects involve co-ordinating with several teams in the U.S. and India. In many cases, the requirements are generated in the U.S. with software development then taking place in both the U.S. and Ireland. In this study we analyse three systems development project cases within Pennysoft (Table 1).

Table 1 describes the three cases highlighting their differences. By including three cases in our study we will get a good conception of how decision-making is carried out in agile IS projects. As the chosen project teams involve several professions, we will get a conception of preferred data presentation as perceived by the team members. In addition, the team compositions reveal that there
are high-ranked people involved, such as a senior vice president (SVP) and vice presidents (VP) and this situation may influence the decision-making (see Stapleton 2008).

While we aim to examine the most appropriate data presentation formats to best support decision-making, we believe that the three cases offer good insight into real-time decision settings and environments.

References


ENABLING OPEN INNOVATION: PROPOSAL OF A FRAMEWORK SUPPORTING ICT AND KMS IMPLEMENTATION IN WEB-BASED INTERMEDIARIES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0569.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Open Innovation, Sensemaking, Absorptive Capacity, Design Science</td>
</tr>
</tbody>
</table>
Abstract

Open Innovation is a model used to describe how nowadays companies source and exploit new technologies, new products and services. Web-Based Intermediaries (WBIs) have entered the emerging innovation market and are expected to dramatically increase the number of innovation exchanges. However there are not yet clear theoretical guidelines supporting the design and management of such intermediaries. We use organizational sense-making theory and relative absorptive capacity (RAC) theory to analyze the factors that still hinder Open Innovation. From sense-making theory and RAC theory we draw directions on the services WBIs need to provide in order to effectively support an innovation market. Since information technology is critical to the success of a WBI, we also give directions on how ICT and KMS can be used in order to support these services.

Keywords: Open Innovation, ICT, Intermediaries, Sense-making, Relative Absorptive Capacity.
1 INTRODUCTION

The concept of Open Innovation (OI) was first proposed by Henry Chesbrough (Chesbrough 2003) and has quickly gained the attention of scholars and practitioners. While in the so-called closed innovation paradigm, research, development, and commercial exploitation of a new technology were performed mainly by large companies within their boundaries, today companies increasingly rely on outside innovation for new products and processes and have become more active in licensing and selling results of their own innovation to third parties (OECD 2008).

OI is strongly driven by globalization. The OECD has recently addressed the issue in a research project on “globalization and OI” (OECD 2008). The findings of this project show that increasingly companies link into global innovation networks with people, institutions, and other companies in different countries to source and/or exploit innovation. According to some authors a global, secondary market for technology is emerging (e.g. Chesbrough 2003, 2006, Lichtenhalter & Ernst 2008a, 2008b, OECD 2008).

Information Systems play a critical role in coordinating innovation markets. Large companies like Procter & Gamble and IBM have made major investments in proprietary platforms supporting the sourcing (Dodgson et al. 2006) or exploitation (Davis & Harrison 2001) of innovation.

In particular, Web-Based Intermediaries (WBIs) for OI provide virtual milieus to bring together buyers and sellers and to support their transactions. Yet2.com, for example, is a technology marketplace where it is possible to exchange patented inventions. Other WBIs, like Ninesigma, Innocentive or Yourencore provide access to broad networks of scientists, researchers and professionals which are potentially able to solve new technological problems proposed by companies (Tapscott & Williams 2007; Chesbrough 2006). These intermediaries are expected to dramatically increase the number of innovation exchanges (Fredberg et al. 2008; Lichtenhalter & Ernst 2008a, 2008b, OECD 2008). In fact, they enable firms to source (as well as exploit) innovation globally with limited investments in proprietary structures. However reducing transaction costs and information asymmetries through a broad availability of information is not sufficient to improve liquidity in the innovation market. Other factors hinder the development of an efficient market. Some of these factors are related to regulatory issues or cultural factors as the NIH syndrome (Laursen & Salter 2006, Lichtenhalter & Ernst 2008b). In this paper we focus on knowledge related factors. That is, we explore what factors hinder the ability of firms to recognize a valuable innovation among the many offered on the market, to find an external application suitable for a technology firms developed or to transfer the innovation from the provider to the recipient.

We argue that the difference between the context in which an innovation is developed and the context in which it is applied is a major cause of the difficulties that hinder OI. We use organizational sense-making theory (Weick, 1995, Weick & Sutcliffe 2001, Taylor & Van Every 2000, Weick et al. 2005) and relative absorptive capacity (RAC) theory (Lane & Lubatkin, 1998; Lichtenhalter, 2008b) to describe and explain the problems that arise when firms which share no substantial previous relation take part in the innovation market. From sense-making theory and RAC theory we draw directions on the services WBIs need to provide in order to effectively support an innovation market. Since information technology is critical to the success of a WBI, we also give directions on how ICT and KMS can be used in order to support these services.

This paper provides a theoretical framework for analyzing the role of WBIs and ICT in a OI context. It is the first paper using sense-making theory and RAC theory together to study OI.

The reminder of this paper is organized as follows: in section two we briefly discuss functions that intermediaries, and WBIs in particular, perform in facilitating OI processes according to the existing literature; in section three we present our theoretical framework; in section four we apply our framework to give directions on the implementation of ICT and KMS by WBIs to support OI processes. Directions for further research and conclusions follow.
2 THE FUNCTIONS PERFORMED BY WEB-BASED INTERMEDIARIES IN THE INNOVATION MARKET

OI comprises all the processes involving the external sourcing or the external exploitation of innovation. Existing research points out that firms are interested in the potentialities of OI and are practicing or are experimenting OI-oriented activities (OECD, 2004; Sheehan et al., 2004; Lichtenthaler & Ernst, 2008b). However, apart from some well-known success cases, the application of the OI principles is still limited (OECD, 2004; Lichtenthaler & Ernst 2008b).

In the literature it is often remarked that when the participating firms share a common background the innovation exchange is easier (e.g. Hertzfeld et al. 2006, OECD, 2008, Stock & Tatikonda 2008). On the other hand, studies on Absorptive Capacity (Lane & Lubatkin, 1998, Lichtenthaler 2008b) underline that dissimilarities between firms hinder innovation exchange. OI involves by definition the exchange of innovations across different contexts and backgrounds. OI strategies are often driven by the intention to get access to valuable sources of knowledge in other countries bringing together companies and institutions with different cultures (OECD, 2008). The increasing integration of different technologies in many industries increases the need for interdisciplinary research; spin-offs and licensing often lead to the commercialization of innovation in different markets (Fredberg et al. 2008, OECD 2008); innovation is exploited in companies with different business models (Chesbrough 2003, 2006). These differences, then, are expected to be a major challenge to the development of OI.

The emergence of intermediary markets for ideas and technologies may facilitate the exchange of innovation. The number of firms interested in buying or selling innovation on an open market seems to be increasing (Arora et al. 2001; Muthusamy & White, 2005; Chesbrough, 2007; Lichtenthaler, 2008a; Lichtenthaler & Ernst, 2008b). Intermediaries entered the market which are able to bring together solution seekers and problem solvers or buyers and sellers of intellectual property (Lichtenthaler & Ernst, 2008a; Lichtenthaler & Ernst, 2008b; Fredberg et al. 2008; OECD, 2008).

Firms may draw on the resources and capabilities of intermediaries to improve their proficiency in exchanging innovation (Makadok 2001; Foss & Ishikawa 2007; Lichtenthaler & Ernst, 2008a). Intermediaries can provide resources and capabilities for identifying exchange opportunities (e.g. through their network of resources) and for supporting the innovation transfer (e.g. through their R&D and market experts).

So far, innovation intermediaries have mainly concentrated in marketing and searching for technologies (Morgan & Crowford 1996), but they perform several other functions: identifying partners, helping package technology, selecting suppliers, providing support in deal making, adapt specialized solutions to the needs of user firms (Howells 2006, Lichtenthaler & Ernst 2008a). The functions performed by intermediaries can be seen as related to either of two phases in the innovation process:

1. The search for innovation phase comprises functions performed by intermediaries such as indentifying partners, selecting suppliers, evaluating alternative options;

2. The innovation transfer phase comprises functions such as supporting deal making, packaging the technology, adapting the innovation to the needs of the user firm, transferring related knowledge.

Several web-based intermediaries operate in the innovation market. Yet2.com, for example, is a marketplaces for IP which also offers intermediary services for the adaptation of technologies to the specific needs of the customer (Lichtenthaler & Ernst, 2008a). Ninesigma, provides access to a broad network of scientists, researchers and professionals which are potentially able to solve new technological problems proposed by companies (Tapscott & Williams, 2007; Chesbrough, 2007). Innocentive works like Ninesigma, but technological problems are more narrowly defined, are posted on the company web-site and can be solved by anyone who registers. Yourencore provides access to a network of retired scientists and engineers.
WBIs were expected to dramatically increase the number of transactions in the innovation market by expanding it to a global scale. They are able to provide wide and ubiquitous access to actors, technologies and information. However, few studies on the performance of these internet platforms have been conducted. The few existing studies suggest that some limits in the way WBIs operate make their services unsatisfying for their users. For example, in their study, Lichtenthaler & Ernst (2008b) on IP commercialization by industrial firms through WBIs, found that even if firms show a keen interest in the potential benefits of web-based technology marketplaces, still their attempts to use them were not satisfactory. The conclusions of the two authors suggest that even if WBIs are able to bring together a large number of potential buyers and sellers of innovation, they offer inadequate support in the selection and integration of innovation.

As a consequence, the research question addressed by this paper is: what kind of services and tools should WBIs provide in order to improve the effectiveness of innovation search and transfer?

This paper focuses on how ICT and KM tools could be used to improve the performance of web-based intermediaries in supporting the global innovation markets. Even if we acknowledge that changes in other areas of WBIs’ business models are needed, for example in the value proposition or in the approach to the value network, we focus on technological aspects since we deem that ICT and KM tools can significantly increase the effectiveness of web-based intermediaries if they are used consistently with the needs and structure of global OI processes.

3 A THEORETICAL FRAMEWORK BASED ON SENSE-MAKING AND RELATIVE ABSORPTIVE CAPACITY THEORIES

In this paper we draw on sense-making theory (Weick 1995, 2001, Taylor & Van Every 2000, Weick et al. 2005) and relative absorptive capacity (RAC) theory (Lane & Lubatkin, 1998, Lichtenthaler 2008b) to build a theoretical framework aimed at supporting WBIs in designing and implementing ICT and KMS based solutions. Our hypothesis is that services and tools would be needed to externally complement sense-making and relative absorptive capacity of firms participating in OI processes.

Probably the most common theoretical perspective on innovation exchange is absorptive capacity (AC) theory. Cohen and Levinthal (1990) argue that AC, i.e. a firm’s ability to acquire new knowledge, depends on its level of prior related knowledge and in the field. This definition implies that firms already possess substantial knowledge about an innovation. Through AC firms can advance their knowledge, but the theory does not explain how firms can create or exploit knowledge related to a completely new technology or market. OI posits a new challenge to firms: how to gain competitive advantage from knowledge developed in contexts different than the one where it is going to be used.

We argue that AC’s shortcomings are due to the fact that it focuses on the exchange of innovation within (more or less) homogeneous contexts. In an OI context it is necessary to understand how two potential, previously unrelated partners can be supported in:

1. Recognizing the opportunity for profitably exchanging innovation;
2. Transferring the innovation and adapting it to its new context of use.

While sense-making theory provides insight on the first problem, RAC theory can be used to address the second one.

When the innovation provider is unfamiliar with the context and/or the recipient is unfamiliar with the innovation, a sense-making effort is needed. Sense-making is a process through which circumstances are turned into a “plausible narrative” which substitutes for a rational decision process in order to take action. In the presence of new and complex problems, in fact, a rational decision process is often not feasible (Uren et al. 2006). As a first step in OI exchanges, the potential provider and recipient of technology need to make sense of the use that can be done of the technology in the new context. Sense-making will lead to a decision whether to undertake the exchange or not. Through the services
they provide WBIs are able to support sense-making. The better WBIs support their customers in making sense of a technology (from the seeker’s perspective) or of its potential context of use (from the provider’s perspective) the more it is likely that an exchange will effectively take place.

Lane and Lubatkin (1998) coined the term relative absorptive capacity to point out how the ability to acquire new knowledge also depends on the similarity between the source and the recipient of the exchanged knowledge. They suggest that two firms are more likely to effectively exchange new knowledge if they have similar: 1) knowledge bases; 2) knowledge processing systems and norms; 3) organizational structures; 4) dominant logics.

In order to successfully transfer an innovation, it is necessary to create RAC between the firms participating in the exchange. In our view RAC is not only a structural characteristic of the dyad of companies which determines if an exchange can be successful or not, but a “temporary capability” the two firms can build as a part of the exchange process. WBIs can provide tools and services which support RAC creation within the participating firms. But they can also provide tools and services which substitute for the capabilities firms are missing.

The simplest situation is that of a firm dealing with a familiar technology to improve well known business processes. In this case the best line of action is probably to develop and exploit the technology internally (see figure 1, quadrant 1). When considering external innovation sourcing or exploitation a firm can face one of the three situations represented in figure 1, quadrants 2, 3 and 4. The need to support sense-making and RAC formation depends on the asymmetries between the dyads of firms potentially interested in exchanging innovation. These asymmetries depend on two factors: familiarity and similarity.

“Familiarity” is, for a potential recipient, knowledge of the technology of interest which stems from direct experience. For a potential provider, familiarity represents direct knowledge of the recipient context of use. Familiarity provides a common ground for the mutual understanding of the potential buyer and supplier of technology. As the level of familiarity of the recipient with the technology the innovation is based on decreases, there is a greater need for a sense-making effort on the recipient side (quadrant 2). Familiarity with the innovation increases with the past experience a company has in using (and not necessarily developing) the technology. Similarly a greater sense-making effort is needed on the supplier side when she is not familiar with the potential recipient context of use. The level of familiarity increases if the two companies had previous contacts or one has previous experiences working with the business sector of the other, if personal relationships exist or if they are geographically close.

<table>
<thead>
<tr>
<th>Level of similarity between the two firms (RAC)</th>
<th>Level of familiarity with the partner/innovation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
</tr>
<tr>
<td><strong>High</strong></td>
<td>1. Closed innovation</td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td>3. Absorption effort</td>
</tr>
</tbody>
</table>

*Figure 1. Sense-making, RAC and OI.*
As maintained by Lane and Lubatkin (1998) two firms are similar if they have similar: 1) knowledge bases; 2) knowledge processing systems and norms; 3) organizational structures; 4) dominant logics. RAC theory implies that the level of similarity between the participating firms affects the need for support by a WBI when innovation has to be transferred. In particular the lower the similarity the stronger the need for RAC formation support (quadrant 3). If the two firms are familiar with each other, but they are not similar, problems are likely to arise when innovation has to be transferred (Lane & Lubatkin, 1998). Operative procedures and organizational structures need to be coordinated and differences in dominant logics can cause different priorities. The most complex situation, but also the most likely when OI takes place, is the one represented by quadrant 4. In this case, both a sense-making effort and a RAC formation effort are needed. The difficulties in the search and transfer phases sum up making the support of a wide range of intermediary tools and services critical.

In particular our focus is on the adoption and implementation of ICT and KMS to support the services provided by WBIs. In the following section the above described theoretical framework will be further developed and its implications for the choice of ICT an KM tools will be discussed.

4 IMPLICATIONS FOR ICT AND KMS IMPLEMENTATION BY WEB BASED INTERMEDIARIES FOR OI

4.1 ICT and KMS for enhancing sense-making

“Sense-making involves turning circumstances into a situation that is comprehended explicitly in words and that serves as a springboard into action” (Weick et al. 2005). In the absence of a single canonical view of the world, people must construct ‘plausible narratives’ to fill in the gaps (Uren et al. 2006). Sense making is about explaining phenomena through narratives, given that logical, causal explanations often fail in complex and dynamic contexts. In Weick’s words “we expect to find explicit efforts at sense-making whenever the current state of the world is perceived to be different from the expected state of the world”. Innovation management, then, is a field in which sense-making should have a prominent role, since by definition innovation changes the “states of the world”. Traditional decision support models are based on causal assumptions. For example, effort has been spent in developing quantitative methods for the economic valuation of innovations. The idea is that causal relations can be singled out which link the characteristics of an innovation with the economic value it is able to produce. Accordingly decision-makers will rationally chose the alternative which maximizes the expected returns. In a sense-making perspective the focus is different. It is important to figure out how the technology can be useful and what impact it is likely to have on the business. Sense-making is about the interplay of action and interpretation rather than the influence of evaluation on choice (Weick et al. 2005). According to Bettis and Prahalad (1995) the key problem for an organization is not to accurately assess scarce data, but to interpret an abundance of data into “actionable knowledge”.

We argue, then, that much of the effort spent by actors participating in OI processes, above all in early phases of the search for innovations or innovation applications, is a sense-making effort. In traditional innovation exchanges (e.g. R&D alliances, technology transfer projects, spin-offs), the sense-making effort can go unnoticed by an external observer, since sense-making processes are often carried out through informal interactions and only their formalized outcomes are visible (e.g. contracts, projects, agreements). In an OI context and in the presence of WBIs, however, these informal interactions are often limited. The sense-making problem must be explicitly addressed. The lack of tools supporting sense-making is, in our opinion, an explanation of the limited impact WBIs had so far on the growth of the innovation market. These tools must be implemented and managed by WBIs even if, to be effective, the involvement of potential buyers and sellers of innovation. Such an involvement can be thought as a pre-requisite that WBIs need to achieve in the earliest phases of their interaction with seekers and providers.
The model of sense-making process we use in this paper is the one proposed by Weick, Sutcliffe and Obstfeld (2005). According to the authors sense-making starts with chaos. People are immersed in a flux of events and activities. They may or they may not extract cues from this flow. If they do, the first phase of the sense-making process, namely *noticing and bracketing*, takes place. With reference to innovation, decision makers experience both a flux of information from inside their company and from external sources. When decision makers notice a potential match the sense-making process starts. The second phase is *labeling*, that is people use their experience and mental models to give a name to what they noticed and, so, to stabilize the “streaming of experience”. One recurrent observation is that ICT is used in OI processes in order to make innovations and innovation needs more visible to the participant actors. For example Tao and Magnotta (2006) describe a web-based system called “Needs Tracker” used within a large chemical company. Through this system employees can post the technological needs they noticed. The Needs Tracker helps to make the needs more visible and to rank them. It is also possible to propose solutions to a need. Similar functions are provided by Procter&Gamble’s website InnovationNet, described by Dodgson, Gann and Salter (2006). The company uses the web-site in the context of its Connect&Develop initiative (Huston & Sakkab, 2006) to foster collaboration among its employees and with external sources of knowledge. Artificial intelligence is used for data mining. The system acts in a similar way to Amazon.com, taking into account users’ interests and sending back information the user may be interested in. Data mining is also extensively used by a staff of 70 specialists systematically harvesting web pages, scientific literature and databases and global patent databases. “The change of the (technological) interface demands a change in the organizational ability to absorb, or assess the impressions from the outside” (Fredberg et al. 2008). In other words the use of ICT in these cases has changed the sense-making process by supporting the *noticing and bracketing* and the *labeling* phases. In general web-harvesting, i.e. using different, more or less intelligent, tools to search the Internet for relevant information and knowledge (Carlsson, 2003) can be understood as a means to support the *noticing and bracketing* and the *labeling* phases in the sense-making process. Internet-based toolkits for idea competitions (Piller & Walcher, 2006) are also a powerful means to make a technology need visible to potential solvers. They are used by WBIs such as Innocentive. From a sense-making point of view WBIs provide several traditional tools (e.g. newsletters or alerts) to address noticing, bracketing and labeling. However, the effectiveness of these phases can be improved by introducing innovative tools, e.g. Web 2.0 tools such as collaborative tagging (Golder & Huberman, 2006).

The third phase is *thinking retrospectively*. During this phase several cues noticed and labeled before are put together. A plausible narrative is created. Tools such as computer supported argument visualization have proven effective in supporting sense-making about ongoing scientific or professional debates (Shum, 2003). Similarly they can be used to trace debates about innovations or innovation applications. Knowledge mapping or knowledge cartography visually display the conceptual structure of ideas (Okada et al. 2008). By introducing these tools WBIs can support customers in understanding how an innovation could be adapted to a new context from several points of view: it is possible to link a technology to several actual or potential applications, to know about needed complementary competences and resources, to understand who possesses the necessary knowledge and to take into account intellectual property issues. This possibility can be very valuable since one of the problems pointed out by Chesbrough (2006) is the difficulty, for firms, of making sense of the interdependencies among different aspects of OI initiatives (technology, business model, intellectual property rights).

During the following phase, *supposing*, a tentative narrative is created to link the pieces together. Also in OI processes there is a need to build a tentative narrative to make sense of the potential applications of the innovation. Only through confrontation with other people, however, the narrative is rejected or becomes an accepted argumentation. *Socializing*, the fifth phase, explains why sense-making is a collective process: the explanation developed by an isolated individual is influenced by the interpretations of other people she is going to interact with if she is going to put her intentions into action. The subsequent actions will become experiences influencing a person mental models and, as a consequence, her future sense-making. As Web 2.0 tools become more common also in a business
environment (McAfee, 2006; Bardhan et al. 2008) the possibility to support the socializing phase of the sense-making process increases. In Web 2.0 tools the supposing and socializing phases are tightly intertwined, since it is possible to propose opinions, interpretations, points of view which are collaboratively discussed. Some WBIs are already introducing features based on the Web 2.0 approach. For example, Innocentive recently introduced a blog with all the (by now common) Web 2.0 functionalities like tags and feeds. In its “Innovation Community” section, YourEncore provides a full range of Web 2.0 tools like Wikis, forums and other tools to help clients collaborate with experts. These tools can be used independently from the other services, and are aimed at creating an ongoing discourse among the participants. A different way to support the supposing phase is through simulation and modelling or virtual prototyping tools (Dodgson et al. 2006). However, in this case a firm has already a good understanding of the innovation and of its applications. As a consequence, in our framework, simulation and modelling tools are better understood as tools for enhancing RAC. Malhotra (2001) suggests that Artificial Intelligence and Expert Systems could be used for supporting the supposing phase if designed to “encourage ongoing and continual re-assessment and modification of practices to ensure dynamic adaptability to the rapidly changing business environment”.

In the last phases of the sense-making process, action takes place and, in an organized context, it is likely to require communication with other people. The last phase, then, is called organizing through communication. In an OI context both socializing and organizing through communication require interaction with external actors. Both Web 2.0 tools and more traditional e-collaboration tools can be used to support the last phase of the sense-making process, that is organizing through communication. In particular e-collaboration tools such as whiteboards, document management systems, collaborative project management systems, file sharing can be used when translating the collaboratively created interpretation of the situation into action (Migliarese & Corvello, 2006; Fink, 2007; Kumar & Becerra-Fernandez, 2007).

4.2 ICT and KMS for enhancing Relative Absorptive Capacity

In Cohen and Levinthal’s approach, AC depends only on the previous related knowledge a firm possesses. As a consequence “a firm has an equal capacity to learn from all other organizations”. Lane and Lubatkin (1998), instead, argue that AC depends also on the similarity between the two firms exchanging knowledge. So it is better understood as relative absorptive capacity. While in Lane and Lubatkin’s paper RAC is described as a given attribute of the dyad, other research suggests that RAC can also be created for a specific exchange (Lichtenthaler 2008b).

If the knowledge exchange takes place without the intervention of an intermediary, only the characteristics of the participants (i.e. organizational form, processes and dominant logics) influence RAC. In the presence of an intermediary, however, the services and structures it provides can influence the capacity of the participating firms to exchange innovation. This is consistent with the idea that firms can draw on the resources and capabilities of intermediaries to improve their proficiency in exchanging innovation (Makadok 2001; Foss & Ishikawa 2007; Lichtenthaler & Ernst, 2008a). There are five modes an intermediary can contribute to the development of RAC:

1. Create a common knowledge base: intermediaries can provide firms with knowledge related to interdisciplinary (i.e. issues common to several technological domains) or complementary aspects (e.g. issues related to problems such as intellectual property rights, regulatory issues, electronic infrastructures) useful in more than one exchange;

2. Create field specific knowledge bases: intermediaries can collect, organize and package knowledge related to each specific domain to be provided to the partners once the exchange has been decided in order to speed up the development of a common domain specific knowledge base. This issue probably requires the intermediaries to be specialized in a limited number of domains;

3. Accelerate knowledge transfer: this is the most intuitive function an intermediary can provide for enhancing AC. The sooner knowledge is transferred, the easier is to proceed in the exchange;
4. Develop standard methods: by using standard methods provided by the intermediary (including standard documents, procedures and technologies) the participants can partially overcome the problem of different organizational processes;

5. Act as a temporary structure for innovation transfer: members of the participating firms and of the intermediary can work together as a temporary structure able to limit the problem of different organizational structures.

There is a fairly broad literature on tools supporting the creation of knowledge bases and on knowledge transfer. In fact these two aspects can be considered the central functions of a KMS (e.g. Robey et al. 2000).

With reference to the creation of RAC the first opportunity for intermediaries is to create a common knowledge base even before an innovation exchange is envisioned. Some pieces of knowledge can be useful for several different innovation exchanges. For example complementary knowledge such as knowledge about intellectual property rights or regulatory issues (Somaya et al. 2007). Also complementary technical knowledge can be used in several projects. For example knowledge related to soldering techniques is needed for different applications. Building searchable databases or preparing documents and tutorials related to these topics (in other words “packaging” the related knowledge) can support companies which at the moment of the exchange will be prepared to manage complementary aspects. Collaborative tools as document sharing, forums, blogs and wikis can also be used in order to create a common knowledge base, in particular with respect to more unstructured issues.

Since intermediaries are exposed to diverse knowledge in different exchanges, they gain relevant experience related to several specialist fields. The availability of expert individuals at an intermediary is much appreciated by customers. For example Yet2.com has recently shifted its services from marketing technologies to assisting its customers from a technological point of view (Lichtenthaler & Ernst 2008b). Such services can be made more efficient through the use of KMS. They can be used, in fact, to organize knowledge by specialist domain and make it available to customers. Automation of data retrieval and use (Robey et al. 2000) is complementary to the creation of knowledge bases. Besides capturing knowledge and making it easily accessible, also providing opportunities and tools for communication and discourse is important to speed up knowledge transfer (e.g. Robey et al. 2000). Collaboration tools such as Lotus Notes are still widely used to support collaborative work. Web 2.0 technologies provide further possibilities to cooperate and exchange knowledge (McAfee, 2006).

Creating standard methods is also a much appreciated feature of WBI. Innocentive, Yet2.com, Yourencore and the other WBI pay great attention in communicating to their customers how the methods they developed are able to make the exchange easier. They continuously modify their methods in order to adapt them to emerging customer needs. The methods developed by WBI are meant to coordinate the processes of innovation seekers and providers. For example Innocentive provides consultancy services and formats to firms seeking a new technology in order to formalize a technology need (which in Innocetive’s language is called “a challenge”). At the same time it provides solvers with interaction procedures consistent with the expectations of the seekers. The interaction takes place in a structured virtual room dedicated to the specific challenge. Structuring the innovation process through ICT and KMS, however, can also have drawbacks, since it is possible that established methods will not be revised and will become not adaptable to changing conditions (Robey et al. 2000).

Finally WBI can enhance RAC by providing virtual organizational structures for managing the innovation exchange. Organizational structures are virtual when they are temporary, geographically dispersed and based on electronic communication (Jarvenpaa & Leidner, 1999). Virtual organizational structures are especially common in R&D, due to the internationalization of research and to the frequent formation of R&D partnerships (e.g. Hagedoorn 2002). WBI, through their web-sites and their permanent structures are able to create temporary groups in which members of the provider and of the recipient organizations are involved. Such groups, guided by the norms and methods set by the intermediary, carry out the innovation exchange overcoming the difficulties created by differences in organizational structures.
5 CONCLUSIONS AND FUTURE RESEARCH

The enthusiasm about web-based intermediaries for OI is due to the promising features they show. In particular the possibility to bring together a large number of potential buyers and sellers of innovation, to act on a global scale and to provide structured and searchable information about innovation and innovation needs. The use of ICT and KMS by WBIs, however, seems to be limited to more or less advanced search techniques.

In the previous section we showed how a sense-making and an absorption effort are needed in OI exchanges. We argue that WBIs should provide tools and services able to support sense-making and the creation of RAC. Recent developments in the services, provided tools and business models of some web-based intermediaries for OI seem to support this hypothesis.

Since sense-making is an unstructured process, which produces unstructured outcomes (it produces sense, exactly), also the tools needed have been described by grouping them according to the phase of the process they support. The same tool, however, can be used in more than a phase. It is the way the user approaches the tool that makes the difference between phases (and, actually, also between sense-making and RAC creation). RAC creation, instead, produces structured outcomes that can be used also in future exchanges. So we grouped the tools for RAC creation according to the specific outcome they contribute to produce.

We expect that WBIs effectively supporting sense-making will experience a larger number of exchanges. On the other hand the percentage of successful exchanges will increase if RAC formation is effectively supported. The importance of sense-making and RAC creation support is likely to vary according to the context. In particular if the involved companies are familiar with each other and/or they show high levels of RAC the role of an intermediary will be less critical from a knowledge related point of view.

In this paper the two theories have been discussed separately. We presented OI processes as comprised of two sequential phases: recognizing the opportunity for profitably exchanging innovation and transferring the innovation and adapting it to its new context of use. While sense-making theory supports the analysis of the first phase, RAC theory supports the analysis of the second one. We have discussed OI processes considering an isolated exchange between firms which are characterized by low familiarity and low similarity. In this case the participating actors need to go through the whole sense making process and to spend a strong effort in building RAC. In practice, however, the two phases overlap. The same tools can be used to support both sense-making and RAC formation and there is not a temporal separation between sense-making and RAC creation. The difference is mainly in the use of the (often common) outcomes produced by the two processes. A database can be created and used both to support sense-making and RAC creation, but it is used differently (and probably using different functionalities) in the two cases. It is used in a structured and systematic way when supporting RAC, it is used in a less structured way when supporting sense-making.

The sense-making process influences the RAC creation process, by posing emphasis on same aspects of the innovation problem rather than on others.

If we consider repeated exchanges it is likely that much sense and RAC will be inherited from previous exchanges. In this case the sense-making process in a new exchange will build on the same knowledge made available during RAC creation in previous exchange which, in turn, had been influenced by the sense-making process.

In repeated interactions, sense-making and RAC creation intertwine and it is not possible to distinguish activities aimed at sense-making from activities oriented at RAC creation. This does not mean, however, that mechanisms to support both processes do not need to be devised and properly designed.
To our knowledge no other study has applied sense-making theory to OI. Decisions regarding OI, however, rely on highly ambiguous and uncertain data. As a consequence managers are more likely to use plausible narratives than rational methods when deciding to externally source/exploit innovation. In our opinion, then, sense-making support by WBIs is often more valuable than other forms of decision support. Few studies also exist on RAC in OI processes (Lichtenthaler 2008b). In this paper we considered RAC as a temporary capability which can be developed as a part of an exchange process relying also on the complementary capabilities of an intermediary. Supporting the rapid development of RAC is both a need for companies involved in global innovation exchanges and an opportunity for WBIs. The use of ICT to globally source or exploit innovation is still limited. In this paper we argue that a proper use of the available tools by web-based intermediaries can provide the innovation market with the needed liquidity, eventually enhancing OI on a global scale.

We suggest two main line of future research: 1) behavioral science research, and 2) design science research (Hevner & Chatterjee 2009). Future behavioral science research will, based on our two underpinning theories, describe and explain how ICT and KMS can enable and support sense-making and RAC formation. The intrinsically unstructured nature of the sense-making process makes the measurement of related constructs especially challenging. In the Absorptive Capacity field, instead, extensive research has been conducted which may provide insights into empirically examining the formation of RAC. Research aimed to the study of the performance of WBIs is needed, for example exploratory research involving interviews to WBIs’ personnel. This kind of research could also support the refinement of our framework. Future design science research will, based on our two underpinning theories, develop practical design knowledge for the design and implementation of WBIs. The design knowledge can be in the form of algorithmic or heuristic design propositions, design exemplars, design models or frameworks, and stories or narratives.

References


DIGITAL LIBRARIES AS INFORMATION ORGANIZATIONS. THE RE-UNFOLDING OF THE MEMORY/INFORMATION PARADOX

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0671.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Digital Libraries, Information Growth, Information, Organizational theory</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
DIGITAL LIBRARIES AS INFORMATION ORGANIZATIONS
THE RE-UNFOLDING OF THE MEMORY/INFORMATION PARADOX

Marton, Attila, London School of Economics and Political Science, Houghton Street, WC2A 2AE London, UK, a.marton@lse.ac.uk

Abstract

Throughout history, libraries have played a key role in remembering the past continuously adapting to changes in societal communication and its technologies. Digital libraries prove to be the next step of unfolding the memory/information paradox which is the foundation of what makes a library a library. Libraries are again changing in order to remain libraries. The focus has begun to shift from being an archive of knowledge containers – the memory side of the paradox - to being an organizer of information taking the form of information organizations. In this sense, a library does not use an information system but rather is an information system. Contemporary developments in ICT seem to harbour the capability to transform the memory/information paradox into a solvable technical problem making libraries themselves an institution of the past. At a closer look, however, the paradox re-emerges provoking the dynamics behind the discussions of what a library actually is and does in the 21st century. The conceptual elaboration of contemporary libraries as information organizations serves as a sensitizing tool for the social science study of digital libraries but also as a key to introducing digital libraries into the Information Systems community.

Keywords: digital libraries, information organizations, information systems, information growth
1 INTRODUCTION

Those who cannot remember the past are condemned to repeat it. Libraries have a long standing tradition of organizing the past by using contemporary technological innovations of their time for, what we call today, data editing and ordering. As information organizations, the changes, libraries have gone through over the centuries, have been tightly coupled to the developments in communication technologies - be it the shift from manuscripts to print books or the diffusion of binary-based communication media into basically every aspect of social life (Bennett 2001). However, harbingers of the rise of the Internet, such as Weinberger (2007), proclaim the end of pre-ante categorizations of knowledge organized by professional librarians; instead, we all will join into running the limitless library of libraries – the Internet. More cautious voices denounce the Internet being The Alexandrian Digital Library as a mere myth arguing “[t]echnological progress has changed how libraries do their work, not why. […] technology will not substantially alter the business of librarians – connecting people with information” (Kuny and Cleveland 1996:1).

In the following paper, we will rephrase the “why libraries do what they do” as a paradoxical form of memory and information. This distinction will serve as a conceptual frame for observing historical as well as contemporary developments within the domain of librarianship which, in addition, will also serve to identify a library in difference to other forms of organizations. According to Luhmann (2006), a social system is observable, hence identifiable, because of its difference to its environment. It is a paradox. A library, for instance, is what it is, because it is what it is not at the same time. Paradoxa, however, are not solvable problems; they can merely be ignored, hidden, made fun of, or be a reason for creativity (Bateson 2000); in more general terms, a paradox can only be unfolded creating a dynamic that might lead into a self-organizational, autopoietic form (Maturana and Varela 1992, Luhmann 1996). Consequently, we shall see the techné of librarianship as ways of unfolding the memory/information paradox that finally led to the emergence of an autonomous organization – the modern library. Digital libraries may be seen as a phase of renegotiating the difference a library needs to maintain, in order to remain a library in a changing environment – the memory/information paradox is being re-unfolded.

The aim of the paper is to refocus the conceptual framework of a research-in-progress on digital libraries by formulating a problem rather than reporting on preliminary results. In order to make a first step from getting-a-sense-of to making-sense-of the data, it is necessary to conceptualize what makes a library observable as a library (and not as an archive, book store, bank, and so forth). Following Weick’s (2007:14) argument that “perception without conception is blind”, the aim of the paper is to enrich the observation of the phenomenon. Theorizing is to be seen as a sensitizing device by making assumptions and expectations explicit. Only with a conceptual map at hand, can an observer appreciate the richness of the observed, be open for surprises that were not captured by the map in the first place.

The second aim of the paper is to open up libraries and librarianship as a fruitful topic for IS research. The key viewpoint is the observation of libraries as information organizations based on following points: 1) libraries have always used information systems in their daily routines – be it shelf-, paper-, or data-based. Constructing and maintaining an information system is the core competency of any library. In other words, it is impossible to separate the organization library from the information system it uses. The library is the catalogue. 2) Opposed to other forms of organizations, like for instance profit oriented firms, libraries have a self-referential approach to what information is and how it should be organized which might contribute to similar discussions on information within the IS discipline (Boland 1987, Bryant 2008).

The paper begins with a brief outline of librarian history as a process of internal differentiation of, sometimes referred to as memory institutions (libraries, archives, museums) in relation to more general developments in communication technologies. Section three introduces the reader to points raised in the Library and Information Science community in relation to digital libraries. In section 4
the memory/information paradox will be elaborated in relation to developments in IT. Section 5 concludes the paper by arguing for the viability of the concept “information organization” in relation to future research efforts on digital libraries.

I would like to thank Jannis Kallinikos, Giovan Francesco Lanzara, Nathalie Mitev, Katerina Voutsina, Eleni Lamprou, and the anonymous reviewers for their constructive criticism and motivating support.

2 A BRIEF HISTORY OF LIBRARIES

A history of libraries can be constructed in relation to developments in communication technologies. In fact, it can be seen as a co-evolutionary relationship in terms of innovation and diffusion, meaning that the diffusion of new communication media (e.g. books) is coupled to librarianship developments (e.g. public libraries). Hence, the following outline should not be read as a unilateral cause-effect relationship but rather as a feedback loop between the library and its communicative environment mediated by certain technologies (Orr 1977).

2.1 Libraries in Ancient and Medieval Times

Following Niklas Luhmann’s communication model (Luhmann 1996), a communication technology separates the direct and mutual observation of communicative participants. Opposed to talking face to face, the utterance is de-contextualized and needs to be re-contextualized by the receiver in order to be (mis)understood. Hence, the written word gains a context-independent existence. That does not yet apply to the very first forms of pictographic writing like the hieroglyphs of ancient Egypt developed around 3000 B.C., since those do not mediate communication but rather representations of things. Pictograms allow for a wide range of phonetic, semantic, and grammatical interpretations that requires adequate knowledge in order to be read and, therefore, are hardly able to mediate something the reader did not already know. Consequently, the very early pieces of writing were memory devices rather than communication technologies (Esposito 2002:45-47). Consequently, libraries were storing what was deemed worth to be remembered. Still, “[…] there was no separation between book collections and archives; books and documents had the same outward appearance and required similar methods of storage” (Dahl 1958:11). At that time, libraries were simply storage compartments without any organizational autonomy. They were as undifferentiated as the papyrus rolls and cuneiform tablets they stored.

A crucial innovation in terms of communication technologies is the phonetic alphabet. The simplification of writing into a digital system\(^1\) of consonants in 2nd millennium B.C. Palestine and Syria, complemented with vowels roughly a millennium later in ancient Greece, allowed the text to become completely context free and autonomous (Borgmann 1999:45-47). While reading ideograms still relies heavily on the contextual knowledge and memory of the reader, the reading of a text based on a phonetic alphabet relieves the reader to have any contextual knowledge except how to pronounce the letters. In fact, a person can read such a text without understanding a single word. Although, alphabetic writing was firstly used as a support for mnemonic techniques (Luhmann 1998:511), it already allowed the mediation of content completely new to the reader, it allowed for the reader to be surprised – a very important aspect of information, as we will discuss further below. The classic example in terms of librarianship is, of course, the Great Library of Alexandria with an estimated amount of 743,000 scrolls collected through 3rd century B.C. Although its largest collection was still part of the temple of the muses – the Museion – we can already distinguish librarian practices and terms. For instance, “bibliotheka” was the word for the wooden or stone jars in which rolls were kept that also became the term for a collection of scrolls (Dahl 1958:18-20). The scrolls themselves were

---

\(^1\) In this sense, digital does not mean binary but rather the rendering of speech into a limited number of discrete characters.
already organized by a professional class of scholars according to subject criteria, however, the catalogue was merely an alphabetic inventory list of knowledge areas and authors (Orr 1977:126). In other words, while the scrolls themselves already harboured the potential for reading to shift from remembering to learning, the catalogue still required a lot of contextual knowledge by the librarian. One had to know what one was looking for. Opposed to being a storage compartment, the library slowly turned into a warehouse with an inventory list.

While the rudimentary communication technologies changed from the usage of papyrus to parchment over the centuries, evolving into standardized documents and codices in the Roman Empire around 2nd century B.C. (Landheer 1957:18), the basic setup of libraries did not change until the final stages of the Medieval era (Dahl 1958). Parchment was increasingly used for administrative records leading to the differentiation of libraries and archives. The import of paper production techniques from China finalized the evolution of the manuscript as a bounded entity. However, throughout the Medieval Ages, libraries (opposed to archives) were still part of religious institutions like monasteries or teaching mosques mostly dedicated to copying and preserving existing manuscripts. There was no need to change librarian routines of keeping an inventory, since a single library had to take care of a steady and rather small number of manuscripts. For instance, one of the largest monastery libraries – the Bobbio monestary library – had a collection of about 700 volumes in the 9th century (Dahl 1958:57). With the rise of universities around 12th century A.D. and the increasing number of quite large private collections throughout the Renaissance, the dominance of monastery librarianship came slowly to an end. However, the first major shift in librarianship was initiated a few hundred years later by the diffusion of the printing press and the rise of mass media.

2.2 Libraries and the Emergence of Printed Mass Media

The invention of moveable type printing by Gutenberg in the middle of the 15th century can surely be called revolutionary. Although this technology was already well established in China, it was not very successful because of the extensive number of ideograms (4-5000 different types) required to print a book (Dahl 1958:84). In comparison, the system of relatively few Latin letters turned out to be an enabler for the mass production of media. Only in the first 50 years after Gutenberg’s invention, 15-20 million books were produced (Weinberger 2007). Of course, this caused libraries to reorganize their items according to size rather than topic to save shelf space. Still, the printing press also brought qualitative changes along. Besides the standardization of the book in terms of size and structure (e.g. usage of titles, paragraphs, or page numbers), it was also necessary to standardize languages in order to minimize variety to reach as big an audience as possible – a process which led to the establishment of national languages from around the 16th century on. Most importantly, printing finalized the shift from what is known to what is new and interesting – from the oral tradition of repetition and devotion (Luhmann 1998:295-299).

The differentiation of books into a communication media rather than a memory device co-evolved with the emergence of libraries as autonomous organizations by gaining an educational aspect in addition to the preservation of knowledge (Landheer 1957:98). This was achieved by opening up the private collections of the aristocracy to the public in the 17th century. Of course, at that time the public was yet limited to a small elite (Dahl 1958:177). However, printing was not only dedicated to the production of books but also of pamphlets, leaflets, and newspapers that addressed a growing literate portion of the population. The final innovation that popularized the reading of books was the novel in the 18th century. Accompanied by the invention of the paper-making machine in 1799 and the power press in 1810, reading became a comparatively cheap undertaking (Dahl 1958:220). In England for instance, so called circulating libraries rented novels out and further promoted literacy in the population to an estimated amount of 5 million until 1850 (Shapiro and Varian 1999:95). Hence, the beginning of an internal differentiation of libraries into public libraries as opposed to national and research libraries can be observed in those times. It is also the beginning of the shift from memory to information.
The establishment of libraries as learning facilities for a wide range of people and the immense growth of books and other print media marks the era of the modern library from the mid 19th century on (Thompson 1982, Svenonius 2000:2). Obviously, books – being the prime collection item of any library – are quite useful for diffusion into a wider population as they are controlled by going through a publication process, mass produced, and easy to use. Still, the notion that books preserve knowledge and hence libraries need to preserve books was and still is very much alive which Ranganathan (1931:2) came to call a “tendency to hoard books”. The user-friendliness of libraries, propagated by the British Museum Librarian Panizzi in the 1850ies, towards open access to the items themselves collided with the need to store the items as efficiently as possible (Dahl 1958, Weinberger 2007). In general, public libraries allow direct access by ways of organizing the items themselves. Books are being shelved according to the categories they are assigned to; hence, the way books are displayed is already informative. The dominant system, especially in Anglo-American libraries, is the infamous Dewey Decimal Classification System (DDC) developed by Melvil Dewey in the 1870ies that classifies all the world’s knowledge into 10 main categories and 100 sub divisions per category (OCLC 2003). Admittedly quite an ambitious task, it produces some biased or even embarrassing classifications because of its Anglo-American roots and the limited number of categories. For instance, the 100 divisions for the main category “Religion” are mainly dedicated to Christianity while there is only one division for Islam. As ambivalent as these open shelf arrangements may be, they allow the user to browse through the collection, therefore, to find something without knowing exactly what to look for. The library, rather being an inventorized warehouse, becomes “walkable” for usage and discovery.

On the other hand, closed shelving applies to the immense growth of books allowing indirect user access via various retrieval systems. The shift from book catalogues to card catalogues is the main result. Card catalogues, first formulated by the Smithsonian librarian Charles Coffin Jewett in 1852 based on the work of Panizzi (Weinberger 2007:59), can organize an unlimited knowledge space since it is not bounded by the limitations of books. However, browsing through a library’s stock is not possible anymore, since the user needs to comply with the librarian way of describing items. The author catalogue is only usable when one knows exactly what one is looking for. From this perspective, a closed shelf library still remains a warehouse with an inventory that only answers questions like “Does the library have the book X by the author Y?” In order for a library to be used by patrons (this is the differentiating factor between a library and an archive) it needs to provide services enabling a user to find what he/she does not know (Ranganathan 1931, Thompson 1982:100ff). Subject catalogues are of limited help as they are mostly based on a controlled vocabulary, meaning that items are indexed according to a standardized set of categories a user needs to familiarize with. From this perspective, a library only provides a limited number of very specific paths for discovering items. In more general terms, librarianship refocuses from the problem of preserving knowledge to a problem of retrieval – of being informative in terms of its items (Orr 1977:125ff, Hjorland 2000).

Be as it may, while industrial mass production finalized the standardization of formats like books or newspapers, the card catalogue standardized how librarian items are to be described. Already to be found in Jewett’s conceptualization, the standardization allowed for the merging of catalogues into union catalogues, first on a national and ultimately on an international level. The consequences were, for instance, that from the early 20th century on, loaning books among libraries became possible. Another consequence was the establishment of national libraries as reference libraries distributing printed catalogue cards in order to maintain the nation wide but also international quality of the catalogues (Dahl 1958:255). Especially among national libraries, cooperation became a standard practice particularly after the Second World War.

The exponential growth in the mass production of books – what today would be called information growth (Kallinikos 2006) – being archived by libraries is, in fact, an immense problem. Various authors came to question whether libraries will be capable to not only store books but also to maintain a certain level of usability or whether we would witness the decline of libraries into unusability (Thompson 1982). Various technological innovations, like microfilming, may have had temporarily
solved the storage problem, but the biggest hope is put into computerization and digitalization not only in terms of storage and preservation but also in terms of access and retrieval (Schwartz 2000). First steps were the image scanning of catalogue cards later followed by OCR scanning accessible via various Online Public Access Catalogues (OPAC). Data-based catalogues are now able to use a wide variety of IT enabled search and retrieval functionalities such as Boolean operators or ranking and relevance feedback (Hahn 1998). However, the ultimate step is the digitalization of librarian items themselves which enables users not only to retrieve descriptions about specific items but rather to access the items themselves via the WWW.

3 ON DIGITAL LIBRARIES

The domain of Library and Information Science (LIS) observes and discusses changes of librarianship attributed to the rise of the Internet (Davis and Lagoze 2000); this includes changing user behaviour (Peterson Bishop 1999), the increasing importance of documentation and preservation of online communication, especially in science and research (Ercegovac 1997), and comparisons with new information service providers of which Google is the most prominent one (Schwartz 2000, Bjorner 2006). An indicator for the ongoing discussion is the lack for a standard definition of digital libraries (Meyyappan, Chowdhury and Foo 2000). Schwartz (2000), for instance, discovered 64 different definitions of digital libraries that vary from very strict characterizations to very loose ones. If we compare these definitions with a traditional definition of a library as, for instance, presented by Oppenheim and Smithson (1999:99); “The traditional library is defined as a specific place with a finite collection of tangible information and it is geographically constrained”, we may come to the conclusion that either the library does not change a lot or it changes tremendously. Either a library simply includes digital works into their collections and offers corresponding services or it leaves “the place” behind offering any document any time to anyone in any place (Covi and Kling 1996:672) – a library without walls.

A second very important issue is the relationship between librarianship and ICT. Digital libraries are often positioned between these two domains (Oppenheim and Smithson 1999, Schwartz 2000). The discussion unfolds along three major dimensions; digitalization and preservation of digital media, accessibility for library users, and interoperability of librarian cataloguing. The first point refers to the difficulties of not only digitizing analogue media but also to keep the digital content retrievable at any time. Decisions made during digitalization (e.g. file structure) have a huge influence on the preservation policy and vice versa. That is because digital documents are unable to “care for themselves” (Russell, Weinberger and Stone 1999:277), for digital media requires the appropriate hard- and software in order to be accessible (Bennett 2001, Kallinikos 2009).

Accessibility mostly refers to the interaction between the library and the library user. ICT does not only enable multi-dimensional search functionalities through the library’s repository but also new meeting points for face-to-screen interaction and collaboration. The discussion covers a wide range from usability in terms of screen- and interface design (Thong, Hong and Tam 2002) and the enhancement of the readability of texts on screen (Greene, Marchionini, Plaisant and Shneiderman 2000), the design of systems to improve the relevance of results of user enquiries (Marcum 2001, Tuominen, Talja and Savolainen 2003), to new tools for especially scientific users in terms of collaborative research and user behaviour (Wilensky 2000). The final point mainly discusses concerns regarding the interoperability between librarian standards of classification and metadata (Suleman and Fox 2001). Metadata standards play a key role in the future development of librarianship (Baker 2006), since they build the basis for search, retrieval, delivery, rights management, and preservation (Russell, Weinberger and Stone 1999).

Instead of observing the discussions surrounding the issue of digital libraries as a problem of developing and implementing the right technologies, we should rather treat them as indicators for the re-emergence of the more fundamental theme of what makes a library a library. In other words, the memory/information paradox needs to be re-unfolded in relation to a changing environment. As it will
be discussed further below, IT is one way to attempt a re-unfolding by transforming the paradox into a technical problem. The rest of the paper will be dedicated to further explore what is actually meant by memory and information in relation to the latest innovation in communication technology – binary-based media.

4 MEMORY/INFORMATION AND IT

The ways traditional libraries organize the items and their descriptions lies in the nature of the items themselves. Books, for instance, are bounded physical objects that can only occupy one place on a shelf at a time. As it is the case with the Dewey Decimal Classification System, the organization of the items turns into a system of either-or categories. Of course, all the books could have been piled up into an immense heap but, as a consequence, finding what one was looking for would turn into an event of luck. Therefore, having any order is better than having no order at all. In other words, one can find something because not everything is findable. Naturally, any classification system hides and distorts. This already begins with the collection policy of each library that is, in fact, a selection policy. In principle, that does not change with the introduction of a separated second-order search and retrieval system – the card catalogue. It only allows answers to very specific questions, on the one hand, by relying on the standardized attributes of books (the name of the author or title) and, on the other hand, by relying on librarian standard procedures of adding further descriptions such as keywords which are based on an existing pre-ante classification system. However, with the rise of ICT and ultimately the Internet, the established formats do not apply to new forms of online communication such as hypertexts.

In order to avert misunderstanding, digitalization should not be understood as the end of books or paper-based media in general. As Brown and Duguid (2000) argue, digitalization does not mean de-paperization. The paper-based domain of books and newspapers, on the one hand, and the binary-based online world, on the other hand, are not mutually exclusive but rather complement one another. Still, librarian 1st order categorization – organizing the books themselves – as well as 2nd order categorization through organizing descriptions of books – the catalogue – are based on the standardization of specific communication technologies as packaged media (Esposito 2002). Books, seen as containers of knowledge, have been organized the same way as any other mass produced good (Svenonius 2000:10). To some extent, this is still applicable to packaged digital media such as CD-ROMs or digitalized versions of books. However, the rise of hypertextual, networked media (Esposito 2002), of dynamic, unstable, borderless online digital media (Svenonius 2000:12) brought librarian collection practices back into the centre of attention2. So what is a library actually organizing – is it really simply books and the knowledge they contain?

4.1 What is a Library organizing?

Roughly in the middle of the 20th century, the incommensurable growth in mass printed media made librarians rethink the role of libraries similar to the situation today. The answer was the introduction of another difference – that is the difference between work and item. As Svenonius (2000:9) reports, the first explicit distinction was made by Julia Pettee – chief cataloguer of the Library of Union Theological Seminary in New York City – in 1936. She described a particular message content as a “library unit” and its embodiment in a medium as a “book”. In 1955, Ranganathan laid the foundation of what came to be the Functional Requirements for Bibliographic Records (FRBR). Ranganathan distinguishes between “work” as an expressed thought and “document” as an embodied thought. Finally in 1995, the International Federation of Library Associations and Institutions (IFLA) developed a four level model of work, expression, manifestation, and item (IFLA 1998). Work is

2 At a closer look, not even a book is as stable as it may seem if we, for instance, think of different editions and translations.
defined as a distinct intellectual or artistic creation – an abstract concept that is realized through an expression in the form of alpha-numeric, musical, or choreographic notation, sound, image, object, movement, etc., or any combination of such forms. The physical embodiment of an expression of a work is referred to as manifestation. Finally, an item is a single exemplar of a manifestation. For instance, Shakespeare’s Hamlet (work) is realized as various English versions and translations into other languages (expressions) which are published as hard-cover and paperback editions (manifestations) and one exemplare of each (items) is available in a library. Now, a peculiarity of modern libraries, since they are dealing with mass-media rather than unique archival records, is the organization of items (things) and works (abstract concepts) (Svenonius 2000:10). As Svenonius (2000:11) further elaborates in reference to Panizzi, this is quite an important distinction, since a user may know the work but not the peculiarities of different editions.

The introduction of IT into librarian services in the form of databases put a lot of emphasis on the item level. Information technology treats information like a thing that can be manipulated, ordered, and categorized at will, or at least so it seems (Buckland 1991). Of course, this notion is based on a very specific definition of information provided by Information Theory founded by the work of Claude Elwood Shannon (1993). Accordingly, information is a selection of a message out of a number of pre-selected messages based on the design of the communication channel between a sender and a receiver. The information value of the message sent derives from the number of messages it was selected from. Therefore, the informativeness of a message lies not in the message itself but in what else could have been sent. The basic measurement unit is a binary digit of 0 and 1. For example, sending the result of flipping a coin has an information value of 1 bit – 0 for head, 1 for tail. Shannon’s model is a basic input/output model of what comes in should come out at the other end as efficiently and undistorted as possible. However, in this model we do not find a differentiation between message or signal, on one hand, and information on the other hand. In other words, information and document – or in other words data and information - are one and the same (Lee 2004). Treating information as a countable thing, however, leaves the receiver completely out of the picture. In order to understand, the receiver needs to distinguish between noise and message which is based on her/his expectations. In fact, when pure noise complements the receiver’s expectations it is meaningful for her/him hence informative without the intention of a sender. Information lies in the eye of the beholder and is not a factual thing with an existence of its own. Consequently, even the absence of a signal may be informative for a receiver or as Bateson (2000:458) puts it: “The letter which you do not write can get an angry reply”.

Accordingly, books and catalogues may not be considered as containing information but rather having the potential of triggering the event of informing a reader and user. What books and catalogues contain is mere data that is organized in a way that is expected (by the author/publisher/librarian) to be potentially informative for somebody (an anonymous mass audience of readers, library users). However, the technological innovations in terms of data processing and ordering propel the librarian abilities to deal with items (things) but fail to notice the importance of organizing works (abstract concepts) as well (Svenonius 2000).

4.2 Memory and Information

The observer-related nature of information has been expressed by Bateson (2000:459) in his renown definition of information as being a difference which makes a difference. This concept has mostly been interpreted as information having an element of surprise or novelty (e.g. Luhmann 1996, Kallinikos 2006). In other words, a difference has to be informative for an observer and only then does information occur (Wildavsky 1983). Information has no existence of its own; it is a no-thing, an event that loses its informativeness for an observer the moment it is observed. This led Kallinikos (2006) to call information non-foundational and ephemeral, hence it can neither be stored nor produced. In short, information is not what triggers change nor what is changed but the change itself. Based on this concept, books contain potentialities for information for somebody. For instance, reading an English text without knowing English is of course possible but for the reader the work remains pure noise
without any meaning. Now, this is the crucial aspect that allows for a further elaboration of the shift of writing/printing from being mostly a memory device to a communication technology that mediates what the reader does not know. Mass media in general (e.g. TV channels, newspapers, radio channels) deal foremost with delivering exactly that (Luhmann 1996). Information, therefore, is closely attuned to learning rather than to knowing (Wildavsky 1983).

The same applies to the ways library items are organized and documented. Beginning with a mere inventory list, the documentation had no potential of informing a user. It only told where to find a specific scroll, manuscript, or book the user already knew about. To be precise, we cannot even think of a library being used at that point. A library was a memory or an archival function of mostly religious and later on secular institutions. Just as we cannot speak of users during those periods, in fact, neither can we speak of libraries as such. Only when the notion of education was introduced did the organization of library items start to turn into an informative system. Walking through an open shelf library or using a subject catalogue allowed the user to be surprised even when, admittedly, in a rather limited fashion. The crucial point, however, is that libraries started to reflect on how to make the immense amount of items or, in other words, data informative. Developments in ICT seemed to be the solution to all the problems (Thompson 1982). The paradoxical form of memory/information seemed to be, after all, only a technological problem soon to be solved by implementing state-of-the-art library information systems (LIS). As already mentioned further above, technology only deals with the problem of managing the immense growth of data but not the problem of information. “What is processed in information systems is only data (or potential information)” (Hjorland 2000:32). The problem of information is how to actualize information out of these immense pools of data (Wildavsky 1983).

At this point, we may return to the notion of memory in order to further specify the term. Memory should not be seen as a storage of past events; it is rather an expression of recursivity (Esposito 2002). It is the capability to observe sameness in difference, to construct repetition and hence what is already known. Information only occurs in comparison to what is remembered as well as forgotten. Memory is not a stock but rather the organization of observing information. Hence, we can formulate the memory/information paradox; memory is necessary to recognize information and information is necessary for the construction of repetition. Without memory everything and nothing is new and surprising. From a temporal perspective, we end up with the chicken and the egg: what was there first, memory or information - the observation of repetition or novelty? In order for an event to be observed as novel, it needs to be distinguished from what is repeated. From a factual perspective, we come to a tautology; in order to describe memory we refer to information and vice versa. By definition, a paradox does not offer a solution but rather the oscillation between the two sides of the form. In our case, the form is the library. It is the difference between memory and information. Consequently, librarianship is the self-observation of the form – the observation of the memory/information paradox based on the paradox – that oscillates between the two sides. In other words, memory is differentiated from information by applying the memory/information difference.

Similar to the argumentation on the potential information further above, books and catalogues do not remember but rather provide a potentiality for repetition and hence remembering as well. Memory and information are two sides of the same coin. Coming back to the difference between works and items, the organization of the fuzzy, abstract level of works by defining different items as being the same, is what makes a library an information organization. It constructs, like in every classification system, categories based on abstract definitions (works or keywords) to which physical items are assigned to. Sameness in difference and difference in sameness is just another way of phrasing the memory/information paradox. However, the self-referential closure of librarianship into an autonomous organizational form due to the mass production of books and the re-orientation towards

---

3 It is hard to imagine to read, for instance, a newspaper with the front-page title “nothing new happened today” or to watch a TV channel broadcasting the same shows, series, and movies every day.
being used is the key aspect that differentiates a library as an information organization in comparison to archives or museums which, so far, have remained mostly on the memory side of the paradox. In short, an information organization organizes the memory/information paradox in order for it to be a difference that makes a difference for an anonymous audience.

5 LIBRARIES AS INFORMATION ORGANIZATIONS – AN OUTLOOK

In this paper, libraries were introduced along the difference of memory and information in relation to developments of communication technologies. The emergence as an autonomous organizational form was outlined as a process of increased differentiation from its relevant environment that led to a shift from a focus on memory to a focus on information and retrieval – a form termed information organization. However, building an argumentation of libraries being information organizations based on the observational form of memory/information needs further development in order to being applicable to contemporary developments in librarianship. So far, the focus was on how libraries do what they do which is not a distinguishing characteristic anymore. Because of the converging, interoperable nature of binary-based media, libraries, archives, and museums become again less and less distinguishable when offering their services online. Additionally, other non-librarian organizations also provide information services online. Google, especially after launching their massive book digitalization project, seems to fulfil an equivalent function to what libraries were doing in the 20th century – connecting people with information. However, if we change the question of how into why a library does what it does, the difference becomes quite obvious. Google, like every commercial company, is a business oriented enterprise; consequently the reason why Google does what it does is simply profit. It seems a very trivial point to make, but rarely considered by harbingers of the end of libraries such as Weinberger (2007). The private sector has different goals than the public sector and fulfilling the role of an archival trustee in terms of provenance is not one of them (Jeanneney 2005). The Amazoogles of the Internet organize information to make money, but they are not information organizations in the sense of this article. In other words, digitalizing and indexing books is not what makes a (digital) library. What is it then?

Another point to be considered is that basically every organization can be observed according to the memory/information difference. Every organization “uses” information to make decisions, hence a relationship to memory is easily established (Wildavsky 1983, March 1991). After all, organizations do store and retrieve documents via an information system. Still, there is a difference between having a library and being a library – between using an information system and being an information system. Libraries are solely based on the dynamic relationship of memory and information. They establish a self-referential notion of the memory/information paradox leading to further differences (such as work and item) which in turn distinguishes librarianship from an IT centred organization of information-as-things (Buckland 1991). In order for libraries to remain libraries a re-unfolding of the memory/information paradox seems to be required which can be subsumed under the rather broad concept of digital libraries which, in fact, is not distinguishable from digital archives and museums anymore. The term information organization is expected to help to appreciate the peculiarities of these public information service providers in the next steps of the research. It directs the analysis towards contemporary changes without leaving historical developments behind.

Finally, the librarian memory/information distinction highlights different aspects than the predominant knowledge/information and data/information distinction in the IS field. The conceptualization of memory/information as a paradox expounds the limitations of IT – i.e. being capable to only solve problems - in a new light. The ways libraries re-unfold this paradox due to changes in societal communication and technological developments (mostly accomplished outside the domain of librarianship) is an excellent example for dealing with the increasing and boundary-crossing data-pollution and noise, basically every organization has to face.
## References


---

Proceedings ECIS 2009
http://www.oclc.org/dewey/resources/summaries/deweysummaries.pdf
A TRANSACTION COST THEORETICAL ANALYSIS OF SOFTWARE-AS-A-SERVICE (SAAS)-BASED SOURCING IN SMBS AND ENTERPRISES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0003.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Adoption, Small and medium sized enterprises (SMEs), Structural Equation Modeling, Transaction Costs Theory</td>
</tr>
</tbody>
</table>
A TRANSACTION COST THEORETICAL ANALYSIS OF SOFTWARE-AS-A-SERVICE (SAAS)-BASED SOURCING IN SMBS AND ENTERPRISES

Benlian, Alexander, Ludwig-Maximilians-University of Munich, Institute for Information Systems and New Media, Ludwigstr. 28, 80539, Munich, Germany, benlian@bwl.lmu.de

Abstract

This study seeks to better understand the factors contributing to the adoption of Software as a Service (SaaS) as on-demand sourcing option. Grounded in transaction cost theory (TCT), we developed a research model for assessing SaaS-sourcing at the application level. Four hypotheses with three TCT-based constructs (application specificity, environmental uncertainty, and usage frequency) and one contingency factor (firm size) affecting the adoption of SaaS-based sourcing were formulated. Survey data of 154 firms in Europe with 284 response items across different industries were collected to test the theoretical model. PLS-based structural equation modeling demonstrated that uncertainty emerges as the strongest factor for SaaS-adoption, while application specificity also contributes significantly. Differentiating between small and medium-sized businesses (SMBs) and enterprises, uncertainty represents the strongest driver for SaaS-adoption in SMBs, whereas application specificity is more important in enterprises. Finally, firm size is significantly and negatively associated with SaaS-based sourcing within, but not across the subgroups of SMBs and enterprises.

Keywords: Software as a service (SaaS), IT outsourcing, IT adoption, SMBs and enterprises

1 INTRODUCTION

On-demand software application delivery models are known since the late 1990s and have come in many forms and varieties such as Application Service Providing (ASP) or Business Service Providing (BSP). The common denominator of all these concepts is that this kind of demand-driven application sourcing model provides users and firms with network-based access to resources and expertise as well as to an integrated portfolio of complex applications that spans the complete virtual value chain of an enterprise (Kern et al., 2002a). While discussions on ASP-based outsourcing have become rather silent in recent IS management literature mainly due to the missing breakthrough success stories, a new acronym, called Software-as-a-Service (SaaS), has gained ground in the attention span of IS executives and is said to become a fundamental pillar of revenues for IT vendors in the next couple of years (Pettey, 2006). Although often discussed as an “old wine in new skins” concept, SaaS is promised to breach last barriers of adoption, as it offers more mature and comprehensive service packages on the customer side and the ability to support service-oriented multitenancy and a shared IT infrastructure to reap significant economies of scale on the provider side (Valente & Mitra, 2007).

Traditionally, the interest in on-demand based sourcing models was driven by a focus on core competencies, the attractive cost model for customers (i.e., turning fixed costs into variable costs), the benefits from cutting-edge technology and faster implementation times, transparent and predictable subscription pricing, and the possibility to flexibly balance the shortage of necessary IT skills (Kern et al., 2002b). Countering these benefits are the acknowledged risks of reliability (how can IT departments ensure that the business can access its applications?), security (how can it guarantee data privacy in line with regulations?), and process dependence (how can it make sure that quality of service is really achieved?) when outsourcing IT applications to a 3rd party (Kern et al., 2002a). Puzzled by these mixed point-of-views, researchers and practitioners have struggled to determine...
which side of the discussion will prevail over time. However, while many research and management papers have theoretically explored the meaning, characteristics, and economic implications of on-demand application outsourcing, only few have examined the drivers of SaaS-similar sourcing decisions on a substantial empirical basis, which may also be due to the immaturity of this research object. Specifically, what is missing in the existing literature is a solid and theory-based research model on the drivers of SaaS-based outsourcing which is assessed based on a broad dataset instead of a few isolated cases. Our study seeks to reduce this research gap. Key research goal of our study is to contribute to a heightened understanding in the drivers behind using SaaS-based sourcing models viewed through a transaction cost based lens. Concrete research questions are: (1) How do transaction cost based factors such as asset specificity, environmental uncertainty, and usage frequency explain the sourcing of applications via SaaS? (2) What is the role of firm size in the adoption of SaaS-based sourcing models? (3) How does SaaS-adoption look like on an application level today and in the future? Based on these questions, our research focus is not to compare SaaS-based sourcing with other types of outsourcing arrangements (e.g., licensing), but to explain why companies would agree or disagree to have their applications operated as service by an IT provider.

The following section reviews the literature from which our transaction cost theoretical framework is developed. A conceptual model and a series of research hypotheses are then formulated, followed by our research methodology, analysis, and the empirical results. The paper concludes with a discussion of findings and limitations.

2 TRANSACTION COST THEORY AND IT OUTSOURCING

Transaction cost theory (TCT) is embedded in the framework of the new institutional economics which was first introduced by Coase (Coase, 1937), who analyzed why firms exist and what firms do. One of Coase’s initial propositions was that firms and markets are alternative governance structures differing in their transaction costs (Coase, 1937, p. 389) – where by transaction cost Arrow referred to the “costs of running the system” (Arrow, 1969, p. 48). Accordingly, the transaction-cost approach is based on the premise that the existence of different organizational forms, whether they are markets, bureaucracies, or clans, is primarily determined by how efficiently each form can mediate exchange transactions between parties (Ouchi, 1980, p. 130). Over the past four decades, Williamson has added considerable precision to Coase’s general argument by identifying the types of exchanges that are more appropriately conducted within firm boundaries than via the market. To evaluate different governance or exchange arrangements, Williamson introduced characteristics or determinants of transaction cost differences of which asset specificity, frequency, and uncertainty are the most important ones (Williamson, 1985). In the IS research literature, transaction cost theory has particularly and comprehensively been applied in the context of IT outsourcing (e.g., Ang & Straub, 1998; Lacity & Willcocks, 1995). Wiener found in his literature synthesis on IT outsourcing that transaction cost theory has been the predominant theoretical approach to understand and explain IT outsourcing decisions (Wiener, 2006). Aubert et al. for example has investigated a transaction cost model on IT outsourcing using asset specificity, uncertainty, business and technical skills as determinants of the level of IT outsourcing (Aubert et al., 2004). Dibbern instead focused on the effect of human asset specificity on the sourcing of application services (Dibbern et al., 2005), while Loh developed an integrated governance model of IT outsourcing including various theoretical approaches (e.g., transaction cost theory and agency cost theory) to test his hypotheses (Loh, 1994). All of these studies provided consistent empirical support for the TCT framework.

Reviewing the IT outsourcing literature more closely focusing on on-demand based IT outsourcing, it is interesting to note though that drivers of on-demand outsourcing decisions have not been studied extensively so far. On the one hand, a number of theoretical research studies exist that propose definitions and classifications, and explore special characteristics (such as service quality or capabilities) of ASP-based sourcing models (e.g., Ma et al., 2005; Smith & Kumar, 2004). On the other hand, existing empirical studies very frequently limit their analysis to providing content-rich, but
less representative case studies on determinants of on-demand based IT outsourcing (e.g., Jayatilaka et al., 2003; Kern et al., 2002a). The number of research studies on the drivers of on-demand IT sourcing which are based on a solid empirical basis is thus rather small and limited to special aspects in the adoption in single countries and industries (Daylami et al., 2005). Based on these observations, the current IS literature is rather scarce on the assessment of more broadly valid and applicable adoption drivers of on-demand sourcing decisions. Drawing upon this empirical evidence, we seek to fill this research gap by examining TCT-based factors that may affect SaaS-based sourcing decisions.

3 HYPOTHESES DEVELOPMENT

Outsourcing of software applications is the handover of the development and maintenance of an application to an IT provider and is an alternative to internal development and maintenance. The essence of the transaction-cost based argument is that using an external IT vendor is not frictionless. When buying or leasing a service (such as SaaS) or a product, one incurs costs. If these costs become too high, relying on self-production is more appropriate. Transaction costs theory has two underlying assumptions. Bounded rationality is the inability of the human mind to find or process all the information about a transaction; therefore, it is conducted with a certain level of uncertainty. Opportunism is more than the simple defense of one’s interest or value maximization; it is self interest seeking with guile (Williamson, 1989). The combination of these assumptions results in information asymmetry. In order to strike a better deal, sellers will hide negative characteristics of their products, and buyers will not reveal how much they are willing to pay. Since both parties know that the other is opportunistic, each will engage in information seeking activities (e.g., having a product tested before buying it or asking for warranties). All these actions generate transaction costs. This is especially true for SaaS-based outsourcing in which an outsourcer’s handover of application data and process ownership may cause lower or higher levels of transaction costs compared to operating the application in-house. According to transaction cost theory, the factors determining the extent of transaction costs are grouped into three broad categories: (1) the specificity of the assets required for performing the transaction, (2) the uncertainty surrounding the transaction, and (3) the frequency of the transaction executed. As firm size seems to be a special factor in SaaS-based sourcing decisions due to the reasons mentioned in the introduction, it will also be included in our hypothesis development (4).

3.1 Application Specificity and SaaS-based Sourcing

Traditional transaction cost economics posit that transactions with high asset specificity (i.e., the degree to which assets can be redeployed elsewhere without losing value) are managed less expensively in-house, while the rest should be more efficiently outsourced (Williamson, 1989). Specifically, the theory suggests that when asset specificity is high, the assets involved in the transaction are so idiosyncratic and customized to the application context, that in-house production will bear lower transaction costs than outsourcing due to lower levels of so-called post-investment opportunistic behaviour (Klein et al., 1978). A number of prior research studies on IT outsourcing have employed asset specificity in their research models to explain outsourcing decisions with rather puzzling outcomes. While Nam et al. did not find any significant link between specificity and outsourcing (Nam et al., 1996), Aubert et al. discovered inconsistent effects of specificity (Aubert et al., 2004). On the other end, Dibbern showed that insourcing is more cost efficient and advantageous in creating strategic benefits through IS, if the provision of application services requires a high amount of firm specific human assets (Dibbern et al., 2005). In the case of SaaS-based application sourcing, application specificity is reflected in the degree that applications are customized to the individual requirements of the outsourcing company and thus comprise highly specific investments. Based on TCT thinking, it can be argued that the higher the degree of application specificity, the lower the level
of SaaS-based outsourcing, as integration and coordination costs for running a highly customized application system on the IT provider side outweigh the transaction costs for running the application in-house. These theoretical assertions lead to the following hypothesis:

**H1:** Application specificity is negatively associated with SaaS-based outsourcing.

### 3.2 Environmental Uncertainty and SaaS-based Sourcing

Analogous to asset specificity, environmental uncertainty surrounding a specific outsourcing relationship is posited to be negatively associated with the degree of outsourcing (Williamson, 1989). By the majority, prior research on IT outsourcing has supported this theoretical proposition. For instance, Nam et al. found that uncertainty towards potential opportunistic behaviour of the IT provider played a significant role in the decision to outsource. As expected, higher uncertainty and thus higher perceived risk lead to less outsourcing (Nam et al., 1996). Similarly, Aubert et al. found in their study that the level of uncertainty is the major deterrent to outsource IT operation activities (Aubert et al., 2004). Dibbern conceptualized environmental uncertainty in the IT outsourcing context as comprising business driven and technology driven uncertainty (Dibbern, 2004, p. 53-54). While business driven uncertainty refers to the extent to which the development of business related issues (such as pricing or processes) may be changed over time in the course of the outsourcing relationship by the IT vendor, technology-driven uncertainty captures the extent to which the required technical functions or features of the outsourced application may be changed over time. Based on these notions of uncertainty, one can hypothesize for SaaS-based outsourcing that when environmental uncertainty is high due to potential opportunistic behaviour of the IT provider regarding business- or technology-related activities, the outsourcer will prefer internal governance for highly risky activities. This leads us to the following hypothesis:

**H2:** Environmental uncertainty is negatively associated with SaaS-based outsourcing.

### 3.3 Usage Frequency and SaaS-based Sourcing

Theoretical ideas about transaction or usage frequency stem from the classical notion that an increasing number of transactions between two parties entails trust-building and routine which can be represented in a hierarchical governance form more efficiently than in a market governance form (Williamson, 1984). This is also backed up by classical organizational theory which argues that the higher the interdependence between organizational units, the higher the necessity to integrate both organizational units due to transaction cost savings (e.g., Thompson, 1967), which in turn can be attributed to routinization effects and economies of scale. Low frequency transactions with less complexity involved are thus likely to be organized through market interactions. In the context of IT outsourcing, Aubert et al. found that frequently used assets or skills in software development were used in-house, whereas infrequently and specialized assets and skills were contracted out (Aubert et al., 2004). Similarly, Hancox and Hackney synthesize from their findings in a study of private and public sector organizations that usage frequency indeed is negatively associated with the use of the market (Hancox & Hackney, 2000). Regarding SaaS-based sourcing, it can be argued that increased usage of an application system that requires different technical and human interfaces, skills and resources, as well as a high number of interactions (e.g., planning, adapting, and monitoring of task completion) will translate into increased coordination complexity between these different entities (Dibbern, 2004, p. 47-48). This will in turn induce the outsourcing decision-maker to rather opt for in-than for outsourcing due to transaction cost advantages. Hence, we can formulate the following hypothesis:

**H3:** Usage frequency is negatively associated with SaaS-based outsourcing.
3.4 **Firm Size and SaaS-based Sourcing**

Firm size is one of the most commonly studied factors in the IS and organization literature (e.g., Gremillion, 1984). Yet, different opinions exist as to the role that firm size plays in the process of adopting new technologies and innovations. On the one hand, large firms often possess more slack resources that can increase the openness for technology experiments – the so-called Schumpeterian resource advantages (Schumpeter, 1950). On the other hand, large firms tend to be less agile than small firms. This comparatively bigger structural inertia associated with large firms may entail more effort and cost for technology adoption (Duncan, 1976). In the context of IT outsourcing, Kern et al. developed the proposition – which is informed by transaction cost and resource based theory – that small and medium-sized companies (SMBs) are especially interested in more flexible application service providing mechanisms. This is mainly due to the fact that they can get access to strategic resources which are often prohibitively costly yet essential for SMBs to remain resource competitive against larger firms (Kern et al., 2002a, p. 171). Based on this theoretical assertion and the hypothesis that large firms show higher structural inertia, we argue that firm size will be negatively related to SaaS-based sourcing and thus deduce the following hypothesis:

**H4:** Firm size is negatively associated with SaaS-based outsourcing.

### 4 RESEARCH METHODOLOGY

#### 4.1 Data and Sample

To test the hypotheses suggested above, we designed a questionnaire and conducted a Europe-wide survey. The survey questionnaire was designed on the basis of a comprehensive literature review in the IT outsourcing context and based on interviews with IT managers. After several rounds of pretests and revisions, the survey was distributed in the biggest 6 European countries (Germany, France, UK, France, Italy and Spain) during the period of September through December 2007. The survey was conducted at the application type level, so that one company had the possibility to rate different application systems (i.e., Collaborative, Content, ERM, Human Capital, SCM, Production, Engineering, CRM applications) on the applicability and drivers of SaaS-based outsourcing.

The questionnaires were sent to 1,200 top or senior IT executives randomly selected from a directory of 3,000 firms in the finance, manufacturing, logistics, high-tech, energy, and TIME industry supplied by a market vendor (Bisnode Business Information Group). After we dropped 18 responses due to missing data and inconsistencies, a total of 284 usable responses from 154 different companies could be used for the analysis. We tested non-response bias with chi-square tests and found no statistically significant differences between early and late respondents on firm size and SaaS outsourcing levels (Pearl & Fairley, 1985).

#### 4.2 Operationalization of Constructs

Table 1 shows the operationalization of variables of our conceptual model. Besides one question, for which a percentage estimate on the SaaS adoption rate in 2007 and 2010 was requested, all other questions were asked from a scale ranging from 1 to 5, where 1 refers to the lowest score and 5 the highest score in the item scale. Measurement items of all variables were drawn from those of previous studies as depicted in Table 1.
4.3 Instrument Validation

Content validity was established through the adoption of constructs that had been used in former studies (see Table 1) and through pilot tests with IS practitioners of different industries. The reflective measurement models were validated using the standard procedures of current literature (Chin, 1998; Straub, 1989). Items of scales in a related domain were pooled and factor-analyzed to assess their convergent and discriminant validity. While convergent validity was determined both at the individual indicator level and at the specified construct level, discriminant validity – the extent to which different constructs diverge from one another and thus are distinct in nature – was assessed by analyzing the average variance extracted and inter-construct correlations. The results are illustrated in Table 2.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Indicators</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>SaaS-based Sourcing</td>
<td>Overall level of SaaS-based outsourcing with respect to the application</td>
<td>Loh, 1994</td>
</tr>
<tr>
<td></td>
<td>Approximate/Estimated percentage of the application’s budget allocated to SaaS-based outsourcing in 2007 and 2010</td>
<td>Dibbern, 2004; Teng et al., 1995</td>
</tr>
<tr>
<td>Application Specificity</td>
<td>The extent to which the application is related to the individual needs of the company</td>
<td>Kern et al., 2002a</td>
</tr>
<tr>
<td></td>
<td>The extent to which the application is customized to individual requirements of the company</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The extent to which the application supports special processes of the company</td>
<td></td>
</tr>
<tr>
<td>Environmental Uncertainty</td>
<td>The extent to which the application outsourcing is subject to technical difficulties in terms of bandwidth and reliability</td>
<td>Smith &amp; Kumar, 2004; Ma et al., 2005</td>
</tr>
<tr>
<td></td>
<td>The extent to which the application outsourcing is subject to economic dependencies in terms of pricing model changes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The extent to which the application outsourcing is subject to process dependencies in terms of quality of service (in SLAs)</td>
<td></td>
</tr>
<tr>
<td>Usage Frequency</td>
<td>The extent to which the application requires interfaces to other systems in and outside the company</td>
<td>Smith &amp; Kumar, 2004</td>
</tr>
<tr>
<td></td>
<td>The extent to which the application is used frequently by different internal and external users</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The extent to which the application usage requires coordination efforts (i.e., skills and resources) among users</td>
<td></td>
</tr>
<tr>
<td>Firm size</td>
<td>Number of employees of company</td>
<td>e.g., Thompson, 1967</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Number of indicators</th>
<th>Range of Standardized Factor Loadings*</th>
<th>Composite Reliability ($\rho_c$)</th>
<th>Average variance extracted (AVE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SaaS-based Sourcing</td>
<td>2</td>
<td>0.920 – 0.941</td>
<td>0.928</td>
<td>0.865</td>
</tr>
<tr>
<td>Application specificity</td>
<td>3</td>
<td>0.797 – 0.860</td>
<td>0.874</td>
<td>0.700</td>
</tr>
<tr>
<td>Environmental uncertainty</td>
<td>3</td>
<td>0.899 – 0.957</td>
<td>0.944</td>
<td>0.849</td>
</tr>
<tr>
<td>Usage frequency</td>
<td>3</td>
<td>0.707 – 0.883</td>
<td>0.834</td>
<td>0.628</td>
</tr>
<tr>
<td>Firm size</td>
<td>1</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

* All factor loadings are significant at least at the p<0.05 level

Table 1. Measurement of Constructs

Table 2. Assessment of Measurement Model: Factor Loadings and Reliability
All standardized factor loadings are significant, thus suggesting convergent validity (Bagozzi et al., 1991). To evaluate construct reliability, we calculated composite reliability for each construct. All constructs have a composite reliability significantly above the cutoff value of 0.70 (Hair et al., 1998). All reflective constructs also met the threshold value for the average variance extracted (AVE > 0.50). For discriminant validity of latent variables, the square roots of AVEs exceeded the inter-construct correlations that were negligibly low between the independent constructs.

5 EMPIRICAL ANALYSIS

5.1 Descriptive Results on SaaS-based Outsourcing on the Application Level

We analyzed the adoption levels in terms of the percentage of an application’s budget allocated to SaaS-based sourcing in 2007 and 2010. SMBs allotted around 5.7% of their IT application budget to SaaS-based outsourcing in 2007 overall and plan to spend as much as 14.8% in 2010. On the contrary, enterprises allocated only around 3.6% of their IT application budget to SaaS-based sourcing models. However, according to the results of our study, the adoption level of SaaS in enterprises will triple until 2010 to a level of around 11.1%. To gain a deeper understanding of current and future SaaS adoption, Table 3 shows the outsourcing levels of different SaaS-based applications in 2007 and 2010.

<table>
<thead>
<tr>
<th>Application type</th>
<th>SMBs</th>
<th>Large Enterprises</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adoption level 2007</td>
<td>Estimated Adoption level 2010</td>
</tr>
<tr>
<td>Collaborative</td>
<td>~15%</td>
<td>~30%</td>
</tr>
<tr>
<td>Content Mgmt.</td>
<td>~2%</td>
<td>~7%</td>
</tr>
<tr>
<td>ERM</td>
<td>~2%</td>
<td>~4%</td>
</tr>
<tr>
<td>Human Capital</td>
<td>~12%</td>
<td>~33%</td>
</tr>
<tr>
<td>SCM</td>
<td>~1%</td>
<td>~1%</td>
</tr>
<tr>
<td>Production</td>
<td>~0%</td>
<td>~0%</td>
</tr>
<tr>
<td>Engineering</td>
<td>~1%</td>
<td>~2%</td>
</tr>
<tr>
<td>CRM</td>
<td>~11%</td>
<td>~41%</td>
</tr>
<tr>
<td>Total average (n=284)</td>
<td>~6%</td>
<td>~15%</td>
</tr>
</tbody>
</table>

Table 3. SaaS-Adoption on the Application Level

Comparing the absolute adoption levels between SMBs and enterprises in 2007 and 2010 along all applications, a significant difference could be identified using t-tests (t-value_{2007}=3.49, p<0.001; t-value_{2010}=2.22, p<0.05). The implications of these descriptive results will be dealt with in the discussion section along with the following explanatory outcomes.

5.2 Analysis of the overall sample

We tested our research hypotheses with actual 2007 adoption rates by using PLS-based structural equation modelling (Chin, 1998; Lohmöller, 1989) based on SmartPLS (Ringle et al., 2005). In contrast to parameter-oriented and covariance-based structural equation modeling, the component-based PLS method is prediction oriented (Chin, 1998, p. 352). It seeks to predict the variations in the dependent variables of the model, which we want to achieve for the SaaS-based outsourcing construct in our study. Since PLS does not account for the covariances of all indicators, but only for those variances that have been specified in the model, it is closer to the actual data than the covariance-based procedure (Fornell, 1989). Due to the partial estimation of single elements in the causal model, fewer empirical cases are needed in PLS-than in covariance-based structural equation modeling to generate consistent and reliable results. To provide an aggregate view on the assessment of PLS-based models, the structural model is evaluated by looking at the percentage of the variance explained (R²) of all...
dependent latent variables. By examining the size and stability of the coefficients associated to the paths between latent variables, hypotheses, which were proposed during the model specification process, are finally analyzed for their significance. Standardized path coefficients and $R^2$ as major model-fit index are shown in Figure 1.

![Figure 1. Results of the Overall Model](image1)

![Figure 2. Comparison of Path Models SMBs vs. Enterprises](image2)
5.3 Comparing SMBs and Enterprises on the drivers of SaaS-based sourcing

Based on sub-samples of 169 SMBs and 115 enterprises, structural equation models were calculated. Analogous to the assessment of the full sample, standardized path coefficients and the share of explained variance ($R^2$) were analyzed and compared (see Figure 2). In the SMB sample, application specificity and environmental uncertainty have both strong negative and significant paths leading to SaaS-based outsourcing. The paths from usage frequency and firm size to SaaS-adoptions are also both negative and significant. A total of around 92% of the variance of SaaS adoption could be explained by TCT attributes and firm size. In the enterprise sample, 76% of total variance could be explained by the independent constructs in our research model. Paths from application specificity and firm size to SaaS-based outsourcing are both negative and highly significant. While environmental uncertainty is also negatively and significantly correlated with the dependent variable, the path from usage frequency is insignificant.

6 DISCUSSION

6.1 Major Findings and Interpretations

Finding 1: Environmental uncertainty emerges as the strongest factor for SaaS-based outsourcing, while application specificity also significantly contributes to SaaS-based application sourcing.

Linking the results of our empirical study to prior research, environmental uncertainty turns out to be the main deterrent of outsourcing on-demand IT applications as was also found out by Aubert et al. (2004). Most apparently, business and technical uncertainties still hinder IT executives and managers to take advantage of on-demand outsourcing options. Although asset specificity showed mixed results in its impact on the level of IT outsourcing in previous research, we found strong evidence that it can explain the level of IT outsourcing in the case of SaaS. This is also illustrated by our descriptive results on SaaS-based sourcing on the application level. While application systems, which supposedly have a lower degree of specificity in the sense that they are easier to transfer from one to another application context (such as CRM applications), show higher degrees of adoption, highly specific application systems (such as Operations & Manufacturing or Engineering applications) indicate very low adoption levels. Usage frequency as factor for SaaS-based sourcing showed inconsistent results. Obviously, the degree of interaction complexity and coordination intensity plays a more important role in the outsourcing decision for SMBs than for enterprises. This may be due to the fact that such organizational efforts attached to an outsourcing relationship engender a comparatively higher resource burden for SMBs than for enterprises. Overall, companies seem to be eager to acquire SaaS for applications that are not mission critical, involve relatively low data security and privacy concerns, and require little integration with on-premise applications. However, they still seem to be sceptical in application areas that involve transactions that are tied to mission critical processes including production and ERP and transactions between buyers and suppliers including ERP and logistics. Software vendors that shift from a traditional licensed model to SaaS will have to track the receptiveness of customers to the service delivery model in order to understand which applications will migrate when and position one’s own offerings early in the uptake.

Finding 2: While application specificity plays the strongest role in the adoption of SaaS-based applications in large enterprises, environmental uncertainty is the strongest factor for SaaS-based outsourcing in small- to mid-sized businesses.

As illustrated in Figure 3 with the comparison between the path models of SMBs and enterprises, the weights of the paths from application specificity and environmental uncertainty to SaaS-based
outsourcing are reversed in their magnitude in dependence of organizational size. Here it can be argued that uncertainty is a stronger driver or barrier in the SMB decision to outsource compared to application specificity and usage frequency. Before a SMB is ready to make a positive outsourcing decision, it wants to feel ensured that future changes in prices, processes, and technical issues are safeguarded through specific (e.g., contractual, reputational) mechanisms so that it is not worse off than before. This reasoning is in line with prior studies where uncertainty and risk factors also played a major role in explaining IT outsourcing (e.g., Aubert et al., 2004; Clemons & Row, 1992). By contrast, large enterprises are rather hindered to outsource by the idiosyncrasies and intricacies of their IT landscape. Large firms tend to have more fragmented IT legacy systems built over a long period of time, which is often further complicated by complex business processes. For that reason, application specificity may play even a more important decision factor in enterprises than in SMBs when it comes to carving out IT applications in the wake of a positive IT outsourcing decision. These results have several important implications for management. On the one side, IT managers of SMBs evaluating the adoption of SaaS-based sourcing models should concentrate on clarifying and safeguarding for potential sources of business and technology uncertainty. Integrating these aspects into contractual modalities (e.g., into the price model or SLA specifications) could be an important step for the success of the entire outsourcing deal. On the other hand, IT managers of larger enterprises should rather focus on addressing the ease or difficulties of integrating application system services delivered by SaaS-application providers into processes and the IT landscape of the outsourcer. An in-depth assessment of the application portfolio for its level of application specificity would therefore be a valuable preparatory step for actual outsourcing evaluations and negotiations. On the other side, SaaS-application providers should develop and establish authentic trust-building mechanisms to reduce outsourcers’ perceived uncertainty. Moreover, they should build up necessary skills and capabilities to help integrate their SaaS-based offerings into the outsourcers’ process and IT landscape.

**Finding 3**: Firm size is significantly and negatively associated with SaaS-based sourcing within, but not across the subgroups of SMBs and large enterprises.

In terms of firm size as a factor influencing the level of SaaS-based outsourcing, we could find that the organizational size had no significant impact in the overall sample. However, when analyzing the sub-samples for SMBs and larger enterprises individually, we could observe a consistent significant effect of firm size on SaaS-based sourcing. This means that within the SMB and enterprise sub-samples respectively, firms with a lower number of employees had a comparatively higher SaaS-adoption rate than firms with a higher number of employees. These results are not entirely in sync with the proposition made by Kern et al. (2002a) that smaller and medium-sized firms are generally more prone to adopt on-demand outsourcing options. Our research findings rather point to a more sophisticated picture which shows that effects emanating from the sub-group level cancel out each other on the full sample level. This phenomenon may suggest the existence of moderator variables which intervene in the relationship between organizational size and SaaS-based outsourcing. Our empirical findings also indicate that the absolute adoption levels in 2007 and 2010 are comparatively lower for enterprises than for SMBs indicating that smaller and mid-sized firm tend to take advantage of flexible software sourcing models more strongly than corporations. However, according to our descriptive analysis, the gap between the SaaS adoption levels in SMBs and in enterprises will not widen but narrow in the future. These results add to the findings above that, against the predictions of many IT providers, SaaS-based sourcing is not only a valid option for SMBs, but also for larger corporations.

### 6.2 Limitations and Future Research

As with any research, this study does have some limitations. First, since our dataset is cross-sectional in nature, we can only show associations, not causality. Furthermore, assessing estimated adoption rates of SaaS-based sourcing models in 2010 is an insufficient proxy for studying longitudinal processes which could probably uncover time preferences in the adoption behavior of companies.
Second, although this study included 6 different countries and industries with consistent results, the low sample size in specific categories did not allow us to apply inferential statistics to explore the differences and commonalities between different countries, industries, and application types. Third, the investigation of SaaS-based outsourcing on the application level turned out to be a relevant and appropriate approach to gain concrete insights into different application markets. Focusing on a small set of special IT applications would foster an even more thorough understanding in selected areas of SaaS-based application sourcing. Accordingly, future research directions are to increase the sample size in the abovementioned categories and to extend our study database over time to pave the way for a longitudinal study. With regard to our theoretical framework and empirical findings, we can derive three further recommendations for future research. First, extending our transaction cost theoretical framework by other theories (e.g., incomplete contracts or property rights theory) would add further insights into the drivers and barriers of on-demand outsourcing. Second, deepening the notions of application specificity and environmental uncertainty as important drivers of SaaS-adoption would add more insights for researchers as well as practitioners. Last but not least, digging deeper into the relationship between firm size and SaaS-based sourcing would uncover more clearly what moderator variables intervene in the relationship between organizational size and SaaS-based sourcing.

References


CONTINGENT DYNAMICS OF IS ALIGNMENT IN SMES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0119.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Strategic Alignment, Small and medium sized enterprises (SMEs), Strategy, IT Strategy</td>
</tr>
</tbody>
</table>
CONTINGENT DYNAMICS OF IS ALIGNMENT IN SMES

Margi Levy, Warwick Business School, University of Warwick, Coventry CV4 7AL, UK, Margi.Levy@wbs.ac.uk

Philip Powell, School of Management, University of Bath, Bath BA2 7AY, UK and University of Groningen, Netherlands, mnspp@management.bath.ac.uk

Philip Yetton, Australian School of Business, University of New South Wales, Australia, phily@agsm.edu.au

Abstract

Performance is a function of the alignment between information systems (IS) strategy and other business domains, including business strategy and organization structure. However, research has focused on the outcomes of that process, rather than the processes by which that alignment is developed. Using multiple case data, this paper investigates alignment in SMEs, explaining why different SMEs follow different paths to alignment. Four paths are identified, with the path chosen contingent on an SME’s market position. The four paths vary from simple, when the focus of the IS investment is on improving control, to complex, when the intent is to reposition the SME. The implications of these findings for management and research are explored.

Keywords: Alignment, small firms, paths, fit, SMEs.
1 INTRODUCTION

Information systems (IS) can automate to reduce costs and improve operational efficiency or IS may be adopted to informate and add value. Most small and medium-sized enterprises (SMEs) begin by automating, with some then seeking differentiation and using IS to add value. To achieve the latter, SMEs need an IS strategy that is aligned to their business strategy. However, for many firms, and particularly for SMEs, IS strategy is reactive to business pressures and the potential advantages of using IS strategically are often not considered.

While there is general agreement about the benefits of alignment, the process of aligning IS strategy with business strategy is problematic for many organizations (Avison et al 2004). In part, this is because little is known about the change processes in large organizations that lead to alignment (Sabherwal and Chan 2001, Smaczny 2001) and even less is known about the equivalent change processes for SMEs. Instead, research has focused on the characteristics of alignment, rather than on the process of aligning. In contrast, this paper investigates the different paths to alignment followed by SMEs. The paper begins by integrating the MIT90s model and the Focus-Dominance model (Levy and Powell 1997). The analysis shows that the alignment paths followed by SMEs are context-contingent.

The data analysis reports both qualitative and quantitative findings. Case descriptions illustrate the different paths to alignment in each of the four quadrants of the Focus-Dominance model. Quantitative analysis shows that the frequency with which those different paths are followed is context-contingent. The paths vary from simple, limited to a relationship between technology and management processes, to complex, involving changes in all five components of the MIT90s model. The reasons for adopting the different paths are discussed. The paper concludes with the implications for management and research.

2 LITERATURE REVIEW

Alignment is the process of developing fit among the key internal activities within an organization and between that internal fit and the external context and it enables firms to be competitive (Porter, 1987; Ciborra, 1997). Alignment, also termed fit, integration, bridge, harmony, fusion and linkage, enables IT investment returns to be maximised. The importance of strategic alignment has been discussed for the last 20 years (Earl, 1996). Yet, it remains a key concern (Luftman et al, 1996) and is consistently ranked among the most important issues faced by IT executives (Papp 2001, Tallon and Kraemer, 2003). A number of models of strategic alignment have been proposed. The two that have attracted most attention from researchers are the MIT90s model (Scott Morton, 1991) and the Strategic Alignment Model (Henderson and Venkatraman 1989). The MIT90s model argues that a successful organization has a high fit among its strategy, structure, roles and skills, management processes and technology, and between that configuration and its business environment.

More specifically, alignment in the MIT90s model is achieving synergy between strategy, structure, management processes, technology, and roles and skills, in order to sustain the quality of ‘interdependence’ and thus to achieve competitive advantage (Avison et al 2004, Hsiao and Ormerod 1998). Initially, it was assumed, at least implicitly, that there is only one transition path from one state of alignment to another. This is the ‘classic’ route of a change in strategy, followed by changes to structure, to which management processes, roles and skills, and technology are then aligned (Scott Morton 1991).

There is limited research on the process of aligning IS strategy and other business domains, including business strategy and structure. To examine this process or path to alignment, consider the following two studies. One analyses the experience of Flower and Samios, a small architecture firm (Yetton et al. 1994). The other analyses four cases in large organisations (Hsiao and Ormond 1998). Both use the
MIT 90s model to define the different paths to alignment. Together, they identify five different paths. The different paths start with different elements of the MIT90s model and move through the other four elements via a variety of routes. For example, Hsiao and Ormerod (1998) find two planned paths emphasizing the centrality of strategy – structural reconfiguration and process engineering, while the other two emergent paths – human renewal and IT transformation - lead with individuals and technology respectively. Yetton et al (1994) demonstrate three paths through the alignment elements.

The presence of different paths suggests that the choice of a path is context contingent. Consistent with this conclusion, Levy et al. (2001) show that the choices among different IS applications by SMEs are a function of their location in the Focus-Dominance model. Inspecting the different IS applications, some are adopted to implement a new strategy, some could not be implemented within the existing functional structures and some would require different competences to leverage them effectively.

The potential existence of multiple paths to alignment motivated the general question guiding this research: Are paths to alignment for SMEs context-contingent? This research begins to address the general observation that the process of alignment needs to be better understood (Sabherwal and Chan 2001, Smaczny 2001). Indeed, Cragg et al. (2002) report that many SMEs ‘have developed a high degree of alignment but we don’t know how that was achieved.’

Successful SMEs plan their growth and the systems required to support that growth (Salmeron and Bueno 2006, Street and Meister 2004). This planning is particularly important for mature SMEs, with owner-manager interest and enthusiasm being the prime driver of both the IS investment and its adoption (Premkumar and Roberts 1994). Whatever the motivation, SMEs invest incrementally in IS due to resource constraints (Levy and Powell 2005). Importantly, this incremental nature of IS investment allows mapping of the paths SMEs follow as a sequence of steps through the MIT90s model to implement that investment. Framed in this way, a path defines the sequence in which components in the MIT90s model are re-aligned. In the extreme, a new desktop environment might be purchased without any other changes. Assuming an SME is in alignment before the change, it would be in alignment after that change. This raises the interesting possibility of sub-paths to alignment in the MIT90s model, with the complexity of paths varying from simple (one or two components) to complex (all components) being re-aligned. The literature on alignment has implicitly, if not explicitly, focused on complex re-alignment. This paper shows that this assumption is inappropriate for the analysis of SMEs.

The Focus-Dominance model, in which two dimensions - the strategic focus on cost or value-add, and the level of customer dominance - defines an SME’s strategic market position, creating four scenarios in Figure 1 for IS investment: Control, Co-ordination, Collaboration and Positioning (Brown and Lockett 2004, Levy et al 2001, 2002, 2003).

<table>
<thead>
<tr>
<th>Customer Dominance</th>
<th>Low</th>
<th>Co-ordination</th>
<th>Positioning</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td>Control</td>
<td>Collaboration</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cost</th>
<th>Value-Added</th>
</tr>
</thead>
</table>

*Figure 1. Focus-Dominance Model for SMEs (Adapted from Levy et al. 2001)*
Control: The focus of IS use in this quadrant is control. IS investment is treated as a cost of doing business, specifically, controlling the internal processes. These are operations-focused organisations in which IS has limited business value (Tallon et al 2000). Typically, their IS includes word-processing and accounting spreadsheets to control business expenses and revenues. Changes in structure are not required to achieve alignment because the new systems are implemented within the existing functional structure. In addition, the absence of any link to strategy is consistent with their internal operational focus (Tallon et al 2000). This behaviour is appropriate for SMEs, whose owners are not interested in growth and, instead, are specifically concerned with retaining control. The IS are simple, stand-alone and user-friendly. Thus, there is no requirement for changes in roles and skills, with the alignment process limited to changes in IS to support the new management processes (Quadrant 1 of Figure 2).

Co-ordination: IS investment is used to support communication and customer relationships. Owners are frequently interested in growth. They invest in IS to manage their customer base, when the number of customers exceeds the capacity of their manual systems. The new customer databases are typically supported by a local area network (LAN). Similar to the Control quadrant, the focus is still internal with no link to, or change in, the SME’s strategy. In addition, no changes to the individual roles and skills component are required. There are two reasons for this. One is that the systems adopted are not technically demanding (Levy et al. 2001). The other is that the IS knowledge of the owner, rather than the employees’ skills, is the driver of IS adoption (Premkumar and Roberts 1994). Thus, the dominant path is likely to be from management processes to technology to structure. To follow this path requires only limited restructuring to realise the benefits, with structural change supporting IS management (Bergeron et al 2001). This restricted path is presented in Quadrant 2 of Figure 2.

Figure 2. Focus-Dominance Model and Contingent Paths to Alignment
Collaboration: SMEs exchange real-time information with a small number of large customers. There is extensive use of EDI, with the SMEs often part of a customer’s Extranet (Levy et al. 2001). Customers are frequently the driving force behind the introduction of new IS. The proposed path to alignment is shown in Figure 2 with strategy driving technology, followed by management processes to support inter-organizational collaboration. This aligns an SME’s strategic information system, defined as the alignment between strategy, management processes and IS (Chatfield and Yetton 2000), with its customer’s strategic IS. The investment is managed within the existing structure and implemented without changes to roles and skills.

Positioning: SMEs integrate their IS and business strategies taking a dual internal and external perspective for the business (Tallon et al 2000). The owners are aware of IS-based opportunities for changes in business processes. The challenges here are the potential cost of the IS investment and the need for flexibility to support on-going change. Owners, who are both knowledgeable about IS and able to integrate it with the business strategy, are rare. These ‘dual focused’ SMEs focus simultaneously on operational efficiency and strategic positioning to improve performance through new market creation (Tallon et al 2000). Essentially, the path is strategy-led and enabled by new technologies (Croteau and Bergeron 2001). Bergeron et al. (2001) report a high correlation between strategic orientation and strategic IS. This relationship is supported by new management processes, to which roles and skills, and structure are aligned to position or re-position the SME. The suggested alignment path in this quadrant (Quadrant 4 of Figure 2) is more complex than in the other quadrants.

The above discussion examines paths to strategic alignment through the MIT’90s model, which are contingent on SMEs’ positions in the Focus-Dominance model. With different IS applications typically adopted in each quadrant of Figure 1 (Levy et al. 2001), the subject of this research is the different paths by which those applications are implemented and become embedded in SMEs. Formally, integrating the four paths identified in Figure 2:

Hypothesis 1: When investing in IS, the paths to alignment followed by SMEs are contingent on their strategic positioning in the Focus-Dominance model.

3 METHOD

This research reports both qualitative and quantitative analysis. The qualitative analysis investigates the paths to alignment followed by 27 SMEs. The quantitative analysis formally tests Hypothesis 1. As Lee (1991) argues, an integrated framework that includes interpretive and positivist approaches can provide robust understanding of phenomena. Lee proposes that observation of a situation can lead to an initial understanding that is then interpreted by the researcher based on their knowledge of the field. Finally, the research creates formal propositions that can be tested using positivist constructs. The outcomes from the propositions should then be reviewed in the empirical situation (Lee 1991).

Case studies are a powerful methodology for conducting the type of exploratory research proposed here. In particular, case research is effective when theory is relatively underdeveloped (Eisenhardt 1989). In general, this is true for research on the dynamics of alignment. Further, when the boundaries of the research are not clear, there is a need to investigate the issue within a real life context, collecting views from different sources (Yin 1994). However, while cases can be used to investigate a situation, generalisations based on one or two case studies are subject to a potential external validity threat. The multiple cases collected here ensure that common patterns are identified rather than risking generalisation from chance occurrences based on a single case (Eisenhardt 1989), with formal modelling to validate the patterns (Lee 1991).

Each case study was conducted over a one week period by a team of researchers and postgraduates, following the protocols developed by Levy et al. (2000, 2001, 2002). The organizations were not selected by industry, as there is little difference in SMEs’ strategic IS adoption across industries (Levy and Powell 2000). During the survey period, the SME owner, the senior management team and other employees took part in a series of standardized, semi-structured interviews, each lasting 1-2 hours.
The interviews cover the history of IS developments, business backgrounds, markets, and future plans for IS adoption. Pertinent issues are reported back to participants to allow them to correct misunderstandings, validate the findings and stimulate further discussion. The interview data identifying the choice of strategy, together with the number and importance of customers, classifies each SME’s strategic focus as cost reduction or value adding, and its customer dominance as high or low. Based on this data, each SME is classified within the Focus-Dominance model. The numbers of SMEs in each quadrant is shown in Table 1.

<table>
<thead>
<tr>
<th>Focus Dominance Quadrant</th>
<th>Number of SMEs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>11</td>
</tr>
<tr>
<td>Coordination</td>
<td>8</td>
</tr>
<tr>
<td>Collaboration</td>
<td>5</td>
</tr>
<tr>
<td>Positioning</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
</tr>
</tbody>
</table>

Table 1. Numbers of case studies in each quadrant of the Focus-Dominance Model

Having located each SME in the Focus-Dominance model, the next step identifies the alignment path. The changes the owners had put in place to implement each IS investment are used to map the paths taken. The interviews identify the nature of the IS, and changes in both the management processes, and IS-based training and recruitment. Changes in organizational structures are identified from organization charts, and the owners’ and senior managers’ descriptions of, and reasons for, structural change. Finally, a case description of each SME is written. Two of the authors inspected the sequence of steps defining the alignment paths followed by the SMEs. Based on that data, they independently coded the path followed by each SME. The two coders agreed on the classification of each SME’s path, demonstrating high inter-rater reliability.

The final step in the data analysis is what Lee (1991) calls third level understanding, where positivist, formal propositions are tested. To test Hypothesis 1, the frequency of agreement between the observed and expected paths to alignment for each SME is compared to the expected frequency under the null hypothesis that paths to alignment are not context contingent. If the four paths identified in Figure 2 are assumed to be both exhaustive (i.e. there are no other paths) and equally likely, then the critical value for the frequency of agreement, \( \theta \), would be \( \theta_c = 0.25 \). This would be a conservative test, given that there are many different potential paths through the MIT90s model. For example, there are 120 different potential five-step paths (Yetton et al. 1994).

In addition, the data in Table 1 suggests that the distribution of SMEs is not random across the four quadrants. Inspecting Table 1, the null hypothesis would be supported for 11/27 cases, when investing in IS, if the expected, non-contingent, path for all SMEs was specified as the path to alignment presented in the Control quadrant of Figure 4. So, \( \theta_c = 11/27 \), where that is the frequency of the most commonly adopted path, namely, management processes followed by technology, is adopted as the critical value to test Hypothesis 1 in the Results section below. This is a very conservative and demanding criterion – the probability that an SME adopts the path to alignment shown in Figure 2 for its location in the Focus-Dominance model must be greater than 40.7%.

4 RESULTS

The expected alignment path for each quadrant in the Focus-Dominance model is presented in Figure 2. The SMEs in the Control quadrant are characterised by a dominant path from management processes to technology adoption. This pattern accounts for 10 of the 11 cases here. The exception is Bird Designs, which needed to upgrade its staff skills to leverage its use of the new IS-based accounting systems. Therefore, its path included a link from management processes to roles and skills.

Chemical Analysis Co. illustrates the path to alignment followed in the Control quadrant. It followed the simple alignment path of management processes to technology. A small, family-run firm that
analyses soil samples for local authorities and utility industries, it was established by the owner in 1983, and grew steadily for 10 years. Growth then declined and it entered a steady state. The market is highly competitive. The owner invested in IS to reduce demands on senior management time, implementing a spreadsheet package for financial analysis, supported by a stand-alone accounting package. A specialist package was added to develop chemical analysis reports. The adoption of technology was reactive, focused on improving control over individual functions.

The dominant path in the Co-ordination quadrant in Figure 2 is management processes to technology, followed by changes in organization structure. This pattern explains seven of the eight cases in this quadrant. The exception is Regional Travel in which there is no link to structure: Its path is consistent with the dominant path in the Control quadrant in Figure 2. Regional Travel had not recognized the benefits from restructuring its informal approach to managing the organization.

Queensway Photographic Designers illustrates the path to alignment followed in the Co-ordination quadrant. It specializes in the development of large-scale photography for window and conference displays. The CEO owns the firm and is supported by a small management team. The firm works in a specialist niche, large-scale photographic printing, although it provides a wide range of photographic services. He has invested in technology to support that market niche, including a small administrative network, deploying accounting, word processing and order processing systems. For example, job costing was automated, using prior experience to quote prices. Job tracking was also implemented, with the system providing job progress information for customers. The owner is only planning sufficient growth to facilitate the sale of the business when he retires.

The complexity of or number of steps in the alignment path in the Collaboration quadrant is the same as for Co-ordination, but the path takes a different form. All cases here follow the dominant path, led by strategy, and followed by changes first in technology and then in management processes. Critical drivers include changes in major customers’ procurement, production and quality control systems.

Car Sign Design Co. (CSD) illustrates the path followed in the Collaboration quadrant. The owner established CSD in 1990 to produce graphic design liveries and signs, specializing in the automotive market. The firm employs 15 staff, with a plan to increase turnover 25% in 5 years. While only 10% of customers are in the automotive sector, they account for 80% of turnover. CSD works closely with major customers to ensure that orders are maintained in this highly competitive industry. The firm competes on quality and design innovation to develop and maintain the relationships. To do that, it has invested heavily in design and production technologies. CSD recognized early the value of the Internet to improve its competitive advantage by developing customer relationship management, particularly for the exchange of design information. New management processes were then developed to integrate order processing, accounts, design, production and delivery. CSD is now able to turn around design information quickly, with the new business processes supporting on-time order delivery.

Finally, the Positioning quadrant in Figure 2 identifies the need to align all five components of the MIT90s framework. In the three cases here, owner knowledge about opportunities from IS and the potential to exploit them is the key driver of strategic alignment. All three SMEs adopted the same path, beginning with their strategic objectives influencing their IS technology adoption, followed by changes in management processes and, finally, changes in roles and skills, and structure.

Mobile Phone Surveyors illustrates this path. The owner and CEO recognized early the potential growth in the mobile phone market. He positioned his firm to survey sites for the major mobile providers, developing a niche market that involves both negotiating with local authorities and recognising the public’s concerns over mobile phone masts. He and his team of surveyors have developed a specialised body of knowledge about the acceptability and viability of phone mast locations in the UK and overseas. His decision to adopt a knowledge management system to provide a repository for the surveyors’ knowledge was explicitly designed to support growth. Management processes were then implemented to co-ordinate and exploit that knowledge management resource. This was followed by changes to the organizational structure to support staff working as a team rather
than as individuals. At the same time, the owner recognized that he needed to train or recruit staff to use the new systems.

Aggregating across the four quadrants, 25 SMEs adopted the alignment path specified in Figure 2 for an SME in their strategic market position. Only two SMEs followed other paths. Hypothesis 1, *when investing in IS, the paths to alignment followed by SMEs are contingent on their strategic market positioning in the Focus-Dominance model*, is strongly supported. Table 2 reports that the probability of adopting the expected contingent paths, $\theta = 0.92$ or 92%, is significantly higher than expected under the null hypothesis that the alignment path is non-contingent, $\theta_c = 0.41$, ($\theta > \theta_c$, $p \leq 0.01$).

<table>
<thead>
<tr>
<th>Quadrant</th>
<th>Agreement with Contingent Path</th>
<th>Non-agreement with Contingent Path</th>
<th>Path Complexity (number of steps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positioning</td>
<td>3</td>
<td>0</td>
<td>4.0</td>
</tr>
<tr>
<td>Collaboration</td>
<td>5</td>
<td>0</td>
<td>3.0</td>
</tr>
<tr>
<td>Co-ordination</td>
<td>7</td>
<td>1</td>
<td>2.9</td>
</tr>
<tr>
<td>Control</td>
<td>10</td>
<td>1</td>
<td>2.1</td>
</tr>
</tbody>
</table>

Table 2. Alignment Paths and Strategic Market Position

5 DISCUSSION

This discussion is organised under four headings. First, the findings are summarized. The relationships among the four alignment paths presented in Figure 2 can be illustrated by overlaying the Focus-Dominance classification upon the MIT90s model. Figure 3 highlights how the different paths to alignment involve different sub-parts of the MIT90s model. This supports Hypothesis 1 that the dynamics of IS alignment are contingent on an SME’s strategic market positioning. Second, the validity threats to the findings reported above are then examined. Third, the implications for research are explored, integrating the extant literature within the contingent model developed here. Finally, the implications of those findings for practice are reviewed.

![Figure 3. Understanding SME Classification through the MIT90s Model](image)

5.1 Findings

Figure 3 shows how the four paths relate both to each other and to the MIT 90s model. At the centre of the model is the basic relationship between management processes and technology to embed the IS investment in the SME. This is a control sub-model within the MIT 90s model. To also capture the benefits from increased co-ordination requires a more complex response by SMEs. Specifically, structural readjustment is required to manage the increased complexity.
Collaboration raises a different challenge. Here, there is a change from a cost- to a value add-based strategy. The control sub-model is developed into a strategic information systems sub-model, in which changes in strategy drive changes in both technology and management processes, and in the relationship between them. Finally, positioning requires changes in all the components of the MIT 90s model. The three sub-models are integrated with up-graded roles and skills. To implement alignment, the existing core competences of the SME would need to be significantly enhanced and new ones developed. This would be a major and high risk strategic change, for which many SMEs and their CEOs would not have the necessary skills and resources.

5.2 Validity Threats

The results are subject to four validity threats. The first threat concerns the representativeness of this sample of SMEs. The sample is not random. Potentially, it over-samples SMEs that are successful, and whose owners are engaged and knowledgeable managers, willing to be involved in a university-based research programme. This bias could affect both the relative numbers of observations across the four quadrants in Table 2 and the frequency with which SMEs select specific paths to alignment.

While a bias in the relative frequency of observations between quadrants is not a threat to the findings supporting Hypothesis 1, a bias in the frequencies with which paths are selected within a quadrant is a potential validity threat. It is reasonable to assume that the SMEs here are both more sophisticated and successful than randomly selected SMEs. In that case, the dominant paths identified above may be less frequently adopted by the typical SME and the findings may be subject to a potential validity threat.

The second validity threat concerns the scoring protocols for identifying the paths to alignment. Two of the authors coded the paths with 100% agreement. A potential validity threat exists as they were aware of the theory being tested, could have been aware of an SME’s location and, therefore, coded that SME’s path to be consistent with the theory. Against this potential threat, first, recall that the authors were coding the paths based on written case reports developed by teams involving others during a 1 week field study of an SME, including interviews with all its senior managers. There is substantial direct information on changes in components and, hence, the paths to re-alignment. Second, issues such as the timing of a restructure or of the acquisition of new technology are not typically subject to interpretation. Similarly, changes in management processes are well documented. However, managers’ descriptions of the timing of changes in roles and skills, and strategy may be open to interpretation.

The third threat is a potential internal validity threat, involving an alternative interpretation of Figure 3. Avison et al. (2004) show that a single large organization followed different paths to alignment for the different IS projects it implemented. Here, this suggests a more complex interpretation of Figure 3. SMEs in the Positioning quadrant could also follow a Co-ordination quadrant path to alignment when investing in IS to manage their customers more effectively. However, SMEs in the Co-ordination quadrant would not invest in Positioning interventions unless they moved to the Positioning quadrant. Similarly, SMEs in the Co-ordination and Collaboration quadrants could follow a Control quadrant path to alignment when investing in IS to improve their internal controls. Essentially, Figure 3 could be interpreted as presenting a hierarchy of paths to alignment, with SMEs in the Positioning quadrant able to follow the other paths as required, and SMEs in the Collaboration and Co-ordination quadrants able to also follow the path in the Control quadrant. This should be the subject of future research.

Finally, the paths identified in Figure 2 are not exhaustive. They do not include the path followed by Flower and Samios (Yetten et al. 1994). That path is technology, followed by roles and skills, structure, management processes and, finally, strategy. This was an emergent, rather than a formally designed path, and maybe infrequently used. Its potential existence does not invalidate the basic hypothesis tested here that the path adopted is contingent on an SME’s position in the Focus-Dominance model. Instead, it implies that an SME’s location in a quadrant does not uniquely define
the path adopted in that quadrant. However, there may only be an emergent and a formal path within the Positioning quadrant, with the paths in the other three uniquely determined.

5.3 Implications for Theory and Practice

Integrating this paper and earlier studies extends understanding of the contingent nature of an SME’s investment in and use of IS, and of the effect of its market position on its management of IS. It also provides guidelines for managers by describing the dominant path to alignment in each quadrant.

In terms of theory, the contingent model developed here allows re-interpretation and integration of the following three sets of findings, which initially appear to be inconsistent. First, benefit realisation depends on alignment between IS and business strategies (Hussin et al. 2002). Second, IS investment is frequently limited to supporting operations and transactions (Foong 1999). Third, organizations with more sophisticated IS tend to perform less successfully than those with less complex systems (Naylor and Williams 1994); the greatest alignment and highest performance are reported for systems to improve efficiency (King et al. 2000); and those that adopt a low-cost approach are unlikely to use IS strategically (Lesjak and Lynn 2000).

Twenty five of the 27 SMEs follow the expected paths to alignment in their quadrants and, therefore, potentially capture the benefits of alignment. However, for SMEs in the Control and Co-ordination quadrants, alignment between IS and business strategy is independent of their paths to alignment in Figures 2 and 3. These null relationships are not included in the typical models (Hussin et al 2002, Spanos et al. 2002) described in the SME literature.

Consistent with the literature (Foong 1999), IS investment is limited to operations and transactions for the 19 SMEs in the Control and Co-ordination quadrants. Also, these SMEs adopt a low-cost strategy, with the relationships between their IS and business strategies unchanged (Lesjak and Lynn. 2000). Finally, given that the complexity of the paths to alignment increases substantially from the Control to the Positioning quadrant, the greatest alignment and highest performance is likely to be achieved by SMEs in the former rather than the latter quadrant. So, findings, which appear on first inspection to be inconsistent, are consistent when integrated within the contingent model developed here.

The patterns in Figures 2 and 3 may also explain Cragg et al.’s (2002) finding that alignment is achieved by many SMEs but is not well understood. It is not well understood because SMEs do not follow the classic path to alignment. Instead, they frequently follow the much less complex paths to alignment in the Control and Co-ordination quadrants. These require only limited change management skills. For those SMEs, the path to alignment is not problematic and, therefore, is frequently realised but not recognised in the literature.

The different paths identify different levels of interventions to achieve re-alignment. It follows that SMEs in different quadrants require different levels of skill in change management to realise the benefits from completing their paths to alignment. However, only in the positioning quadrant is the path to re-alignment likely to prove a challenge to SME management. So, many SMEs are in alignment and capture the related benefits.

In contrast, the literature on large organizations’ IS use focuses on value-adding strategies and the repositioning challenge in the positioning quadrant. This raises the question of whether research on large organizations is limited in scope, and whether it should be extended to investigate organizations investing in IS to support existing management processes rather than focusing almost exclusively on the problems of strategic repositioning.

Turning to practice, Figure 3 documents the path to alignment for each quadrant. The appropriate intervention in the Control quadrant is shallow (See table 2), requiring limited changes in technology and management processes. In contrast, the intervention in the Positioning quadrant is deep, requiring significant change management skills and influencing all factors in the MIT90s model. Therefore, the risk to organization performance is low in the former and high in the latter quadrant.
Finally, few SMEs have the motivation and/or capacity to introduce the integrated systems needed to support IS-based innovation, even though strategic use of IS has become more important in flexible and dynamic markets. Improvisation is one way to respond to challenges in which the innovative use of resources, including IS, is critical (Ciborra 1997). However, while SMEs are typically assumed to be flexible and innovative, their use of IS is relatively inflexible (Levy and Powell 1997), with IS typically used as a means of support. Consistent with this, the dominant use of IS, in 19 of the 27 SME cases, is to support existing management processes.

6 CONCLUSIONS

While it is generally accepted that alignment is a critical enabler of organization performance, there is limited research into the actions or paths by which alignment is achieved. The research that does exist relates almost exclusively to large organizations. This paper adds to the limited literature on SMEs and their paths to alignment and, specifically, it maps the paths to alignment of 27 SMEs. It is hypothesized that the dominant path to alignment is a function of an SME’s market position in the Focus-Dominance model. Nineteen of the SMEs follow paths that are restricted to links among management processes, technology and structure in the MIT90s model. Five of the SMEs work closely with their major customers, whose IS systems strongly influence the SMEs’ IS options. Finally, only three SMEs follow paths to strategic repositioning, the focus of much of the IS alignment research in large organizations. This last finding suggests that either IS strategic choices have more limited impacts on SMEs than on large organizations, or that research on alignment in large organizations should consider other alignment mechanisms, including the less complex interventions identified in Figure 2 for the Control, Co-ordination and Collaboration quadrants versus the Positioning quadrant.

REFERENCES


Smaczny, T. (2001), Is an alignment between business and IT the appropriate paradigm to manage IT in today’s organization? *Management Decision*, 39, 10, 797-802.


SLACK-ENABLED INNOVATION VERSUS PROBLEMISTIC SEARCH: FINDINGS FROM CASE STUDIES AMONG SME

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0283.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Behavioural theories, Resource Based View, Innovation Driver / Enabler / factors / process, Small and medium sized enterprises (SMEs)</td>
</tr>
</tbody>
</table>
SLACK-ENABLED INNOVATION VERSUS PROBLEMISTIC SEARCH: FINDINGS FROM CASE STUDIES AMONG SMES

Heinz-Theo Wagner, heilbronn business school, Bahnhofstr. 1, D-74072 Heilbronn, Germany, wagner@hn-bs.de
Kristina Ettrich-Schmitt, heilbronn business school, Bahnhofstr. 1, D-74072 Heilbronn, Germany, ettrich@hn-bs.de

Abstract

Innovations are important for the development of firms and whole economies. Several theoretic approaches are therefore dealing with innovation and the motivation of firms to motivate. The Resource-based View of the Firm views innovations to be motivated by the use of slack resources while the Behavioral Theory of the Firm predicts problemistic search as an important motivation for innovative maneuvers. Recently, Pitelis proposes an integration of both theories to better explaining the motivation for innovation. This paper empirically tests the predictions from these theories using multiple case studies among small and medium enterprises. The results show that firms’ motivation to innovate is best explained using a combination of both theories.

Keywords: Slack Resources, Problemistic Search, Innovation.
1 INTRODUCTION

Innovations are an important driving force for the prosperity of economies and firms alike. Rapid technological change, increasing globalization, and shorter product cycles render the ability to innovate crucial for firms competing at the global marketplace. Therefore, lots of studies are engaged in innovation research for many years. Within this strand of research, several seminal contributions found problemistic search and use of slack resources as basic motivations for innovations.

The Austrian School of Economics (Schumpeter 1934) provides the central idea of the exploitation of opportunities by creating and adopting innovations that make a competitor’s position obsolete. Superior firm performance, in turn, can only be achieved by continuous innovation. Building on that, the Resource-based view of the firm (RBV) (Penrose 1959) proposes the achievement of a sustained competitive advantage by employing novel services of resources controlled by a firm. Thus using organizational slack to innovate and employing novel services of resources is a key consideration of the RBV. In contrast, the behavioral theory of the firm (BTF) (Cyert & March 1992) is engaged in innovations and focuses on problemistic search as the driving force behind the creation of innovative solutions.

Pitelis (2007) in a recent article discusses the approaches of the RBV and the BTF in terms of their assumptions and predictions regarding the motivation for innovation and the followed-up search strategies for innovation. RBV and BTF provide quite different explanations for innovations while Pitelis proposes an integration of both theories. He calls for the test of the theoretical predictions, among others, with case studies to provide more insights. We follow this call and ask the following research question:

- How to explain different search strategies for innovations and which theory or combination of theories best explains the emergence of innovations?

To address this question, we used case studies among small and medium-sized enterprises (SMEs) that are known as very innovative and successful in their markets.

The remainder of the paper is structured as follows: Section 2 deals with the theoretical foundation that is followed by section 3 explaining the used methodology. In section 4 the case study environment and the case study sample is described. Section 5 presents the case study results. Key findings, limitations, and future research are summarized and critically discussed within section 6.

2 THEORETICAL FOUNDATION

This section first introduces the Behavioral Theory of the Firm, then the Resource-based View and finally the derived propositions.

2.1 Behavioral theory of the firm (BTF)

The classic work of organizational theory, A Behavioral Theory of the Firm’ of Cyert and March (1992) aims to reveal the perception of the business firm being a so-called black box. They seek to provide insight within the internal processes of economic decision making of contemporary organizations as a unit faced with an ever-changing environment (Prietula & Watson 2000).

Cyert and March assume that search within organizations is mainly stimulated by a certain problem. Thus it is strongly related to their concept of problemistic search (Pitelis 2007). Problems are either identified by failure to achieve a certain goal or by the threat of failure in the upcoming future. As long as a specific problem is not solved, an organization will continue to search for a satisfying answer as the declared goal. Thus problemistic search is undertaken as long as the actual performance is below a certain aspiration level. This so-called negative attainment discrepancy induces problemistic search.
problematic search can be defined as search that is motivated by a
certain problem internal or external of an organization and aims at finding an appropriate solution. The
concept of problematic search implies that firms that have not been successful in the past generate
more innovations than more prosperous ones. This assumption could not be verified concerning
important technological improvements in particular. Therefore organizational slack resources were
included in the attempt to predict innovation decisions as these have proven to result in major
technological changes and are in many cases available in successful firms that do not suffer from
shortness (Cyert & March 1992, p. 189). Thus slack resources provide a source of excess funds
enabling search for innovations that would not be realized in times of shortness.

As a consequence the authors distinguish between two types of innovation. The problem-oriented
innovation, developed due to short-run problems and those that are designed to solve major
organizational problems on a long term basis out of slack resources. In conclusion, both, successful as
well as lesser successful firms are innovating (Cyert & March 1992, p. 189). In both cases
(problemistic search and slack search) the existence of a problem (i.e. performance below aspiration
level) is the engine that drives innovation. This is also indicated by the theoretical model forwarded by
Greve (2003a, p. 686) who puts “performance minus aspiration level” as explanans for problematic
search and in a “dotted line” connection also for slack search. The latter also points to the fact that
slack search was introduced into the BTF in a further step to better predict innovation decisions and as
such sometimes appears as sort of an exogenous variable that is not readily explained by the existence
of a problem. In contrast organizational slack is a central tenet of the RBV depicted in the following
section.

2.2 Resource-based view of the firm (RBV)

The starting point for RBV is the assumption of the heterogeneity of resource endowments across
firms that explain performance variations (Peteraf 1993). Thus, a firm’s competitive position results
from bundles of unique resources (Rumelt 1984) that are not perfectly mobile (Spanos & Lioukas
2001). The RBV, therefore, explains sustainable competitive advantage (SCA) by means of the
resources controlled by a firm. The firm itself is viewed as a collection of productive resources guided
by the administrative function (Penrose, 1959) required for market competition. These resource
bundles are found to be immobile, firm-specific, and are developed in a path-dependent way (Teece
1993). The growth of a firm depends on exploiting existing resources and developing new ones
(Penrose 1959; Wernerfelt 1984). “Resources are defined as those tangible (or intangible) assets that
are tied semi-permanently to the firm” (Spanos & Lioukas 2001, p. 909) where “the subdivision of
resources may proceed as far as is useful for the problem at hand” (Penrose 1959, p. 74).

Regarding the core concept “resource”, a basic idea of Penrose (1959, p. 67) is that the same resource
can render many different services. Although important, it is not the resource per se but its use in
productive processes that leads to eventual value. With subsequent operations knowledge of the
services a resource can render is accumulated. With this knowledge the exploitable opportunity set of
the firm is enhanced and excess resources can be put into different services (Penrose 1959). These
slack resources, then, enable innovation and even motivate its use for innovative purposes as long as
the management pursues profit maximization and excess resources can be put into service at no extra
cost (Pitelis 2007) which is a different engine driving innovation than depicted by the BTF.

2.3 Propositions

Pitelis (2007) in his effort to combine BTF and RBV identified several similarities as well as
differences between both theories. The BTF starts from the observation that intrafirm conflict is
paramount (Pitelis 2007, p. 478 f.). This leads to search activities to resolve the conflict. This
problemistic search, then, leads to the discovery and implementation of innovative solutions.
Therefore it can be hypothesized that relatively unsuccessful firms are more likely to innovate (Pitelis
Innovation takes place both in case of immediate problems to be solved and as a means to alleviate conflict in other cases (Pitelis 2007, p. 483). Therefore we formulate:

Proposition 1: Both intrafirm conflicts and the presence of immediate problems lead to innovations to resolve the problem or conflict, respectively.

In the RBV intrafirm conflict is not in focus, and innovations are enabled by slack resources through newly discovered and exploited services rendered possible by the resources of a firm (Penrose 1959). That points to the importance of intrafirm learning for the RBV. Through learning new uses of resources are discovered and exploited. Thus, in growing firms, excess resources are present leading to perennial search for innovation (Pitelis 2007, p. 483). In contrast BTF predicts the use of slack resources to solve major organizational problems on a long term basis (Cyert & March 1992, p. 189). Following the argumentation of Penrose and Pitelis we state:

Proposition 2: There is a continuous search for innovative solutions through learning new uses of resources.

The integrated view of Pitelis (2007, p. 479) “predicts excess resources/slack-motivated and problemistic search-induced innovation, depending, however, on moderating factors such as the immediacy of the problem faced, the degree of maturity of the line of business, and the firm’s productive opportunity”. Thus we formulate:

Proposition 3: Both problemistic search and continuous search for innovations can be observed simultaneously.

3 METHODOLOGY

One form of empirical studies is case studies which are suited for an in-depth investigation of a phenomenon and thus ask how- and why-questions (Yin 2003). In the following sections, case studies carried out among small and medium-sized enterprises (SME) are presented, focusing on the emergence of innovations. Case studies have to be prepared and carried out thoroughly in order to achieve the necessary rigor. During design and preparation it is important to make the research question, propositions and unit of analysis explicit (Yin 2003; Dubé & Paré 2003). The research question employed for this research is: How to explain different search strategies for innovations and which theory or combination of theories best explains the emergence of innovations? How and why questions are considered appropriate for case studies (Yin 2003). The propositions used in the cases are theoretically grounded as depicted in the previous section.

After the definition phase the case study partners were identified by using the assessment of the general managers of two local investment promotion agencies of the most innovative SMEs of their area of discretion of Heilbronn-Franconia. In order to carry out a case study we developed case and interview protocols, and discussed the approach within the research community prior to the actual interviews. Afterwards, the adapted documents are used for the case studies (Yin 2003; Eisenhardt 1989). After this preparation phase the contact phase was commenced. This phase started with calling the identified companies by phone to verify the contact information, to provide preliminary information regarding the objectives of the planned case studies, and to schedule an appointment.

After scheduling the meetings in each case two researchers visited the firm and performed the interviews. The procedure was as follows:

- The objective and the procedure of the case study were explained.
- The semi-structured questionnaire was applied to cover a wide variety of contextual variables.
- The interview was tape-recorded.
- Data was complemented by company reports and academic literature.
- Performing a wrap-up session among the researchers to summarize the essential items.
- Transcription of the tape-recorded interviews.
The interviews lasted in average two hours. The transcripts of the cases were compiled into a case protocol comprising 61,563 words in total. The unit of analysis is the firm, its resources and the approach towards innovation. Therefore the Chief Executive Officers of the firms were interviewed. To evaluate the transcribed interviews NVivo was applied. NVivo is software designed to support qualitative data analysis and simplifies rich text handling through facilitating sorting, arranging and classifying information. A main feature of the software that was utilized in the actual cases is the creation of specific categories, so-called tree nodes. NVivo allows to directly attributing phrases or paragraphs abstracted from the transcriptions to those nodes, following a certain rule for codification developed in advance. This functionality makes it easier to compare answers to specific questions across the case study partners and to trace back constructs and hypotheses to individual answers. After finishing the case studies a final presentation was created and both sent by e-mail to the contact partners and presented within a workshop.

In the following paragraph, the case studies presented are discussed in terms of the requirements of a positivist method. According to Yin (2003) construct validity can be improved by several tactics. The first is the use of multiple sources of evidence, which provides multiple measures for the same phenomenon. In our case studies we carried out interviews in five different firms using top manager as key informants who represent the firm and who’s perceptions can be assumed to be closely related to objective measures (Tallon & Kraemer 2007). In addition, we had access to several documents related to company background. Lastly, Yin recommends maintaining a chain of evidence. This was done by deriving constructs and hypotheses from literature, and employing NVivo which allows us to trace back from conclusions to initial research questions. Internal validity means that a relationship between two variables may be inferred as causal or not. To ensure internal validity, pattern matching may be used (Yin 2003), which involves qualitative but logical deduction (Lee 1989). In our case studies we compared the statements collected with propositions derived from literature in order to detect support or non-support for theoretical propositions. External validity means the generalizability of the findings. Case studies can be used to test a theory. A case study can confirm, challenge or extend a well-articulated theory and is similar to the use of an experiment, because one instance of both methods is sufficient to reject propositions (Lee 1989; Yin 2003). To support a theory several instances are needed. Reliability means the consistency and stability of the study across researchers and time. Yin proposes a case study protocol and a case study database to ensure reliability. A protocol is needed to guide the researcher in a standardized manner. In our case studies this protocol consisted of a proposal containing focus areas, research questions and propositions derived from literature. Furthermore, it consisted of a semi-structured questionnaire, and a set of slides covering research objective, motivation, research model, unit of analysis, and methodology. Besides the protocol, NVivo was used to cover the interview transcripts which allow easy checking for cross-references to other interviews.

4 CASE STUDY DESCRIPTION

4.1 Case study environment

In total we conducted five case studies with interviewees of small and medium-sized enterprises that operate within different classic industries (see Table 1). All five companies of the sample are located within the region of Heilbronn-Franconia, a part of Germany’s third biggest federal state Baden-Wuerttemberg down south. Heilbronn-Franconia measures 4,765 square kilometers and is home to about 900,000 inhabitants. Baden-Wuerttemberg as a whole hosts a well-diversified domestic industry, even though two emphases can be noticed measured by the number of employees: the process engineering sector and the automotive sector. SMEs dominate the overall corporate landscape. On average around 99 per cent of all companies within Baden-Wuerttemberg (and Heilbronn-Franconia as well) can be attributed to this category. SMEs generate about 40 per cent of the volume of sales and provide 50 per cent of the jobs within Baden-Wuerttemberg (Brenner 2008).
An extraordinary feature of Heilbronn-Franconia is a remarkable density of world niche market leading companies. In Kuenzelsau, a town with a population of about 15,000 people situated within the area, the highest concentration of this type of firms in relation to the number of inhabitants within Germany could be identified by a recent survey of a German manager magazine (WirtschaftsWoche, 2008, 19: 53-83). The largest town of the Heilbronn-Franconia district is Heilbronn with about 120,000 people living there. Although being rural and agrarian-oriented Heilbronn-Franconia features good transport connections as being part of the Stuttgart metropolitan area.

4.2 Case study sample

Table 1 summarizes some characteristics of the sampled firms. While TransDrive, InVitro and ProtoType were proposed interview partners of the local investment agency working within the wider Heilbronn area, SafeEx and TechDat were suggested by the agency responsible for the northern part of the region investigated. The managing directors of both agencies were asked to rank their selected firms based on their subjective impression on the innovativeness of the firms considered. Although TransDrive and SafeEx exceed the upper boundary for SMEs in terms of employees as defined by the European Union (i.e. 250), these firms are considered as SMEs by their CEOs because still operated in a family tradition with strong local roots. All figures relate to the headquarters of the firms investigated located within the region Heilbronn-Franconia.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>TransDrive</th>
<th>InVitro</th>
<th>ProtoType</th>
<th>SafeEx</th>
<th>TechDat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branch</td>
<td>transmission and drive train systems</td>
<td>laboratory equipment</td>
<td>construction models and prototypes</td>
<td>safety engineering</td>
<td>technical data information systems</td>
</tr>
<tr>
<td>Established</td>
<td>1935</td>
<td>1964</td>
<td>1964</td>
<td>1975</td>
<td>2004</td>
</tr>
<tr>
<td>size (workforce) 2007</td>
<td>740</td>
<td>110</td>
<td>12</td>
<td>300</td>
<td>80</td>
</tr>
<tr>
<td>Type</td>
<td>LLC</td>
<td>LLC</td>
<td>LLC</td>
<td>LLC</td>
<td>Inc.</td>
</tr>
<tr>
<td>shift in sales (’03-’05)</td>
<td>+ 128%</td>
<td>n.a.</td>
<td>+/- 0%</td>
<td>+ 17.7%</td>
<td>n.a.</td>
</tr>
<tr>
<td>Interviewee</td>
<td>managing director</td>
<td>managing director</td>
<td>managing director</td>
<td>divisional director</td>
<td>CEO</td>
</tr>
<tr>
<td>ranking by MD of local investment agency</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1. Case study sample

5 CASE STUDY RESULTS

5.1 Problemistic search

“Problemistic search occurs as a response to an organizational problem” (Greve 2003b, p. 54) so that “an innovation is a solution that will be implemented if it is matched with an organizational problem” (Greve 2003b, p. 96) in both cases: immediate problems and intrafirm conflicts (Pitelis 2007).

Proposition 1 state that according to the BTF both intrafirm conflicts and the presence of immediate problems lead to innovations to resolve the problem. An example for immediate problems being a driver for a firm’s innovative activity is SafeEx, a provider for high-quality industrial safety technology for applications in explosion-hazardous areas. With Airbus as an important customer SafeEx strives to meet the special requirements of the aviation industry. When Airbus urgently needed to have a possibility to connect cables in a way that guarantees watertightness even at temperatures at
around minus 60 degree Celsius, the engineers of SafeEx managed to combine well-known materials and technologies in an entirely new way. The result was a cable connection system designed to meet the demands put on the material by the conditions occurring at maximum heights. The related quote from SafeEx is presented in Table 2. Additionally quotes of the managing directors of InVitro and ProtoType on proposition 1 are listed:

<table>
<thead>
<tr>
<th>Case</th>
<th>Exemplified quotes</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>SafeEx</td>
<td>“The challenge to meet was to modify the material of the cable for the specifications of Airbus […] We have managed to combine materials that are normally impossible to combine.”</td>
<td>External event: Customer request</td>
</tr>
<tr>
<td>InVitro</td>
<td>“Some customers wanted to have transparent ones (= bottle top dispensers, the authors) so that they could see the interiority. We developed them as a sampling but in the end the material wasn’t chemically resistant.”</td>
<td>External event: Customer request</td>
</tr>
<tr>
<td>InVitro</td>
<td>“Between the manager of R&amp;D and the operations manager an increased tension could be noticed. […] This got out of control leaving me highly dissatisfied. […] As a consequence we have redesigned the organization chart.”</td>
<td>Internal event: Intrafirm conflict</td>
</tr>
<tr>
<td>InVitro</td>
<td>“Sometimes we are forced by external suppliers. […] The electronics is already 10 years old and out-dated and we can’t buy it anymore. Because of that we have bought ahead enough components for our restage. […] And now we have to figure out if our product (= electronic pipette, the authors) needs to be up-dated.”</td>
<td>External event: Environmental change</td>
</tr>
<tr>
<td>ProtoType</td>
<td>“Pricing pressure has never been a driver of innovation. By experience I know that it’s different. […] you notice a question or a problem e.g. raised from production and then you reflect about a possible solution.”</td>
<td>Internal event: Problem detection</td>
</tr>
</tbody>
</table>

Table 2. Immediate Problems

The statements of our case study partners show, that an important driver of a firm’s innovative activity is the pressing need to search for solutions for internal and/or external problems. Being faced with a certain problem all forces are concentrated on turning a deficiency fastest possible into a benefit for the organization, it’s customers, and/or further stakeholders. We could detect four types of “immediate problems”.

(1) External event – customer request:
External problems can be formulated in the sense of a typically urgent customer-specific order that involves finding solutions for specific requirements. This is termed “customer project” and is an externally initiated impulse to search for innovation. It is categorized as an immediate problem because the formulation and acceptance of customer-specific orders almost always comprises a short timeline, because the customer presses to get e.g. the product as fast as possible. In particular, if the customer is an important one, failing to find a solution jeopardizes e.g. market shares and sales of a firm imposing a major problem.

(2) External event – environmental change:
Also externally stimulated are changes in the market or technological environment of the firm causing immediate problems to be solved. Typical examples are shortages in specific raw materials or components due to sudden increases of demand exceeding supply, technological progress rendering obsolete specific components that are no longer produced, technological progress pressuring for reorganize the implemented production technology, and market exit or severe production problems of key suppliers without having direct substitutes. In some cases, in particular regarding technological progress, the problem becomes an immediate one, because a firm did not react in time and now faces an immediate problem. All this requires immediate action to replace e.g. a certain technology by another within a short timeframe to not disrupt production.
(3) Internal event – conflict:
Internally stimulated are immediate problems caused by conflict situations. Similar to the statement regarding technological progress mentioned above, the situation might become critical because there was no appropriate reaction in time and then suddenly must be resolved. The conflict situation in our case was resolved by changing the organizational design which in contrast to the other cases described is not primarily directed towards the innovation of a product.

(4) Internal event – problem detection:
In some cases problems along the business process are detected by firm-internal stakeholders. These are immediate problems if they cause a disruption of the process and thus jeopardizing its outcome. In contrast to simply “repair” a situation by restoring it to the previous state and thus do not change anything, an innovation comes up if the immediate problem is resolved not by “repair” but by coming up with an alternative solution.

The motivation to innovate by resolving immediate problems in any of the above mentioned types of immediate problems is based on negative attainment discrepancy as described in section 2. Furthermore those problems are “immediate” in a sense that they will cause negative effects on performance if unresolved in time. Intrafirm conflict as stated by BTF is one type of immediate problems but not the only one, and externally as well as internally motivated events could be revealed.

5.2 Continuous search for innovation

Superior firm performance can be achieved by continuous search for innovation (Schumpeter 1934) that builds on using organizational slack to innovate by employing novel services of resources (Penrose 1959) through learning. This is expressed in proposition 2 (“There is a continuous search for innovative solutions through learning new uses of resources”). An example for proposition 2 is TechDat, a firm providing data base services for technical data for the automobile industry. One idea of TechDat was not just to check incoming invoices by an information system for the management of vehicle fleets, but to add additional functions to this already available system. Building on both the knowledge of the information system including possible extensions and on market knowledge already available, these additional functions were laid out to streamline the process of car maintenance by avoiding the print-out of different forms at each maintenance step, automatically listing recommended maintenance measures according to the specifics of the car, and to provide a report covering the complete maintenance process. Thus TechDat developed an online tool for managing all maintenance procedures of a fleet without any printing required. This includes the registration of a certain vehicle at a garage, plausibility checking according to the car’s manufacturer’s specifications as well as the approval. At the end of the process the fleet manager can generate a detailed report for a review if needed.
The following table depicts specific quotes from TechDat and other cases centering on the search for innovative uses and therefore referring to proposition 2:

<table>
<thead>
<tr>
<th>Case</th>
<th>Exemplified quotes</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>TechDat</td>
<td>“We asked ourselves, how the printing of an invoice could be avoided. We’ve called up the warranty systems of motor vehicle manufacturers where everything works well without any bills.”</td>
<td>Internally initiated search for improvement</td>
</tr>
<tr>
<td>TechDat</td>
<td>“Information on accredited tires or wheels is not recorded within the registration certificates of new vehicles anymore. Car mechanics nowadays do not even know about the permitted tire-size for a specific car model. […] We have integrated the information the legislator has removed from the registration certificate into software and sell this product to garages.”</td>
<td>Internally initiated search for improvement</td>
</tr>
<tr>
<td>SafeEx</td>
<td>“Around already developed key components we build a product portfolio in order to provide new solutions for our customer.”</td>
<td>Internally initiated search for improvement</td>
</tr>
<tr>
<td>ProtoType</td>
<td>“Related to our business innovation is the utilization of well-known techniques and material in a new way. E.g. things always being made out of aluminum we create out of plastics as a new variant with different properties.”</td>
<td>Internally initiated search for improvement</td>
</tr>
<tr>
<td>TransDrive</td>
<td>“We are already today thinking about what sort of products we could offer in about 10 to 15 years. These products are often beyond the scope of transmission systems.”</td>
<td>Internally initiated search for improvement</td>
</tr>
</tbody>
</table>

Table 3. Search for innovative uses

At first look the data collected from the cases revealed that most firms, independent from the specific industry, looked for new combinations of their resources in order to better serve their markets. However, reading the transcriptions thoroughly shows that this search for innovative uses was not always rigorously and formally applied. Rather the search for innovative uses was internally initiated and based on ideas derived from recognizing potential for optimization of current practices, for example.

Furthermore the observed innovations are all incremental in nature. Depending on the degree of newness and the pace innovations can be classified as either being incremental or as radical if achieving noticeable breakthroughs. Whilst the term incremental relates to only slight changes such as simple product improvements breakthrough innovations can be defined as novel and unique advancements that have the potency to substantially shift the patterns of a market or even create a new one (Zhou & Yim & Tse 2005). However, the innovations mentioned by the interview partners of the sampled companies were incremental in that the firms exploit their deep knowledge in certain technologies and markets to create improvements. This is close to Penrose’s assumption that novel services of already available resources are implemented and thus innovations are created.

All firms in the sample follow a long-term strategy of growth that favors benefits in the long run over short-term profits. This does not mean that short-term profits are not important but they do not play an outstanding and dominant role.

To sum up, in the cases reported above learning about new uses and extending products and services to new application fields without having an immediate problem or conflict is the main driving force of the innovations as predicted by the RBV. Thus proposition 2 is supported.
5.3 Both problemistic search and continuous search for innovations are present

The third proposition deals with both problemistic search and continuous search for innovations which can be observed simultaneously. The term “simultaneously” can be understood twofold:

(1) As simultaneously occurring within the same firm which will be addressed in the next section and
(2) as simultaneously occurring regarding one innovation which will be addressed in this section.

An example for simultaneous problemistic and continuous search is InVitro, the producer of laboratory equipment. It was one of the first firms offering a battery operated pipetting aid for all pipettes from 0.1ml to 200ml. Instead of pipetting manually laboratory technicians now attach glass made pipettes directly to this innovative device. The product features a 5-stage speed key for the selection of the ideal pipetting speed as well as an easily readable color display. It resembles a bit like a pistol nowadays. When it was introduced to the market the shape was similar but much more angular and unhandy. Nevertheless the product was quite successful on the market. Every once in a while users mentioned that the shape of the instrument is not beneficial to fit all different sizes of hands. However stimulated by a certain problem but without high pressure of time the developers of InVitro reassessed the design of the pipetting aid within a slack period in order to provide a better fit for all laboratory technicians. The developers of InVitro conducted several ergonomic studies to find out, how to change the shape of the instrument in order to make pipetting more convenient. Finally they managed to improve the design of the device for a faster, safer and absolutely precise pipetting. The new product shape was even awarded a prize for its exceptional design.

Table 4 contains some quotes on the topic of the combination of problemistic and continuous search for innovation within the development process of the sampled firms.

<table>
<thead>
<tr>
<th>Case</th>
<th>Exemplified quotes</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>InVitro</td>
<td>“In the past it (= the pipetting aid, the authors) was an angular shaped part, a bit like a pistol. […] We have conducted a few ergonomics studies. Now it fits both female as well as male hands, small and big hands.”</td>
<td>Internal initiative triggered by external sensing</td>
</tr>
<tr>
<td>SafeEx</td>
<td>“Our task is to strategically develop the product portfolio initiated by suggestions from the customers. […] it has to become more and more complete. There are quite a few useful components that do not exist yet.”</td>
<td>Internal initiative triggered by external sensing</td>
</tr>
<tr>
<td>ProtoType</td>
<td>“We redesign our internal development process on a systematic basis when we detect an error occurring repeatedly at the same stage. Then we think about how to stop this problem. That is one way how an innovation can come up.”</td>
<td>Internal initiative triggered by internal sensing</td>
</tr>
<tr>
<td>TransDrive</td>
<td>“Our customers force us to innovation by asking for lower costs or the like.”</td>
<td>Internal initiative triggered by external sensing</td>
</tr>
</tbody>
</table>

Table 4. Simultaneous search

The data showed, that in some cases both, problemistic search as well as continuous search for innovations, is present. All the cases present internally initiated innovative activities that refer to detected problems either externally occurring such as usage problems, or internally occurring such as repeated errors e.g. in development or production. The difference to the innovations mentioned in section 5.1 is that these problems are not immediate in a sense that they have to be resolved within a
short time to avoid performance problems. The difference to the innovations mentioned in section 5.2 is that innovations of this section are problem-induced, either internally or externally articulated.

6 CONCLUSION AND FURTHER RESEARCH

Before concluding we discuss the limitations of our study. First, we employed the theoretical lenses of BTF and RBV thereby acknowledging that those aspects of innovation addressed by these theories are visible to us while other aspects are faded out. Second, we used the abstract definition of innovation offered by Schumpeter to allow for a cross-sectional investigation, thus neglecting the more micro-level of innovation with its manifold characteristics and complementarities.

From the cases presented in the previous section we have found support for the propositions to several degrees.

While all firms of our sample used slack resources to innovate either on a continuous basis or triggered by a detected problem, there is no firm in the sample that exclusively used problemistic search to resolve immediate problems.

We have found two firms that are ProtoType and SafeEx, reacting to immediate problems, using slack resources for innovations and combining both approaches. Thus these firms cover the whole bandwidth of motivations and for explaining this behavior both RBV and BTF arguments are needed.

For TransDrive we did not find evidence for problemistic search in case of immediate problems but the use of slack resources for continuous innovation and the combined motivation (see section 5.3).

In contrast, at InVitro we found the combined motivation (see section 5.3) plus the problemistic search in case of immediate problems.

Finally, TechDat was found to only make use of slack resources for continuous innovation.

These findings point to the argument of Pitelis that a combination of RBV and BTF approaches to explain innovation is best suited to analyze the innovation behavior of firms. Only applying BTF or RBV arguments, respectively, are not sufficient.

Nevertheless, all firms of our sample economize on their long-year experiences with certain technologies whether they resolve immediate problems, searching for new applications, or looking for innovations triggered by usage problems, for example. That is, according to Penrose, firms look for novel services of their resources regardless what the original motivation was.

Although a statistical generalization is not possible when using case studies our cases provide evidence for Pitelis’ argument to combine both approaches. The following table summarizes the results.

<table>
<thead>
<tr>
<th>Research question</th>
<th>Indication</th>
</tr>
</thead>
<tbody>
<tr>
<td>How to explain different search strategies for innovations and which theory or combination of theories best explains the emergence of innovations</td>
<td>a) Immediate problems have to be resolved regardless of the context to minimize the negative attainment discrepancy and to avoid severe performance losses</td>
</tr>
<tr>
<td></td>
<td>b) Use of slack resources whether continuously investigated or triggered by concrete challenges is a response to expected mid- and long-term requirements and seem to be positively influenced and motivated by the long-term orientation and strategy of the firms.</td>
</tr>
<tr>
<td></td>
<td>c) As proposed by Pitelis a combination of BTF and RBV best explains the emergence of innovations</td>
</tr>
</tbody>
</table>

Table 5. Summary
For this study we decided to select the top performing firms based on their perceived innovativeness (see section 4) rather than to restrict the sample to only IT innovators to test the general arguments of the theories involved. The next step in research will apply the insights of this study to a sample of IT innovators to test for possible differences in the explanation of search strategies. In this context more in-depth studies are needed to better understand in which situations and under which conditions firms use problemistic search, slack resources, or both. In particular, the time horizon of strategies and also the ownership structure is a promising field of future research to further clarify the call of Pitelis to investigate the contexts and motivations of innovation.

References


A QUANTITATIVE EVALUATION OF NFC BASED CONTACTLESS PAYMENT SYSTEMS IN RETAIL

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0497.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Mobile systems, E-commerce (B2B / B2C / B2G / G2C), Ubiquitous systems, Cost-benefit analysis</td>
</tr>
</tbody>
</table>
A QUANTITATIVE EVALUATION OF NFC BASED CONTACTLESS PAYMENT SYSTEMS IN RETAIL

Wiechert, Thomas J. P., ITEM-HSG, University of St. Gallen, Dufourstrasse 40a, 9000 St. Gallen, Switzerland, thomas.wiechert@unisg.ch

Thiesse, Frédéric, ITEM-HSG, University of St. Gallen, Dufourstrasse 40a, 9000 St. Gallen, Switzerland, frederic.thiesse@unisg.ch

Fleisch, Elgar, ITEM-HSG, University of St. Gallen and D-MTEC, ETH Zurich, Dufourstrasse 40a, 9000 St. Gallen, Switzerland, elgar.fleisch@unisg.ch

Abstract

Near Field Communication (NFC) technology allows for the implementation of novel contactless payment systems in stationary retail. In this paper, we quantitatively analyze the impact of such systems on a retailer's payment costs on the example of real-world data from a Swiss food retailer. Our results indicate that the introduction of contactless payment under current card fee models would in virtually any case significantly increase the payment costs due to the substitution of low cost cash payments for expensive card payments. This increase might be balanced out by a substantial growth in sales, a reduction in operating costs, or a reduction of card transaction fees.

Keywords: Near Field Communication, Retail, Contactless Payment, Mobile Systems.
1 INTRODUCTION

The term “Near Field Communication” (NFC) denotes a wireless communication technology that operates at 13.56 MHz and is compatible to the international industry standards ISO/IEC 14443 and ISO/IEC 18092. The integration of a NFC module into mobile phones or PDAs enables these devices to act as contactless smartcard as well as to read from and write onto such cards. Possible applications include the use of NFC devices for mobile payments, as electronic tickets, for the participation in loyalty programs, and for the storage of rebate coupons. The fact that NFC compatible devices can hold a large number of virtual smart cards in the form of secured applications allows for providing their owners with access to a wide variety of contactless services without the need to carry numerous plastic cards. In recent years, most manufacturers of mobile phones have developed corresponding prototypes. The Nokia 6131 NFC phone has, to date, been the only NFC-compatible device freely available on the market, while all others were available to technology trial organizers only.

In recent years, numerous trials of NFC-based payment have been conducted world-wide. The first publicized example was started in August 2005 in the Dutch town of Kerkrade (KPN 2005). In November 2006, the first trial to implement an EMV-compliant payment process and to use the SIM as secure element for the storage of the NFC application was conducted in the French Strasbourg. This brought the contactless payment process closer to the conventional payment process called “Chip and PIN” as it is practiced in Europe (NFC Forum 2006). In November 2007, the French cities of Caen and Strasbourg saw the launch of the first trial to include all mobile network operators and all banks of an entire country (Balaban 2007).

Against this background, our contribution specifically considers the use of NFC as a novel technology for the realization of payment systems in stationary retail. Retailers are currently confronted with several challenges in regard to their relationship with their customers. According to Chu and Morrison (2002), 75% of retailers view waiting lines at the checkout as the most important factor for improving the customers’ shopping experience. Measurements by large payment system providers seem to indicate that NFC might be able to significantly speed up the payment related part of the checkout process. The payment system providers claim to have measured contactless payments as being 20% to 63% faster than cash transactions and 20% to 53% faster than magnetic stripe based card transactions (cf. Table 1). Additionally, Visa states that the implementation of their contactless solution has lead to a waiting line reduction of up to 23% (Visa 2008).

<table>
<thead>
<tr>
<th>Measured by</th>
<th>Cash</th>
<th>Magnetic Stripe</th>
<th>Contactless</th>
<th>Acceleration through Contactless Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Express</td>
<td>33.7</td>
<td>26.7</td>
<td>12.5</td>
<td>63%</td>
</tr>
<tr>
<td>MasterCard</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>20%</td>
</tr>
<tr>
<td>Visa</td>
<td>34.0</td>
<td>24.0</td>
<td>15.0</td>
<td>56%</td>
</tr>
</tbody>
</table>

Table 1: Transaction times for the payment by cash, magnetic stripe cards and contactless cards

According to Visa and MasterCard their contactless solutions also leads to a growth in sales for adopting retailers. This would come about by means of an increased number of transactions and larger average expenditure by those customers that pay contactless. MasterCard states that PayPass users visit stores accepting contactless payment cards 33% to 52% more often than other customers, and that they spend 28% to 42% more per purchase than they did previously (MasterCard 2008). Visa, on the other hand, claims that customers using their contactless payment solution PayWave spend 22% more than those using cash (Visa 2008). Both companies attribute this increase in spending to the elevated transaction convenience and the fact that the shoppers’ expenditure is not curtailed by the amount of cash they carry.
It is understood that the before mentioned figures should be interpreted with care since the detailed data underlying these published transaction times has not been made available for scrutiny. Furthermore, the studies by payment service providers ignore the financial impact that the implementation of a new payment technology would have, due to the occurring transfer of payment transactions between payment methods. Since each of these methods (e.g. cash, debit cards and credit cards) is subject to its own individual cost and fee structure these transfers would mean a change in the payment costs that retailers have to bear. The most significant transfer would probably occur from low cost cash payments to more expensive card payments, which would lead to an increase in payment costs for retailers. This increase in costs might however be recouped in part or neutralized entirely by the growth in sales promised by the payment system operators. Other means of preventing a cost increase for the retailers are a general reduction in card fees or a special rate for small payments.

Our research concerns itself with the following question: how would the implementation of NFC-based contactless payment affect the payment costs of stationary retailers? For this purpose, we assess the financial impact of novel payment technologies with the help of a mathematical model. We subsequently illustrate its practical use on the example of a data set of payment transactions provided by a Swiss operator of small supermarkets and convenience stores. Moreover, we analyze the influence of a growth in sales on payment costs as well as the impact of reduced transaction fees.

2 RELATED WORK

In their recently published review of literature focusing on mobile payment, Dahlberg et al. (2008) classify 73 relevant publications, in accordance to the stakeholders and issues that they focus on. While 29 publications focus on the technological aspects of mobile payments and 20 on the consumer, only five papers focus on mobile payment providers (Kreyer, et al. 2003, Vilmos and Karnouskos 2003, Karnouskos 2004, Vilmos and Karnouskos 2004, Zmijewska and Lawrence 2005) and four center on retailer-related issues (van der Heijden 2002, Ondrus and Pigneur 2004, Mallat and Tuunainen 2005, Teo, et al. 2005). The literature review’s authors state that the number and diversity of mobile payment publications focusing on retailers are disappointing, and further claim that quantitative studies are needed in order to contribute to a better understanding of merchant adoption. This paper contributes to filling this gap by analyzing the financial impact that the implementation of NFC-based mobile payment will have on retailers and by indicating to mobile payment providers how mobile payments will have to be priced so as to make this technology a viable alternative for retailers.

Not all of the work on mobile payment is however related to this paper without qualification. The term “mobile payment” not only describes payment by means of a mobile device at a point of sale in a retail store, which is the focus of our work. It also includes three other usage scenarios: the use of mobile devices for the payment of mobile content (e.g. mobile phone ringtones), the settlement of purchases made in online stores and its use for consumer to consumer payments (Kreyer, et al. 2003). As this paper is focused on payments in stationary retail, the publications that focus on the other mobile payment scenarios loose some of their relevance.

Mallat (2004) provides a review of literature dealing with information technology adoption and acceptance and discusses the usability of the applied theories in explaining the adoption of mobile payment. While Dewan and Chen (2005) isolate factors which will influence the adoption of mobile payments, Mallat (2007) discusses the inherent advantages of mobile device based payments in comparison with other payment technologies, and how these will advance mobile payment adoption. Amin (2007) studied the attitude of Malaysian bank customers towards mobile credit cards, while Kristoffersen et al. (2008) researched the consumer perception of mobile-based micropayments. Klee (2006) analyzed the time-saving of switching from one payment method to another based on large quantities of POS data. Her results show that debit card payments take on average 30% less time than check payments. A comparable study for contactless and mobile payments is not available.
Several publications have discussed aspects of the costs of payment. Worthington (1996) compared the costs incurred through payment by various methods and emphasized the importance of costs in the adoption of new payment methods. Ardizzi (2004) discussed cost structures and efficiency of payment systems on the basis of the Italian credit card system. Bean (2006) analyzed past developments of card payment fees and made suggestions on how to lower them. Huchzermeier and Van der Heyden (2007) concluded that payment costs are too high due to the inefficiency of the payment card market. They suggest that more retailers should become payment scheme operators, thereby becoming competitors to the established payment systems. Guibourg and Segendorff (2007) made an estimate of the Swedish bank sector’s payment costs and concluded that there are large cross subsidies between the profitable card payments and the unprofitable cash payments. Finally, Papaefstathiou and Manifavas (2004) evaluated the transaction costs for micropayment systems and underlined the importance of transaction costs for the success of electronic micropayment systems. There were, however, no publications on the financial impact that the implementation of a payment instrument or transfer between payment methods would have on a retailer.

We are not aware of any publications that describe the financial impact of payment method and payment technology adoptions on retailers. This seems to constitute a neglected field of research. The publications closest to the methodology of this thesis are a series of articles which analyzed the financial impact that discounts for shoppers who pay cash, instead of using a credit card, would have on retailers (Gordon, et al. 1977, Ingene and Levy 1982, Grant, et al. 1985). The fact that confounded the matter was that retailers would have to give discounts to all cash-paying customers, regardless of whether they had originally intended to pay with credit cards or not. Comparable publications dealing with the financial impact of new payment methods or payment technologies do not exist.

3 METHODOLOGY

We developed a mathematical model in order to measure the financial impact that the decision to adopt a new payment technology would have on a stationary retailer, which we shortly describe in this section. The main input parameters of the model are the concerned retailer’s revenue structures, its current payment method specific costs and fees, and the anticipated transfers of transactions and revenue between payment methods caused by the implementation of the new technology.

Revenue structures. This paper refers to the allocation of payment transactions and revenue to the different amount ranges and the various payment methods accepted by a retailer as revenue structures. The empirical data that our analysis relies on was gathered from a survey among Swiss retailers in autumn 2007. In the following, we will set the focus on one specific data set from a Swiss operator of small supermarkets and convenience stores. Its large share of micropayment transactions – defined as all payment ≤ 15 CHF1 – makes its industry segment particularly suitable for the implementation of a new payment technology that specifically addresses small value transactions. This choice was made, because it is the card scheme operators declared intention to use contactless payment to target the market for small amount payments (MasterCard 2008).

Payment method specific costs. Each payment method is subject to its own individual cost and fee structure. Cost elements can be based on the number of transactions and the revenue. Card payment fees are considered confidential by, both, retailers and payment scheme operators. Approximate average card fees for the Swiss market were, however, obtained from the Swiss Association for Electronic Payment Transactions (cf. Table 2). The cash costs that we use in our calculations originate from company-internal figures of two large Swiss retailers, which both calculated cash payment costs independently of each other and came to similar results.

---

1 The separation between micro- and macropayments is not consistent in prior academic studies and usually varies between 5 and 20 €. We therefore chose 10 € (approximately 15 CHF) as mean value.
Table 2.  

<table>
<thead>
<tr>
<th>Costs per transaction</th>
<th>Costs per revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash</td>
<td>0.00 CHF</td>
</tr>
<tr>
<td>Debit Card</td>
<td>0.30 CHF</td>
</tr>
<tr>
<td>Cash Card</td>
<td>0.02 CHF</td>
</tr>
<tr>
<td>Credit Card</td>
<td>0.15 CHF</td>
</tr>
<tr>
<td>Credit Card (Food Retail)</td>
<td>0.15 CHF</td>
</tr>
<tr>
<td>Private Label Card (retailer operated)</td>
<td>0.00 CHF</td>
</tr>
</tbody>
</table>

Substitution effects. This paper refers to the transfers of payment transactions and revenue between payment methods, caused by the implementation of a new payment method, as “substitution effects”. The substitution effects are an unknown factor in the mathematical model which cannot be determined in advance. Additionally, they will differ between different countries and amongst retail industry segments, due to various influence factors. For this reason, we developed three distinct scenarios that cover various possible outcomes of the adoption process (cf. Table 3).

- Scenario 1 illustrates the transfer of payments from cash to credit cards in the micropayment ranges between 0 and 15 CHF. It assumes that 25% of all cash transactions for less than 15 CHF will be replaced by credit card payments.
- Scenario 2 represents the transfer of payments from cash to debit and credit cards in all amount ranges. 30% of the cash-based micropayments are assumed to be transferred to debit and credit card payments in equal shares. A smaller share of cash-paid macropayments (defined as all payment > 15 CHF) are also expected to be transferred: 20% of the cash payments between 15.01 and 50 CHF and 10% of those above 50 CHF are also assumed to be replaced by debit and credit card payments in equal shares. The transfer rates are assumed to be lower for macropayments, because these amount ranges already feature a higher card payment share than micropayments.
- Finally, scenario 3 includes the same substitution effects as scenario 2, however, it incorporates the possibility of retailer-operated store cards becoming more prominent due to the added convenience of contactless technologies. In contrast to the first two scenarios, scenario 3 not only includes the transfer of cash payment to other payment methods, but also the transfer payments from debit and credit cards to retailer-operated store card schemes.

Table 3. Specification of substitution scenarios

Payment costs consist of two elements: The payment method specific transaction-based costs \(C_{\text{Pay}}\), which have to be paid for each transaction and the payment method specific revenue-based costs \(C_{\text{Re}}\), which are proportional to the amount of a transaction. The multiplication of the transaction-based costs with the number of transactions settled with a payment method \(i\) \((\text{Tr}_{\text{old}_i})\) and the multiplication of the revenue-based costs with the revenue paid for with that payment method \((\text{SR}_{\text{old}_i})\)
results in a retailer’s costs for its use. The sum of payment costs for all payment methods accepted by a retailer equals that retailer’s current total costs of payment \( C_{\text{old}} \).

\[
C_{\text{old}} = \sum_{i=1}^{n} \left( C_{\text{TB}_i} \cdot Tr_{\text{old}_i} + C_{\text{RB}_i} \cdot Sr_{\text{old}_i} \right)
\]

The application of the substitution scenarios permits to give estimates of the future distribution of transactions onto payment methods and the new payment costs \( C_{\text{new}} \) caused by the implementation of the new payment technology. This model uses substitution effects, in the form of percentages to articulate which share of transactions \( ST_{\text{ji}} \) and revenue \( SSR_{\text{ji}} \) are transferred from one payment method to another after the introduction of the new payment technology. A payment method does not necessarily have to be solely a beneficiary or benefactor of a new payment technology. It is possible that a new payment technology may cause it to receive a share of the transactions \( ST_{\text{ji}} \) and revenue \( SSR_{\text{ji}} \) from some payment methods, while at the same time forfeiting a share of its own transactions \( ST_{\text{ji}} \) and revenue \( SSR_{\text{ji}} \) to others. The difference between the old payment costs and the new payment costs constitute the financial impact of the new payment technology:

\[
C_{\text{new}} = \sum_{i=1}^{n} \left( (C_{\text{TB}_i} \cdot Tr_{\text{old}_i} + \sum_{j=1}^{n} (Tr_{\text{old}_i} \cdot ST_{\text{ji}}) - \sum_{j=1}^{n} (Tr_{\text{old}_i} \cdot ST_{\text{ji}})) + (C_{\text{RB}_i} \cdot Sr_{\text{old}_i} + \sum_{j=1}^{n} (Sr_{\text{old}_i} \cdot SSR_{\text{ji}}) - \sum_{j=1}^{n} (Sr_{\text{old}_i} \cdot SSR_{\text{ji}})) \right)
\]

When interpreting our results, it should be noted that our approach comes along with a number of modeling assumptions, which might limit its applicability in practice. First, we arbitrarily separate micropayments \( \leq 15 \text{ CHF} \) and macropayments \( >15 \text{ CHF} \) and restrict the growth in sales to the first. It is, however, possible that a new payment technology could lead to increased sales among large amount transactions as well, even if it might do so to a smaller extent. Secondly, we assume that payment costs increase linearly to the growth in sales, i.e. the transactions caused by the implementation of the new payment technology have the same average value than the previous micropayment transactions. Thirdly, the model does not account for infrastructure investments that are necessary on the part of the retailer. These are excluded for two reasons: On the one hand, the investment costs are complex and strongly dependent on a retailer’s existing point-of-sale and payment terminal infrastructure, and can thus not be generalized. On the other hand, card fees constitute a more significant amount of money than hardware does. Finally, the model does not account for cost savings which might be possible through the acceleration of the check-out process. If a new payment technology could accelerate this process, it might lead to a reduced need for cashiers and checkouts. This could decrease the retailers operating costs and refinance a part of the possible additional costs. These potential savings were not integrated into the model due to the high uncertainty with regard to their realization and to the level to which they are dependent on store formats and sites.

4 EVALUATION

4.1 Case Description

The retailer that we consider in the following is one of the largest food retailers in Switzerland with approximately 600 small supermarkets and convenience stores, and an annual revenue of 1 billion CHF. In their case, the 40 million annual sales transactions are mostly micropayments, i.e. 61.61% of all transactions are valued at less than 15 CHF, while only 2.8% are above 50 CHF (cf. Figure 1). With regard to revenues, on the other hand, micropayments amount to 24.93% only, whereas macropayments account for 75.07%. The large number of micropayments represents a significant potential for a novel payment technology focused on the convenient settlement of small value
transactions. While a large share of the customers’ purchases are probably planned based on specific household needs, supermarkets and convenience stores can also incite their customers to make spontaneous purchases. A new convenient payment method which frees the shoppers from having to carry cash could thus lead to a significant growth in sales through higher spending per purchase.

Figure 1. Distribution of transaction volume across payment ranges

As Figure 2 illustrates, the shares of cash payment in the micropayment and the macropayment area differ strongly. 82.84% of the sales transactions in the respondent’s stores are paid in cash, 16.81% with debit cards, and only 0.33% with credit cards since the supermarket operator only accepts credit cards in those stores that are located in tourist areas. Cash cards are not accepted. The retailer neither offers gift certificates nor a self-operated store card. While 99.6% of the micropayment transactions are settled in cash, the same holds for only 55.94% of the macropayment transactions. With higher payment amounts, the card payment quota increases amongst micropayments and decreases amongst macropayments. This is in part due to the fact that some commercial customers, such as restaurants, are allowed to pay on account (i.e. 0.01% of transactions).

Figure 2. Payment method use by payment range

4.2 Payment Costs

The company’s current average payment costs amount to 0.09 CHF per transaction, and 0.44 CHF per 100 CHF of revenue. The financial impact with regard to the implementation of NFC based payment that we calculated from our model in the three substitution scenarios is given in Table 4. The figures show that the implementation of NFC based contactless payment would cost the retailer additional money in the occurrence of any of the three proposed scenarios. Scenario 1 causes the lowest costs, because cost increases are limited to micropayments. Scenario 2 causes the highest additional costs since a significant share of large value payments are moved from cash to the more expensive debit and credit cards. An occurrence of scenario 3 would be less expensive than scenario 2 for two reasons: Firstly, some of the cheap large value cash payments are transferred to the even cheaper private label payment card. Secondly, some of the more expensive credit and debit cards are considered to be transferred to a private label card as well, and thereby reduce the cost increase.
In all three scenarios, the costs for micropayments would increase by between 290% and 366% when compared to the company’s current payment costs. The costs for macropayments on the other hand would not increase in the occurrence of scenario 1, and increase by only 22.16% and 12.95% in the occurrence of scenarios 2 and 3, respectively. This is due to the fact that (a) our scenarios assume a higher substitution rate for lower amount payments, and (b) that the current card payment share is higher in the macropayment range, thus already causing higher payment costs for macropayments.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>All Payments</th>
<th>Micropayments</th>
<th>Macropayments</th>
<th>Increase</th>
<th>Total payment cost [CHF]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current</td>
<td>0.09</td>
<td>0.02</td>
<td>0.20</td>
<td>-</td>
<td>4,398,329</td>
</tr>
<tr>
<td>Scenario 1</td>
<td>0.13</td>
<td>0.08</td>
<td>0.20</td>
<td>43.26%</td>
<td>6,301,066</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>0.16</td>
<td>0.10</td>
<td>0.24</td>
<td>73.46%</td>
<td>7,629,130</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>0.15</td>
<td>0.10</td>
<td>0.22</td>
<td>63.38%</td>
<td>7,186,101</td>
</tr>
</tbody>
</table>

Table 4. Payment cost before and after NFC implementation

According to these results, it seems that the implementation of NFC based contactless payment would, in any case, substantially increase payment costs. Adopting retailers would thus either have to achieve additional sales that enable them to recoup some of these costs, or be willing to relinquish a share of their return on sales to the payment scheme operators. Otherwise, the retailer would have to make the implementation of contactless payment depend on a reduction of payment fees. These reductions could be applied globally onto all payments or be restricted to small amount payments.

4.3 Growth in Sales

One way of neutralizing increased payment costs is additional earnings achieved through growth in sales. The model underlying the presented results restricts the growth in sales to the micropayment area. This assumption is based on the premise that micropayments are much more likely to be characterized as spontaneous purchases whose number could increase due to a more convenient shopping environment than macropayments. The growth in sales necessary to cover for higher payment costs is dependent on a retailer’s margin, which determines his earnings. The lower a retailer’s margin, the greater the additional sales need to be in order to produce the additional earnings necessary to outbalance additional payment costs. The contour lines (i.e. isobenefit curves) in Figure 3 describe these interrelations for the three substitution scenarios. All growth in sales/return on sales combinations below the lines would lead to a negative financial impact on the retailer, while all combinations above the lines indicate a positive impact of the NFC implementation.

- On the occurrence of scenario 1, a return on sales of 3.3% or lower would not enable the respondent to cover the increased payment costs through a growth in sales at all. At a margin of 4% the retailer, for instance, would have to increase his micropayment sales by 112%. A 50% growth in sales, which surmounts the sales increases proposed by MasterCard and Visa (MasterCard 2008, Visa 2008), would only suffice to cover the additional cost of payment if the retailer’s margin is 4.8% of revenue or more.
• In the event of scenario 2, a return on sales of 3.4% or lower would not enable the respondent to cover the increased payment costs through a growth in sales at all. At a margin of 4.5% the retailer would have to increase his micropayment sales by 126%. A 50% growth in sales would only suffice to cover the additional cost of payment if the retailer’s margin is 6% of revenue or more.

• On the occurrence of scenario 3, a return on sales of 2.3% and lower would not enable the respondent to cover the increased payment costs through a growth in sales at all. At a margin of 3% the retailer would have to increase his micropayment sales by 165%. A 50% growth in sales would only suffice to cover the additional cost if the retailer’s margin is 4.5% of revenue or more.

Figure 3. Growth in Sales Necessary to even out Payment Cost Increase

The curves representing scenarios 2 and 3 resemble each other because both scenarios influence payment transactions in all amount ranges and have similar substitution effects. The difference between them is the enhanced role of the store card in scenario 3, responsible for the lower necessary growth in sales rates for all possible return on sales rates. On the other hand, the curve representing scenario 1 stands out due to its different shape and the fact that it intersects the curve representing scenario 3. This is due to the fact that scenario 1 aims only at micropayments thus representing a very different setting. When interpreting the necessary growth in sales rates in Figure 3, it has to be kept in mind that micropayments represent only 24.93% of the respondent’s revenue. A growth in sales of 100%, restricted to micropayments, thus only equates to a total growth in sales of 24.93%.

4.4 Card Fee Reduction and Micropayment Discount

The costs caused by the implementation of NFC could also be partially or fully balanced out by an overall reduction of payment fees or a special discounted fee restricted to micropayments. Should the implementation of contactless payment not lead to a growth in sales at all, the supermarket operator would have to negotiate fee reductions of 84.03% (scenario 1), 51.42% (scenario 2), or 46.26% (scenario 3). If the implementation of contactless payment were to lead to a growth in sales, the answer to the question concerning necessary card fee reductions becomes more complex. In this case, the reduction necessary is, as is the growth in sales discussed beforehand, dependent on the retailer’s margin. This is due to the fact that a growth in sales also leads to an increase in payment costs, due to the higher number of transactions and the increased revenue.

If the retailer’s margin is lower than the average payment costs of the new transactions, the reduction of fees that is necessary increases when compared to a situation without growth in sales. If, on the other hand, the retailer’s margin is higher than the average payment cost incurred by the new transactions, the reduction necessary to balance out the cost of contactless payment implementation becomes lower. Figure 4 shows the necessary card fee reductions in a highly optimistic case of 50% growth in sales – restricted to micropayment sales – for all three substitution scenarios.

If we assume that fee reductions are given only for micropayments, the retailer would have to negotiate for discounts of 92.47% (scenario 1), 123.13% (scenario 2), and 106.47% (scenario 3) if the implementation of contactless payment does not lead to any growth in sales at all. On the occurrence
of scenarios 2 and 3 without any growth in sales, the payment scheme operators would actually have to subsidize the retailer for each micropayment transaction, so that the retailer is able to pay the additional costs for macropayments with this income. This clearly does not seem to be a realistic outcome for negotiations between payment scheme operators and retailers. The cost increases on the occurrence of scenarios 2 and 3 could however be neutralized by a micropayment discount, provided that the implementation of contactless payment leads to a growth in sales. As depicted in Figure 5 (also in case of 50% growth in sales), the necessary micropayment discounts become dependent on the retailer’s return on sales in the case of additional sales incurred by the implementation of NFC based contactless payment.

![Figure 4](image-url)  
**Figure 4.** Necessary card fee reduction in case of 50% growth in sales

![Figure 5](image-url)  
**Figure 5.** Necessary micropayment discount in case of 50% growth in sales

5  
**CONCLUSION**

This paper concerned itself with the question of how the implementation of NFC-based contactless payment would affect the payment costs of stationary retailers. As our numerical results indicate, the introduction of NFC based contactless payment under current card fee models would in virtually any case significantly increase the payment costs for retailers. This is due to the transfer of payment transactions and revenue from low cost settlement with cash to more costly card payments. Ceteris paribus these increased costs would lead to a reduction of the retailer’s profit. While the results in this paper are based solely on the analysis of one retailer’s data, the authors validated the results using the data from three further retailers which led to consistent results.

There are, however, various means to recoup a share of the additional costs. Firstly, the implementation of NFC-based payment could cause a growth in sales due to the elevated transaction convenience and the fact that the shoppers’ expenditure is not curtailed by the amount of cash they carry. However, the growth rates necessary to neutralize the increased costs are substantial.
Secondly, a reduction in card fees for all transactions or, alternatively, a discount restricted to small amount payments could reduce or neutralize the payment cost increase as well. In the event of the proposed substitution scenarios occurring, the overall reduction in card fees would have to lie between 46.4% and 84%, while the micropayment discount would have to amount to between 92.5% and 123.1%. A discount beyond 100% implies that the payment scheme operators would have to subsidize the retailer for each micropayment transaction, so that the retailer is able to pay the additional costs for macropayments. This clearly does not seem to be a realistic outcome for negotiations between payment scheme operators and retailers. If the implementation of contactless payment were to lead to a growth in sales, the answer to the question concerning necessary card fee reductions becomes more complex. In this case, the reduction necessary is, as is the growth in sales discussed beforehand, dependent on the retailer’s margin. If the card scheme operators are not willing to agree to lower card fees, the implementation of a retailer operated store card constitutes an alternative.

While lower transaction fees would make the processing of small payments a less attractive source of revenue for the payment service providers, the launch of such a service could become a means to save costs, especially for the issuer banks, which retain a large share of the card fee revenues. In general, a cardholder’s issuer bank is also the institute at which he has his bank account and where he obtains cash at ATMs. The transfer of payments from cash to NFC-based cards would very likely result in less cash withdrawals, thus leading to a reduction in cash provisioning costs for issuer banks. Banks could seek further savings by delivering payment cards to NFC devices over the air instead of sending plastic cards by postal service. These possible savings should be a part of the considerations when pricing NFC-based payment services.

Additional costs that cannot be compensated by a growth in sales, lower fees or operating cost savings have to be considered as the retailer’s price for the implementation of NFC based contactless payment. Each retailer has to decide individually, whether the implementation of contactless payment is worth this price. Criteria favoring retailers to spend money on the new payment technology are increased customer convenience due to faster payment transactions, quicker check-out processes and possible added value through other contactless customer services such as loyalty cards, coupons or seamless combinations of different services.

References


NFC BASED SERVICE INNOVATION IN RETAIL: AN EXPLORATIVE STUDY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0587.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Mobile systems, Service innovation, Ubiquitous systems, M(obile)-banking / M-commerce</td>
</tr>
</tbody>
</table>
NFC BASED SERVICE INNOVATION IN RETAIL: 
AN EXPLORATIVE STUDY

Wiechert, Thomas J. P., ITEM-HSG, University of St. Gallen, Dufourstrasse 40a, 9000 St. 
Gallen, Switzerland, thomas.wiechert@unisg.ch

Schaller, Andreas, Technology Consulting, Schulstrasse 11, 95676 Wiesau, Germany, 
andreas.schaller@t-online.de

Thiesse, Frédéric, ITEM-HSG, University of St. Gallen, Dufourstrasse 40a, 9000 St. Gallen, 
Switzerland, frederic.thiesse@unisg.ch

Fleisch, Elgar, ITEM-HSG, University of St. Gallen and D-MTEC, ETH Zurich, 
Dufourstrasse 40a, 9000 St. Gallen, Switzerland, elgar.fleisch@unisg.ch

Abstract

This paper presents the results of a survey conducted amongst European retailers on their plans and 
perceptions with regard to novel applications based on Near Field Communication (NFC) technology. 
Whilst the survey showed that retailers did evaluate the possible applications differently, none of them 
clearly stood out as the most beneficial one. NFC based services were on average conceded to be able 
to accelerate the checkout process at the point-of-sale. This is an important acknowledgment, as 65% 
of the respondents consider waiting lines which can occur at the checkout to hold the largest potential 
for an improvement of their customers’ shopping experience. While the rate of agreement is somewhat 
smaller than in the case of process acceleration, retailers also concede that NFC applications could 
make shopping more convenient for their customers. On the other hand, none of the NFC-based 
applications seemed to yield cost saving potentials to the responding retailers. The most surprising 
result of the survey was the respondents’ low expectations in regards to customer acceptance. This is 
in clear contrast to the reports on NFC trials which generally describe participants as enthusiastic 
about the technology.

Keywords: Near Field Communication, Retail, Payment Systems.
1 INTRODUCTION

1.1 Background

Retailers are currently confronted with several challenges in regard to their relationship with their customers. Shoppers expect retailers to sell quality products and to provide them with good service levels within their stores. However, they also expect the retailers’ prices to be as low as possible. The last several years have seen fierce price competition, which has forced retailers to cut costs wherever possible. Apart from the optimization of supply chain activities, through initiatives such as Efficient Consumer Response (ECR), retailers have also reduced the number of employees in their outlets. This cutback has led to a lower employee to customer ratio in retail stores, which in turn created new problems e.g. longer waiting lines at the point-of-sale (POS) and difficulty for customers in finding store personnel for assistance.

Today, 95% of retailers consider waiting lines at their checkouts as the most serious problem to be solved in order to improve service for their customers (Chu and Morrison 2003). A long term study conducted by Zenith Management Consulting states that 92% of retailers have problems meeting their customers service expectations (Cowgill 2006). This assessment is confirmed by a survey of 1,000 US retail customers, conducted by the Verde Group, according to which the biggest perceived problems that shoppers encounter in retail stores are a lack of support by store personnel and long lines at the checkouts. A third of the respondents stated that they could not find store representatives when they needed help, or that representatives did not have enough product knowledge to be able to assist them. 6% of the respondents declared that they have left retail stores without purchasing anything, precisely because of these problems (Verde Group 2007).

These problems can lead to various negative consequences in the short, medium and long term: Crowded stores, and the necessity to wait at the checkout, cause shoppers to take a place in the waiting lines as soon as possible. Additionally, they reduce their shopping to that which is absolutely necessary, so as to be able to leave the store at the time they intended to. This can directly lead directly to a decline in sales for the retailer. On the other side, negative experiences have an influence on shoppers in their future choice of a store (Harrell and Hutt 1976). Long waiting lines and bad service can thus cause a retailer to lose customers permanently.

One cause for the long waiting lines at checkouts is the slow payment process, which prevents the cashiers at checkouts from starting to scan the next customers purchase (Chu and Morrison 2003). The most time-consuming payment process is cash, which to some extent is promoted by retailers through restricted card acceptance and the instatement of minimum amount rules for card payments. A wide-ranging switch from cash to electronic payment could increase the tempo of the payment process and shorten waiting lines in retail stores. In addition, offering customers the opportunity to use the payment instrument of their choice constitutes a service contribution. A study conducted by the German EHI Retail Institute in 2003 and 2004 found that card acceptance significantly influences the shoppers’ decision on where to shop. According to this study, 56.4% of debit and credit cards owners stated that they were influenced by card acceptance when choosing a store. Card acceptance also leads to enhanced spending. 22% of the credit card users stated that they spend more when paying by credit card instead of cash (S-CARD Service 2004).

A further problem for retail customers is the phenomenon of ‘consumer confusion’, a catch phrase for the fact that consumers can be overwhelmed by the range of products that retailers offer in their stores (Rudolph and Schweizer 2004). While the German hard discounter Aldi has only 200 products (Brandes 2004) in each of its stores, a Wal-Mart Supercenter has 142,000 products (Wal-Mart 2007) to choose from. Consumer confusion, which can result from this abundance, might in turn limit the shoppers’ willingness to spend money. This problem could be tackled either by reducing the number of products in the stores, thereby combating the problem at its roots, or by improved assistance for shoppers with their buying decision. Since assistance by store personnel already seems to be too
scarce, mainly because of cost reasons, other means of assistance, e.g. with the help of new technological solutions, could alleviate the problem.

1.2 Research questions

Against this background, this contribution concerns itself with the use of 'Near Field Communication (NFC)' technology in retail to address the before-mentioned issues. NFC stands for the integration of contactless smart card technology into personal devices, such as mobile phones, PDAs and personal computers. The integration of NFC hardware into a personal device enables it to act as a contactless smart card, as well as to read from and write onto such cards. These abilities make it possible for NFC compatible devices to be used as replacement for physical plastic cards (Ecma International 2005; NFC Forum 2006). The possible applications include the use of NFC devices as payment cards, electronic tickets, for the participation in loyalty programs and for the storage of rebate coupons. When compared to contactless smart cards, NFC devices have some advantages which are due to their user interfaces (keyboard and screen), their connection to mobile networks, their large memory and the close connection that people have built up towards their personal devices.

The implementation of NFC-based payment services could help retailers to tackle the problem of long waiting lines at their checkouts by accelerating the payment process. The adoption of NFC-based loyalty applications and rebate coupons could further contribute to this effect. In addition, NFC devices could be used by customers to obtain information on products without the need for assistance by store personnel. This could represent an interesting solution for retailers, since it would reduce the need for assistance by store personnel, which retailers have to pay for, using instead mobile devices, which would be provided by their customers. On the other hand, however, the adoption of NFC would still require significant investments in infrastructure, such as compatible payment terminals.

The standardization of NFC is being driven by the NFC Forum, which was founded by NXP, Sony, and Nokia in 2004. According to its website, the NFC Forum had 146 members, as of November 2008; only 1 (Groupe Casino, France) of these is a retailer. This could indicate that retailers are insufficiently represented in the organization which could in turn lead to an inadequate observance of their expectations in the standardization process. This paper aims to increase the knowledge on the attitude and expectations of the retail industry towards NFC technology. For this purpose, this paper considers the following research question: How do stationary retailers evaluate NFC?

In the following, we present the results of a survey among European retailers. The respondents were, amongst others, asked to evaluate various NFC applications in regard to their ability to increase the speed of the check-out process, improve the shoppers’ convenience, and increase the customers’ loyalty towards retailers. Section 2 provides an overview of our methodology. Section 3 discusses our survey results. The paper closes with a summary of our main findings.

1.3 Related work

The search for scientific publications on Near Field Communication does not lead to a large number of results. The technology has as of yet not received a lot of attention from researchers due to the recency of its emergence. There is however a considerable amount of works on the closely related subject of mobile payment. In their recently published review of literature concerning mobile payment, Dahlberg et al. (2008) classify 73 relevant publications in accordance to the stakeholders and issues they focus on. While 29 publications focus on technological aspects of mobile payments and 20 on the consumer, only five papers focus on mobile payment providers (Kreyer et al. 2003; Vilmos and Karnouskos 2003; Karnouskos 2004; Vilmos and Karnouskos 2004; Zmijewska and Lawrence 2005) and four center on retailer related issues (van der Heijden 2002; Ondrus and Pigneur 2004; Mallat and Tuunanen 2005; Teo et al. 2005). The literature review’s authors state that the number and diversity of mobile payment publications focusing on retailers are disappointing and claim that quantitative studies are needed in order to contribute to a better understanding of merchant adoption. This paper
contributes to filling this gap by surveying the retailers’ evaluation of NFC based applications including mobile payment.

While there are numerous publications on NFC trials and on the potentials of the technology published by proponents of the technology, these sources usually do not meet an academic level of rigour and completeness and therefore suffer from a lack of credibility. Additionally, since their authors have a stake in the success of the technology they might also be regarded as potentially biased.

2 DATA COLLECTION

Based on an extensive review of articles in trade magazines, press releases, and technology white papers on NFC and its applications in the retail industry, a survey was conducted among European retailers between October 2007 and April 2008. The research question was further operationalised, which resulted in the following sub-questions:

• How do retailers evaluate the different possible NFC applications?
• Do some of the applications seem more appealing to retailers than other applications?
• Which changes do retailers expect NFC to bring to their stores?
• Which NFC applications are retailers planning to implement?
• Which payment methods would retailers like to see implemented into NFC devices?

The survey was conducted by means of a questionnaire consisting of five parts (technology adoption, accepted payment methods, payment method preferences and costs, customer services offered today, and customer services planned for the future). It was originally designed in English, but was later translated into German, French, and Italian in order to facilitate the respondents’ participation. The questionnaire was sent to potential respondents by e-mail if possible, by ground mail and fax if requested. Due to feedback concerning the questionnaire format from the first respondents, the questionnaire was later also migrated to a specialized online platform (www.unipark.de) which made the participation more convenient for the further respondents.

In order to generate contacts to be sent the questionnaire we made use of the 2007 version of the Deloitte Global Powers of Retailing report (Deloitte 2007), which is published annually and, among other data, contains a list of the world’s 250 largest retailers. Of the 89 European retailers among the top 250, 6 are not stationary retailers, but operate mail order businesses and were thus excluded from the survey. A further 16 retailers were not contacted due to language barriers. The remaining 67 retailers were contacted by phone. 14 of the 67 companies stated to not participate in any research activities whatsoever, while 12 companies refused to participate in this survey, due to time issues, confidentiality issues or lack of interest in the topic. The remaining 41 companies asked to be sent the questionnaire. After up to two reminders, 16 (39%) of the companies sent the questionnaire back, while 25 (61%) did not answer.

In order to broaden the scope of our explorative study, a further 22 companies were contacted, whose type of business was not covered by the top 250, e.g. an operator of duty free shops at airports and a fast food chain. Of these 22 contacts, 5 refused to participate in the survey due to time issues or a lack of interest in the subject. The remaining 17 contacts were sent the questionnaire. After up to three reminders, 4 (23.5%) of the contacts sent back the questionnaire, while 13 failed to do so (76.5%).

<table>
<thead>
<tr>
<th>European Retailer Ranking</th>
<th>Total Revenue (US$ mil)</th>
<th>Number of Respondents</th>
<th>Respondent Share</th>
<th>Respondents’ Revenues (US$ mil)</th>
<th>Revenue Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 10</td>
<td>518,816</td>
<td>4</td>
<td>40.0%</td>
<td>214,339</td>
<td>41.3%</td>
</tr>
<tr>
<td>Top 25</td>
<td>814,348</td>
<td>7</td>
<td>28.0%</td>
<td>273,220</td>
<td>33.6%</td>
</tr>
<tr>
<td>Top 89</td>
<td>1,178,979</td>
<td>16</td>
<td>18.0%</td>
<td>329,583</td>
<td>28.0%</td>
</tr>
</tbody>
</table>

*Table 1. Statistics concerning questionnaire respondents (based on Deloitte, 2007)*
Overall, 57 questionnaires were sent out to different retailers, and 20 were returned. This represents a response rate of 34.5%. The 20 respondents include 4 of the 10 largest European retailers, 7 of the top 25, and 16 of the 89 included in the mentioned Deloitte ranking (cf. Table 1). While 50% of the respondents’ companies have annual revenues of more than €5 billion, 40% are between €1 billion and €5 billion, and the remaining 10% between €100 million and €1 billion.

As depicted in Figure 1, the number of stores operated by the respondents varies between 1-20 and more than 1,000. A strong majority of the respondents (75%) operate more than 1,000 check-outs. The adoption of a technology which requires new hardware at each check-out would entail significant investments on the part of these retailers. While most of the respondents’ companies operate stores of different formats, supermarkets are the most common among them with a share of 50%. This seems to give those companies that during the interviews showed most interest in NFC solutions which can accelerate the check-out process a disproportionate weight in the survey’s results. However this is attenuated by the fact that 35% of the respondents operate department stores, which on the other hand showed a greater interest in NFC solutions that support the shopper on the store floor (cf. Figure 2).

During the initial phone calls with retailers, it was always attempted to contact the person that is, or that would be responsible for implementing NFC based solutions in the company’s stores. The respondents were all heads of departments or managers of relevant projects. While some of the companies linked NFC technology directly to payment terminals and the departments responsible for these devices, others classify it as a technology related to RFID and the departments driving these technologies. The responding companies have attributed the responsibility for the operation of their payment terminals to different departments. The most common responsible is the IT department followed by the finance department and general store management. Due to NFC’s closeness to RFID, two of the respondents designated their logistics department as closest to NFC. Finally, one of the responding retailers disposes of a dedicated innovation group, which is responsible for all major innovations that are tested and implemented by the company. The head of this department also answered to the questionnaire (cf. Figure 3).
3 RESULTS

3.1 NFC Application Evaluation

The respondents were asked to evaluate five NFC-based applications that are also discussed in the majority of related works on NFC: payment, loyalty applications, rebate coupons, the retrieval of product information, and mobile device based self scanning. Each application was evaluated in regard to six criteria on a five point Likert-type scale. The scale enabled respondents to answer to statements concerning the NFC applications in a spectrum reaching from “strongly disagree” to “strongly agree”. During the evaluation, these answers were replaced by points:

- strongly agree: 5 points
- agree: 4 points
- neither agree nor disagree: 3 points
- disagree: 2 points
- strongly disagree: 1 point

The criteria for the evaluation of the NFC applications were their ability to (a) accelerate the check out process, (b) make shopping more convenient for the customer, (c) increase the user quota when compared to conventional (e.g. barcode based) implementations of the application, (d) reduce the retailers’ costs when compared to current solutions, (e) increase customer loyalty, and lastly (f) the expected consumer acceptance. These criteria were chosen, because they represent the possible influences that seemed most crucial to the interviewed retailers.

In two cases, the questions in regards to the procurement of product information deviated from those of the other NFC applications, due to the facts that it would not constitute a part of the check-out process, and that, according to interview results, it was expected that product information solutions were less established than the other applications. Instead of its effect on the speed of the check-out process, the respondents were asked to judge the potential of NFC based product information to reduce the customer’s need for store personnel assistance. Furthermore, instead of being asked to judge its ability to increase the user quota when compared to current solutions, retailers were asked whether they thought the solution could increase the consumers’ knowledge of products they buy. An overview of the respondents' evaluation of NFC application is given in Figure 4.

Ability to Accelerate the Check-Out and Payment Process

On average, the respondents agreed with the statements that NFC applications could benefit their stores by accelerating the check-out and payment process. In the case of NFC coupon applications, that agreement is considerably weaker than in the cases of payment, loyalty applications, and self scanning applications.

As stated before, the fact that NFC based product information procurement would support the shopper on the store floor instead of becoming part of the check-out process gave reason to deviate from a check-out related question. Instead, the questionnaire asked, whether NFC based product information
procurement was expected to reduce the customers’ need for assistance by store personnel. The respondents were on average unsure whether this would be the case or not.

**Ability to Reduce Costs**

The respondents were asked whether NFC based solutions would help them to reduce payment transaction costs and issuance costs for loyalty cards and coupons. The questionnaire also asked whether NFC device based solutions would constitute cheaper alternatives to current self scanning and kiosk based product information solutions.

On average, the respondents disagree with the statements that NFC based payment, loyalty applications, and coupon solutions are going to reduce their costs of providing the respective services. Self-scanning and the procurement of product information based on NFC enabled mobile devices receive a less negative assessment. However, these two applications also do not seem to represent big cost savers to the respondents.

**Ability to Increase Customer Loyalty**

The respondents answers concerning their opinion on the ability of NFC applications to increase their customers’ loyalty show, that they are on average unsure whether NFC application can have such an effect or not. The average answers differ less than in the previous two questions concerning check-out acceleration and cost reduction.

**Ability to Increase Customer Convenience**

As the results show, the respondents’ belief that NFC solutions could make shopping at their stores more convenient for their customers is slightly weaker than their belief in their ability to accelerate the check-out process. However, with the exception of NFC based product information retrieval, retailers on average tend to agree that NFC solutions could make shopping more convenient.

**Ability to Increase the Shoppers’ Use of Customer Services**

The respondents tend to agree, that NFC based payment and self scanning could increase the number of electronic payments and the use of self-scanning solutions. On the other hand, they tend to disagree, that NFC based loyalty application and coupons could increase the participation ratio of loyalty programs or the use of coupons by customers.

In the case of the NFC based product information procurement application, the questionnaire again deviated from the standard question. Instead, the questionnaire asked, whether NFC based product information procurement was expected to be able to increase the customers’ knowledge of the products they buy. The respondents, on average, tended to agree to the statement. However, the agreement was not very strong.

**Expected Customer Acceptance**

The retailers’ opinions as to the acceptance of the different NFC application by their customers differ only slightly. Retailers are on average neither convinced that NFC based services would be greatly appreciated by their customers, nor that they would be rejected. This clearly contradicts the statements made by companies that have conducted NFC trials. These statements generally speak of enthusiastic consumers and high approval rates concerning NFC technology. This divergence means that either the statements issued by trial organizers are too optimistic, or that the retailers that participated in the survey lack the appropriate enthusiasm, because they have an inaccurate picture of their customers’ preferences. Anyhow, retailers and trial organizers which usually include NFC hardware vendors and service providers clearly have different views on the shoppers’ preferences. If these hardware vendors and service providers wish to make these retailers their customers, they will have to convince them that customers will be fond of NFC.
1.0 2.0 3.0 4.0 5.0

NFC
Product
Info

NFC
Self‐
Scanning

NFC
Coupons

NFC Loyalty

NFC Self-
Scanning

NFC Product
Info

\( \Box \) Customer Acceptance \( \Box \) Increase Use \( \Box \) Convenience

\( \Box \) Increased Loyalty \( \Box \) Cost Reduction \( \Box \) Speed

Figure 4: NFC application evaluation

3.2 Plans in Regards to NFC

When asked which of the listed NFC applications they were currently planning to implement, 40% of the respondents stated that they were currently planning to implement NFC based payment. NFC based self-scanning and coupons applications followed with two respondents each. An NFC based loyalty application and a product information service based on the technology are currently only planned by one respondent each (cf. Figure 5). This significant lead in planned implementations of NFC based payment is rather surprising when the previous evaluations of the different applications are taken into consideration. While NFC based payment received relatively good evaluations throughout all criteria, there was no exceptionally high rating when compared to the other applications.

<table>
<thead>
<tr>
<th>Application</th>
<th>Planned Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFC Loyalty</td>
<td>5%</td>
</tr>
<tr>
<td>NFC Product Info</td>
<td>5%</td>
</tr>
<tr>
<td>NFC Coupons</td>
<td>10%</td>
</tr>
<tr>
<td>NFC Self-Scanning</td>
<td>10%</td>
</tr>
<tr>
<td>NFC Payment</td>
<td>40%</td>
</tr>
</tbody>
</table>

Figure 5: Planned NFC services

A possible explanation for this predominance of payment could however be found in the answers to another question: The respondents were asked to point out, which of the possible NFC applications they considered to be the killer application that would drive the technology’s introduction in the retail industry. A clear majority of 70% think of payment applications as the driving force of NFC in the retail industry. Two of the respondents on the other hand think that loyalty applications will be the
initial driver. One respondent stated that both payment and loyalty applications together would drive the technology’s adoption. Coupon Applications and the support of internal logistics applications were picked by one respondent each, and finally one respondent stated that NFC would be driven into the retail industry by another application, but did not state which. NFC based self-scanning and product information solutions on the other hand were not mentioned by any of the respondents (Figure 6). The role as NFC killer application that is attributed to payment could in part be due to the fact that retailers would only have to upgrade their payment terminals in order to accept contactless and NFC based payments. The remaining infrastructure would be provided by the payment system operators. Loyalty and coupon applications on the other hand would require higher investments from the part of the retailers, because the issuance of loyalty cards and coupons would require them to also invest into infrastructure for the delivery, the processing, and the authentication of loyalty cards and coupons.

![Figure 6: Potential NFC killer application](image)

3.3 Payment Method Preferences

Finally, the questionnaire asked for the payment methods currently accepted by the respondents, their payment method preferences, which payment methods they would like to see implemented as NFC based payment methods, and whether costs to the retailer or convenience for the customer are more important in the decision to accept a new payment method. As Figure 7 illustrates, the only payment method accepted by all respondents is cash. While debit cards and credit cards are also accepted by a large majority of 95% and 85%, only 35% of the respondents accept the stored value cash cards and a private label card operated by their own company. While, 40% of the respondents allow for on account payments, this method of payment is restricted to regular, commercial customers.

![Figure 7: Accepted payment methods](image)

The respondents’ preferences as to the payments methods their customers settle their purchases with are dependent on the amount that the customer has to pay for. Cash is clearly the preferred payment method for payments up to 10 EUR, while debit cards are the preferred method of payment for amounts above 10 EUR. Whilst cash payments decrease in popularity with increasing amounts, all
card payment increase in popularity with increasing amounts. The only exceptions to this rule are the prepaid stored value cards, whose popularity decreases for amounts above 100 EUR (cf. Figure 8).

![Graph showing payment method preferences for different amounts](image)

**Figure 8: Payment method preferences**

When asked which payment methods they would like to see implemented as NFC enabled payments methods, most respondents favored debit cards (80%) before credit cards (75%). As Figure 9 illustrates, the preferences of the respondents as to which of the payment methods they would like to see NFC enabled seems related to the share of respondents that currently accept them based on conventional cards in their stores.

![Graph showing accepted payment methods and favored NFC payment methods](image)

**Figure 9: Accepted payment methods and favored NFC payment methods**

When asked, whether transaction costs or the effects on their customers’ convenience would influence the decision on adopting a new payment method in their stores stronger, most answers stated that both factors were equally influential. As Figure 10 illustrates, the other answers are almost equally distributed over the remaining options. This seems to indicate, that for a new payment method to reach wide adoption the payment method in question has to provide both, a reasonable transaction price to the retailer, and the ability to increase the customers convenience when shopping.
CONCLUSIONS

This paper presented the results of a survey amongst European retailers conducted between October 2007 and April 2008 in order to find out how retailers evaluate NFC. The respondents agreed that NFC based payment had the potential to speed up the payment part of the check out process. They did however attribute this capability to NFC based loyalty applications and self scanning as well. Coupon applications were also attributed the same quality though the rate of agreement was lower in its regard. None of the proposed NFC applications is seen as a potential cost saver by the responding retailers. The respondents seem to expect that the implementation of NFC will rather cause additional costs. This feeling is particularly strong in the case of loyalty applications and payment. The respondents are unsure whether the implementation of NFC applications in their stores will lead to an increase in their customers’ loyalty towards their company. The answers differ only slightly in between the different applications. The responding retailers, on average, agree that NFC based payment constitutes a more convenient way for their customers to settle payments with. However, NFC based self-scanning, loyalty applications and coupons are evaluated to be almost as beneficial to the shopper’s convenience. NFC based product information applications are perceived to be the least beneficial for the customers’ convenience.

The respondents are not convinced that their customers will be enthusiastic about NFC based services. They only expect a moderate acceptance of NFC on the part of their customers. Retailers are, on average, neither convinced that NFC based services would be greatly appreciated by their customers, nor that they would be rejected. This clearly contradicts the statements made by companies that have conducted NFC trials. Retailers and trial organizers which usually include NFC hardware vendors and service providers clearly have different views on the shoppers’ preferences. If these hardware vendors and service providers wish to make these retailers their customers, they will have to convince them that customers will be fond of NFC.

None of the NFC applications clearly stand out as the NFC application yielding the largest to the improvement of the shopping process. While payment is highly ranked among the different applications in regards to its potential to increase the speed at the check-out and to bring additional convenience to the shopping process, other applications are evaluated similarly. The NFC application which the highest number of the respondents is currently planning to implement is payment. 8 respondents (40%) stated that they were currently planning to implement NFC based payment. This number is four times higher than that of the next most current answers. This result was somewhat surprising due to the fact, that payment had not been evaluated significantly more beneficial to retailers than the other applications. A possible explanation for this predominance of payment could however be found in the answers to another question: A clear majority of 70% think of payment applications as the driving force of NFC in the retail industry.

The respondents’ preferences as to the payment instruments that their customers use to pay for their purchases with are clearly dependent on the amount to be settled. While cash is clearly preferred for
the settlement of small amounts, debit cards are the preferred payment method for larger purchases. The majority of the respondents would like to see both debit and credit cards implemented as NFC payment instruments. In order for such NFC-based payment services to be adopted by a large majority of the responding retailers, such services will have to be convenient for shoppers and feature an acceptable price tag.

References

AN EXPLORATORY INVESTIGATION OF CRITICAL SUCCESS FACTORS IN WIRELESS FIELD FORCE AUTOMATION PROJECTS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0636.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Criticial Success Factors (CSFs), Exploratory framework, Mobile systems, Mobility</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
AN EXPLORATORY INVESTIGATION OF CRITICAL SUCCESS FACTORS IN WIRELESS FIELD FORCE AUTOMATION PROJECTS

Khoubeib Djemai, CERAG Laboratory, CNRS – UMR 5820, University of Grenoble, France, khoubeib.djemai@upmf-grenoble.fr
Marc Favier, CERAG Laboratory, CNRS – UMR 5820, University of Grenoble, France, marc.favier@iut2.upmf-grenoble.fr

Abstract:
Wireless Field Force Automation (WFFA) is a wireless solution designed to support field staff such as service personnel, repair or installation employees, technical teams, or medical workers. Certain companies try to provide their field workers with this sort of solution to improve their productivity, their effectiveness and to reduce administrative costs, but more than 50% of these projects fail. This paper discusses the Critical Success Factors (CSF) contributing to the success of these projects. To do this, an exploratory qualitative study was conducted with ten experts in the field. The results of this study show that eight CSF are important in this kind project and that conversely to traditional, sedentary Information Systems, successful WFFA requires taking into consideration a new CSF called “management of mobility” that integrates device management, helpdesk support and mobile device fleet management.

Keywords:
Field Force Automation, field workers, Mobile business, Mobile B2E applications, Critical Success Factors, mobility
Introduction

Advances in mobile technologies, particularly the development of data transmission bandwidths in wireless networks\(^1\), mobile applications\(^2\) (Alag, 2006) and mobile devices\(^3\) have prompted the development of mobile technologies for business\(^4\). Despite their slow growth, when compared to consumer mobile technologies practitioners and researchers predict that business applications will have an enormous impact in the future (Lehmann et al., 2004). The consulting firm Strategy Analytics (2005) predicts that business use of wireless data will grow by 20 percent in North America, Western Europe and Asia/Pacific to surpass $22 billion in 2009.

Mobile business technologies target several diverse populations such as sales force, executive staff, and field workers. This last group represents those individuals that can most benefit from mobile technologies (Barnes, 2004; Rodina et al., 2003). Mobile technologies that support the field force are called Wireless Field Force Automation (WFFA). In 2004, Redman estimated that the number of WFFA users would reach 4 million by 2008 and that the market for wireless FFA would be worth roughly $400M in North America alone. The same 2005 Strategic Analytics study predicts a 16 percent WFFA growth in North America, Western Europe and Asia/Pacific with the market eclipsing $1 billion in 2009. The benefits claimed by WFFA systems are improved field force productivity, reduced costs, increased revenues and improved customer satisfaction (Evans, 2002; Rodina et al., 2003).

Despite the growing popularity of wireless FFA in organizations, the failure rate for these projects is high, above 50% (Bush et al. (2005) cited in Scholl et al., (2007)), and surprisingly few researchers have studied the Critical Success Factors (CSF) in WFFA project implementations. Existing studies concentrate on very specific or limited CSF factors such as lack of buy-in or user implication, inappropriate technological choices (Djemai et al., 2008), lack of training, or poor project management.

The goal of this research is to simultaneously and globally investigate the CSF of WFFA projects by using an exploratory study resulting in a description of the different factors contributing to success.

The contributions of this paper are both theoretical and practical. At the theoretical level, this paper extends classical critical factors (traditionally adapted to sedentary systems) to the mobile context, particularly to WFFA projects, and proposes a new critical factor called “management of mobility”. At the practical level, the CSF, identified in this study, provide useful guidelines for project managers, IT managers, integrators and consultants in order to successfully implement WFFA.

This paper is organized as follows. The first section presents WFFA by defining it and presents those CSF found in the existing Literature. The second section presents our

---

\(^1\) The bandwidth has evolved from 9.6Kbits/s in generation 2 to 171Kbits/s in generation 2.5 and to 2Mbits/s in the third generation.

\(^2\) Mobile application is the software installed in the mobile device. This software has evolved from a disconnected mode to connected “always-on” mode.

\(^3\) Mobile devices have evolved from mass-market devices to rugged-devices, adapted to the job conditions of field workers.

\(^4\) Mobile technologies for business refer to the concept of mobile business or mobile e-business.
research methodology, followed by results and discussion, and the final section presents our conclusions.

2 Wireless Field Force Automation

Terms such as “Fully Mobile Wirelessly Connected” (FMWC) or “Mobile Workforce Management” are used to identify Wireless Field Force Automation (WFFA) and these terms are inter-changeable. There are also several definitions of wireless FFA that vary from author to author (for more details please refer to the appendix). A review of these definitions highlights the different angles from which the subject is approached. Some focus on technological aspects (Gorlenko and Merrick, 2003; Rodina et al., 2003), others on their characteristics (Barnes, 2004; Olofsson and Emborg, 2004) others on their target users (Barnes et al., 2006; Rangone and Renga, 2006; Rodina et al., 2003), and still others on their objectives (Olofsson and Emborg, 2004). We propose the following definition that synthesizes all of these perspectives:

Wireless FFA is a class of information systems functioning in real time applied in order to support and manage (scheduling, supervising and reporting by) a field force, remote from their base of operations and who do not carry out any sales activities.

2.1 Critical Success Factors for WFFA implementations

(Rockart, 1979) defines CSF as “the limited number of areas in which results, if they are satisfactory, will ensure successful competitive performance for an organization”. Unlike the crowded Information Systems research field where much ink has been devoted to investigating CSF (Plant and Willcocks, 2007), little research has concentrated on CSF for WFFA. In order to address this lack we have chosen to analyse the existing literature using Nah and Delgado’s (2006) classification of CSF for ERP systems as a framework for identifying the critical success factors for WFFA projects. Change management, technological factors and communication emerge as being particularly important, change management having been widely studied as compared to the other factors.

Change management: Front and back office staff reductions and changing roles are the principle areas of potential WFFA impact on the employee and they may incite people to resist change (Vas and Vande Verle, 2000; Westelius and Valiente, 2006). According to Vas and Vande Verle (2000), resistance may occur at different levels (managers, technicians, dispatchers, team managers and regional directors) and may be expressed either tacitly (by inertia, covert acceptance) or explicitly (sabotage, argumentation). These results make change management an important factor for success and thus have been studied from both static and dynamic perspectives. Static perspective research tries to identify factors that can facilitate transition. These factors include:

---

3 The synthesis of the body of existing literature relating to CSF in WFFA projects is based on (Nah & Delgado, 2006)’s classification of CSF for ERP systems. They proposed an interesting classification that groups CSF for ERP projects into seven categories: (1) business plan and vision, (2) change management, (3) communication, (4) ERP team composition, skills and compensation, (5) project management, (6) top management support and championship, (7) system analysis, selection and technical implementation.
• Training both field and back office employees is an important factor for managing change (Westelius and Valiente, 2006; Scholl et al., 2007).

• Business Process Reengineering (BPR): Scholl et al. (2007) identified several factors that may inhibit BPR efforts such as: extant process design inadequacy for work in the field (i.e. manager and/or team exclusion in the conception of new work processes), inappropriate compensation schemes and levels in fieldwork, lack of standard operating procedures, and managerial concerns and issues with regard to field operations.

• Getting people’s Buy-in: according to Bush et al. (2005), success or failure depends mostly on the “Buy-in” of field workers. Two factors may lead to this; first, the degree of process change and workers understanding the associated evolution of work processes due to WFFA implementation, and second, the perception of technology enablement as a means to satisfy worker’s individual needs and the resulting advantages for their work.

In contrast to the static perspective, the dynamic perspective responds to questions related to conducting change in WFFA projects. (Vas, 2005) adopted a longitudinal study for evaluating preponderance factors and their evolution before and after the adoption of wireless FFA. He found that organizational context facilitates the implementation of WFFA in the pre-adoption stage, but inhibits its acceptance in the post-adoption stage.

**Technological Factors:** technology can contribute to the success of WFFA projects and neglecting these factors may cause problems such as lack of flexibility, synchronization, and local decision-making authority, and may lead to cumbersome work order management reporting, or ineffective information organization and poor ergonomics. Effective technological choice can explain the failure or success of WFFA projects. Technological choices are studied according to strategic and functional perspectives. From the strategic perspective, choices are made according to their degree of process transformation and modification of existing procedures within the company (Chen et Kamara, 2008; Scholl, 2005). However, from a functional perspective, choices are made according to technical criteria (Djemai et Favier, 2008). Technical choice includes the choice of architecture (data-based or voice-based), software and its package, wireless network and mobile devices.

**Communication:** Scholl et al. (2007) underline the importance of alliances between managers and field workers who are the stakeholders in wireless FFA projects. They point out how their separation leads to frustration and conflicts for both groups. Synthesizing this body of literature, we conclude that it underlines three factors: change management, technological factors and communication. Among these, change management has been widely studied as compared to other factors.

### 3 Research Methodology

The purpose of this study is to investigate the critical factors effecting WFFA system implementation. To address this objective, we adopt a qualitative exploratory approach. Two factors motivate this choice. First, mobile technology is a young technology in its initial phase of development (Lyytinen and Yoo, 2002). Second, there is little empirical research on WFFA (Barnes et al., 2006; Scornavacca et al., 2005).

---

6 (Hammer & Champy, 1993) define BPR as “the fundamental rethinking and radical redesign of business processes to achieve dramatic improvement in critical, contemporary measures of performance, such as cost, quality, service, and speed”.

---

Proceedings ECIS 2009
When these two conditions are fulfilled (new area and little empirical research), interviewing experts is desirable (Evrard et al. 2003). For our study ten WFFA experts were interviewed: one Business Development Manager, one Mobile Solutions Leader, one CEO of an editing company, 4 CEOs of WFFA integrators, and 3 stakeholders in WFFA projects (2 Directors of Operating Methods and one Telecommunications Manager). The interviewees were asked about WFFA technology, evolution, WFFA impact and barriers. The interviews lasted between 40 and 60 minutes. Audio recordings of the interviews were made and the data was later transcribed. Data analysis was carried out using thematic analysis following the suggestions of Miles et al. (2003). We use Nah and Delgado (2006) classification of CSF to codify data.

4 Results and discussion

The Critical Success Factors found in this study are classified into eight themes: action plan and vision, change management, project team composition, project management, top management support and sponsors, communication, technological factors, and management of mobility.

4.1 Action plan and vision

In this factor group it is important to determine the objective and ROI project before launching the WFFA project.

4.1.1 Objective setting

Clearly established objectives are considered a key to success in WFFA projects. The objectives may be quantitative or qualitative and may vary from one firm to another. For one of the companies studied, the objectives are improved productivity and increased technician responsibility. As stated by one Director of Operating Methods “It’s important for us to provide our technicians with mobility tools to avoid returning to the central station and eventually making a round-trip and for the workload.” At another company, the objective is to improve customer satisfaction.

4.1.2 Justification of ROI

It is important to set operational indicators that facilitate ROI calculations. The most popular indicator is an increase in the number of interventions done per day. In the case of Elyo the number of intervention is increased from 8 to 12 per day and the ROI is attained in 6 months. In some cases ROI is difficult to determine, particularly when objectives are qualitative, for example when the objective is to increase customer satisfaction. “As we prepare the specifications for a project, we analyse each operational process to identify the potential gains. In parallel to the technical solution, we build an administrative tool that allows us to access a maximum of information so that we can track ROI indicators. We only finished deploying the system in February of this year, so the feedback is relatively recent, but we already have very interesting data. These are rough indications, but at the same time they serve as benchmarks, dare I say, for comparison amongst our different operations units. They allow us to
measure where we are in relation to our initial desired target. The trend is very
good.” Director of Operating Methods.

4.2 Change management

The majority of respondents argue that managing change in WFFA projects is relevant
because of potential impact on field workers and state that training, business process
reengineering, and involving users in the project enhances successful change
management.

4.2.1 User training

The majority of respondents agree on the importance of training in managing change,
for example one Telecommunications Manager states that: “Everything is
accompanied by training, for instance, when we deploy an application, it’s always
proceeded by a day of training.” In WFFA projects, there are two kinds of training,
familiarisation with a specific mobile device and software application training. The
first is necessary when a field force has paper process experience but has never used
mobile equipment. The later is always necessary. “What kind of training should be
pursued? If you know that you have field workers coming from a paper culture, you
need to not only provide applications training, but also hands on instruction for using
the (mobile) terminal...Typically, with portable telephones there are occasional
breakdowns. How can I correct the problem? This sort of event isn’t a pure
applications issue, so it needs to be dealt with in the training.” Mobile Solution
Leader.

Training quality can be measured by the number of post-training hotline calls coming
from field workers; the better the training, the fewer the calls. “A poorly trained...
person...will call us everyday and will slow us down, (because) they will require extra
help. That’s proof that the training was ineffective, or that there is a problem in
transferring the information... If the training is successful, there won’t be calls to the
hotline...however, (there is always the possibility that) despite the training, the
number of calls increases (and this is a problem)” CEO of WFFA integrator.

4.2.2 Business process reengineering

WFFA significantly impacts existing business processes. Firms must examine and
adapt their processes in relation to this technology, otherwise the project will fail.
"Repeating existing procedures doesn’t work, changes must be accepted." clarifies one
Business Development Manager. Rewriting or creating processes can also facilitate the
development of specifications lists, as mentioned by one Director of Operating
Methods “When we rewrote our processes we also wrote the specifications list...
100% of our processes were revisited”.

4.2.3 Implication of users

As proved by previous research, user implication is highly relevant to WFFA project
success. “In the appropriation of tools, if we don’t listen (to workers) they reject the
tool, and if they reject it the system falls apart.” explain one Director of Operating
Methods. In addition, these users must be integrated at all stages of a WFFA project, as mentioned by another Director of Operating Methods “The users were integrated in absolutely all stages of the project, even before the specifications list was developed. The technicians were involved in the choice of the final solution after they tested two operational prototypes. These (prototypes) had been field tested for three weeks under real work conditions. The choice was made based on feedback from these experiences. The impact of this test period was huge and gave the project enormous visibility. It really excited the workers who ended up using the system.”

4.3 Communication

The majority of participants emphasize the importance of communication in the success of WFFA. Several classical tools such as internal supports or internal newspapers, meetings…may be used to implement an effective communication but the creation of a network of representative (persons designated in each operational entity) who build a good internal communication appear to be the most efficient tools.

4.4 Team composition

The roles, processes and responsibilities of field workers are determined by Operating Methods Direction. Thus when a WFFA project is undertaken (a project that will strongly impact field work), the project team needs to include representatives from both Information Systems and the Operating Methods Division. Levels of responsibility and involvement vary depending on the phase of the project. At the beginning of a project the two teams interact, but once deployment is finished, the Operations team is left in charge of follow-up and the evolution of the applications.

4.5 Project management

Deployment methodology and financial guarantees are the two project management factors stressed by the participants.

4.5.1 Methodology of deployment

All participants emphasize the importance of deployment methodology in the success of a WFFA project. Globally, they stress five phases:

1. Buy or develop: in this phase the enterprise must choose between an off-the-shelf or a tailor-made solution. With off-the-shelf solutions, the company buys a pre-existing mass marketed application. A tailor-made solution allows the company to develop a specific application that fits their field worker’s tasks. According to Chen et Kamara (2008) these alternatives fall under the category of mobility strategy.

2. Integration or specification: This selection is driven by the “buy” or “develop” choice. In the first case (buying off-the-shelf software) the company opts for gap analysis software to integrate a non-custom-made solution into their existing Information Systems. In the second custom software case, they opt for functional specification software.

3. Pilot phase: This generally takes from two to four weeks. The objective is to provide the solution to a selected group of users who test it and verify its usability.

4. Pre-deployment phase: This generally lasts anywhere from several weeks to a year. At this stage, it is important to select the appropriate population that will test the solution. The goal is to validate all of the WFFA solution components before deployment to the global population.
5. Final deployment: This phase entails global deployment of the solution. A progressive approach is important in order to identify errors or defects and to correct them. As suggested by one participant, “progressive deployment will detect a problem with ten terminals, will redetect new things when we go to fifty terminals, will redetect new hazards at 500 terminals and at a thousand. This is an approach that increases the volume progressively ... while mastering (the system) and reclassifying the process.” CEO of WFFA integrator.

4.5.2 Financial guaranties
Like Information Systems, mobility projects run the risk that estimated budgets may be exceeded and projects will be abandoned. Guaranteeing that the budget estimated in the early stages of the project will not be exceeded is critical to the success of WFFA projects. “(It is vital in order to) assure the operational life of the project because that’s where there’s a major rejection risk. In addition, there are market zones that allow us to be sure that things won’t change economically. We are committed to ... maintaining a fixed price.” CEO of WFFA integrator.

4.6 Top management support and sponsors
Top management implication is also necessary in WFFA projects. “(This aspect) is fundamental because it is truly a project of managing change” explains one Director Operating Methods. The importance of this implication is proportional to the size of the future user group, especially when there are more than 50 users, and when the WFFA project affects the firm’s core business.

Technological sponsors also play a crucial role in the success of WFFA projects especially when there is a large population of field workers that need to be federated. “For example, team leaders help motivate the others (concerning the project) and can be helpful (in passing on messages from top management)” explains a Mobile Solution Leader.

4.7 Technological factors
Effective choice of technological components is considered as the most important aspect in a WFFA project. These choices concern system architecture, software, middleware, mobile devices and telecommunications network (for detailed discussion about these choices and decision criteria please refer to Djemai et Favier (2008) study). “When we put together a project like this, we decide to use a certain number of different components and even if we think that everything is moving in the right direction, it’s important, when choosing middleware and applications, to choose, for example, the right terminal and not make a mistake. After that we have to choose the appropriate telecommunications operator, the correct transmission speed and why, etc...” Mobile Solutions Leader. This aspect can, however, be problematic, as one participant states, “The problems we most often encounter are technical...otherwise stated, the rejections (on the part of the field workers) that we experience are linked to the computing process” Mobile Solutions Leader. This may be explained by the fact that the technology is still maturing.

Owing to the heterogeneity of technology in WFFA solutions and differing development standards, successful communication between purchased components is...
vital. Integration management is an alternative that facilitates and optimises the choice and integration of technological components. “We define each offer in regard to the (existing/available components), they are totally interdependent” CEO of WFFA integrators. In addition, integration management assigns responsibility for and guarantees the correct functioning of the chosen solution(s). If this role is not fulfilled, it is difficult, if not impossible, to identify the origin of function errors and the responsible party.

4.8 Management of mobility

This emergent and increasingly important factor is the result of the growing mobility of field workers who spend the majority of their time outside of their organization. They work alone and remotely from their base of operations, making receiving support difficult if a problem should emerge in the field. This problem does not arise for sedentary workers or workers operating in a local mobility context as they can get support at any time from their colleagues or from their Information System Department. As explained by one CEO of WFFA integrators: “We are working with people who... are all alone (in the field). With a traditional IS, when a problem arises the head of IS is available (to help). (In a remote context) the technician is alone and faced with a malfunctioning machine. He doesn’t have the time to determine what works and what doesn’t, he must respond quickly and make it work...he’s got to figure out the solution by himself.” Several services exist to remotely support field workers and to manage problems. Helpdesk support, device management and applications support are critical factors that are proper to the mobility context and deserve further examination, as follows.

4.8.1 Mobile device management

When mobile devices break down, field workers cannot receive information related to their work and as a result are unable to do their jobs. Quickly providing the fieldworker with functional replacement mobile devices is critical in this kind of project. This activity is called “mobile device management” and integrates several other services like “managing breakdowns, returns, transport and forwarding, (and) management of stock security” according to the CEO of an editing company. These activities can be accomplished by creating an internal maintenance and repair service for mobile devices, or by externalizing the function to integrators who take charge of device management. “We get calls until 16h (four o’clock) in the afternoon. The next morning, before 9 am, the on-site user should have a terminal that is in the final phase of configuration and is being synchronized to pick up from the moment the system broke down. We try to limit this kind of correction, preferring rather to make an “anticipated exchange” containing advanced specialized software that allows the user to avoid breakdown in the event that there is a technical problem... (and this reduces the down time to) virtually zero” CEO of WFFA integrators.
4.8.2 Helpdesk support

Corporations must provide helpdesk support for their field force users on those occasions when they encounter use, login or start-up problems. As one respondent stated, “if the (field worker) can’t enter data into an application, he can call the hotline”. The participants agree that 70 to 80 % of the problems encountered in the field consist of workers who have pushed the wrong keys, mishandled their equipment, or who encounter login or start up difficulties. These problems can generally be resolved remotely in a few minutes by a specialist in the back office.

Two kinds of helpdesk exist, a start-up helpdesk and a recurrent helpdesk. The first supports users for six months following system deployment. The later provides ongoing support for users and is particularly helpful during the release and installation of new software. In addition, the helpdesk can anticipate and react to lingering problems in the field and warn mobile device service management about their existence.

4.8.3 Mobile device fleet management support

When a company wants to install a new application or release a new software version, historically, they needed to physically recover all the devices concerned and install the upgrades. This was very difficult and costly with field force workers in far-flung regions or in several countries. In addition, bringing mobile devices to a central location implied that field force work halted for the several days required to collect the devices from their field locations.

This problem is addressed by device fleet management support or tele-distribution7. “It’s imperative to create new applications for deployment that are capable of verifying what’s happening on a terminal (and to enable the installation or upgrade of existing software). For us this is the biggest challenge for the mid and long term, and it’s extremely important.” explains a Mobile Solutions Leader. This kind of service is relevant to the success of WFFA projects, especially when the company has equipped a large number of field workers with WFFA and this field force is geographically dispersed. “Device management is a critical factor of project success if you have hundreds of regional or international collaborators” Mobile Solutions Leader.

5 Conclusion

In our research we explore critical factors for success in WFFA projects. The eight most prominent keys to success that we identify in WFFA projects are action plan and vision, change management, project management, top management support and sponsors, communication, technological factors, and mobility factors.

This paper reveals that, in addition to classical CSF outlined by Nah and Delgado (2006), “mobility factors” are also critical to WFFA project success due to the dispersed localization of field workers. Mobility factors integrate three sub-elements; helpdesk, mobile device management, and mobile administration. These sub-factors

---

7 Tele-distribution is a kind of middleware (a sort of software) that acts as a communications and command center for multiple applications, content types, and devices (Sabat, 2002).
can be integrated either separately or simultaneously by a company wishing to undertake a WFFA project. The paper provides useful guidelines for IT consultants, firm IT managers and project managers in the implementation of a WFFA system. If the factors presented in this study are taken into account, the likelihood of a successful implementation is greatly improved.

As with any research, this paper has certain limits that can serve as an starting point for future research:
- The CSF in WFFA are studied without consideration to company size or sector. A comparative study of projects carried out in companies of differing sizes and in different sectors is an interesting avenue for further research.
- In this paper the factors aren’t classified by degree of importance, however, this is an interesting subject for further research. Quantitative research is necessary in this type of investigation.
- This study has not explored the pre-adoptions or post-adoptions pattern of CSF importance or shifts in importance between these stages. (Vas, 2005) has undertaken this sort of analysis, but only for change management factors. Thus a longitudinal investigation of the evolution of all factors identified in this study is an interesting area for future research.

References


Appendix

Some definitions of WFFA proposed in the existing literature:

• FFA is the association between software and hardware, which are used by a specific number of staff, who are responsible to build relationships with customers, maintain expensive corporate and government assets and gather critical data that keep organization running (Rodina et al., 2003).

• FFA is a generic term for mobile applications used in real-time support orders, scheduling, supervising and reporting in the field (Olofsson and Emborg, 2004).

• Wireless FFA is usually used to describe mobile employees, remote from their base of operations, utilizing wireless technologies to perform their specific business tasks (Barnes et al., 2006).

• FFA applications are used by employees such as field service teams, technical teams, or more generally, staff that do not carry out any sales activity (Rangone and Renga, 2006).

• Compared to traditional FFA applications, Wireless FFA offers a substantially increased scope for benefits – particularly through real-time, location independent network connectivity (Barnes, 2004).

• Fully Mobile Wireless Connected (FMWC) is defined as Information and Communication Technologies are assumed to have high potential for improving field operations Gorlenko, and Merrick, 2003) cited in (Liu et al., 2007)).
REQUIREMENTS ON IT BUSINESS VALUE MEASURES FOR MOBILE-INTEGRATED BUSINESS PROCESSES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0141.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Mobility, Business value of IT, Business Process Improvement, Multiple case studies</td>
</tr>
</tbody>
</table>
REQUIREMENTS ON IT BUSINESS VALUE MEASURES FOR MOBILE-INTEGRATED BUSINESS PROCESSES

Kai Habermann, wi-mobile Research Group, Business Informatics and Systems Engineering, University of Augsburg, 86135 Augsburg, Germany, kai.habermann@wi-mobile.de

Key Pousttchi, wi-mobile Research Group, Business Informatics and Systems Engineering, University of Augsburg, 86135 Augsburg, Germany, key.pousttchi@wi-mobile.de

Abstract

The IT productivity paradox raised the issue of finding appropriate measures for capturing the business value of information technology. While the topic is much discussed on the general IT level, there is little research on the specific investments in innovative mobile and ubiquitous technologies. In this paper, we apply a case study research methodology in order to identify the specific characteristics of mobile and ubiquitous technology investments and the subsequent requirements on appropriate measures for the IT business value of mobile-integrated business processes. The results are discussed with respect to the existing propositions from general IT business value frameworks. The outcome is a set of requirements on IT business value measures tailored to the specific characteristics of mobile-integrated business processes. We see the result also as a first step in the development of a suitable measurement framework which could be applied by both researchers and practitioners, especially ex ante in mobile business process reengineering.

Keywords: Mobile-Integrated Business Processes, IT Business Value, Performance Measurement
1 INTRODUCTION

While in the past the adoption of mobile and ubiquitous technologies has mostly followed a technology-driven approach (Köhler and Gruhn, 2004), organizations increasingly expect measurable business returns from their mobile technology investments (Basole, 2004). In order to demonstrate the returns that the mobile technology delivers to business, companies are required to define appropriate measures to capture the realized business value outcomes. While the topic of IT impact on organizational performance has been extensively discussed in literature in connection with the *IT productivity paradox* (e.g. Brynjolfsson 1993; Brynjolfsson and Hitt, 1994; Dewan and Kraemer 1998), little research has been conducted on appropriate measures for capturing the specific returns of mobile and ubiquitous technology in business process reengineering (Van der Heiden and Valiente, 2002).

The research objective of this paper is to contribute to theory building in the area of *mobile-integrated business processes (MIBP)* by deriving a set of requirements on appropriate business value measures for MIBP. For this, we analyse the findings from four case studies. In order to relate the business value measures to the specifics of MIBP, we explore typical characteristics of MIBP investments and derive the requirements in relation to each characteristic. The outcome is a set of requirements that need to be considered when assessing the business value of MIBP.

In the context of this paper, the term *mobile business processes (MBP)* refers to any business process which is partly or completely mobile and thus cannot be fully supported by the use of stationary IT (Gumpp and Pousttchi, 2005). The term *mobile-integrated business processes (MIBP)* refers to any MBP that is fully supported by mobile or ubiquitous IT (Pousttchi and Thurnher, 2005). Typical examples of MIBP in this respect are mobile service, mobile sales and logistic processes where applications on mobile devices, RFID-based systems, dispatching software, backend systems or whatever is needed for full support are all seamlessly integrated and allowing in the best case a bidirectional real-time information exchange from any part of the process.

The paper is structured as follows: In section 2 we examine the existing literature on IT business value measures and MIBP. In section 3 we describe our research approach and the context of the applied case studies. In section 4 we then identify the specific MIBP characteristics based on the findings from the case studies, derive the requirements on the MIBP business value measures and compare them to propositions from general IT business value models. In section 5 we draw conclusions and discuss the limitations.

2 LITERATURE

2.1 Research on IT business value measures

The need for appropriate measures of IT business value has been a major topic in IS literature for almost twenty years. The topic received particular attention in conjunction with the debate on the IT productivity paradox which claimed that a direct relation of IT investments and productivity outcomes cannot be empirically shown (Brynjolfsson, 1993). While the original paradox actually addresses the macroeconomic level, it also initiated the IT business value research on the organizational level (e.g. Mahmoud, 1993; Barua et al., 1995). Most studies on the organizational level are able to prove a positive correlation of IT investments and organizational performance, but show at the same time a significant variance over the complete sample (Brynjolfsson and Hitt, 1994; Mukhopadhyay et al., 1995).

The process models of IT value explore this effect by looking at different levels of business value (Soh and Markus, 1995; Weill and Broadbent, 1998). They conclude that only through internal user
adoption, the application to business processes and a positive external environment, the IT expenditures will ultimately lead to positive organizational performance (Soh and Markus, 1995, Kohli and Hoadly, 2006). Based on these findings, they suggest assessing business value outcomes at each level separately with a particular focus on the intermediate level (Seddon et al., 2002). Measures at the intermediate, as opposed to firm-level measures, have the advantage to provide insight on how value is created (Davern and Kauffman, 2000) and they are less affected by external factors such as competitor actions or general market developments (Kohli and Hoadly, 2006).

In contrast to the models that focus on the organizational level of measurement, Mooney, Gurbaxani and Kraemer look at the degree of transformation that is created (Mooney et al., 1996). They distinguish three dimensions of IT business value based on the degree of process change: automation, information and transformation. Based on this distinction, they define different measures and demonstrate how business value increases as IT permeates the organization. Their approach basically follows and summarizes an earlier model from Venkatraman who distinguished seven different categories of potential benefits from IT-enabled business transformation (Venkatraman, 1994). Two other central aspects with respect to business value measures are explored by Barua, Lee and Whinston (Barua et al., 1996). First, the IT impact can typically not been measured in isolation and, second, there may be dependencies and conflicting business value impact. They maintain that without the understanding of the complementarity, of the (inter-)dependencies and the conflicts, wrong measures may be applied or business value outcomes may be wrongly measured.

IT business value measurement frameworks have also been developed in the corporate environment. One of the most recognized ones is published by Svard and refers to the value dials concept established through Intel's IT Business Value program (Svard, 2006). In this program, over 350 IT projects were reviewed over a three year period from 2002 to 2004. The result is a set of 18 general measures for IT business value in four categories: headcount management, expense avoidance, revenue increase and working capital. Especially notable is also Svard's definition of business value which is explicitly requesting to represent IT business value in monetary terms. He refers to business value as the benefit for business groups represented in dollar terms which are a result of IT solutions or services. Academic papers in turn define business value typically more general as the contribution of IT to firm performance (Tallon et al., 2000).

While these general requirements on IT business value measures certainly include important aspects that are also valid for MIBP, they, however, do not consider the specific characteristics of MIBP investments and its consequences on particular requirements for MIBP.

### 2.2 Research on IT business value of MIBP

Dedicated research on the specific characteristics of MIBP investments and IT business value measures is rare up to now. As with static IT, the business value of implementing mobile and ubiquitous technologies can typically not be assessed in isolation but only in conjunction with respective process improvements (Van der Heiden and Valiente, 2002). The MIBP literature mainly describes two broad beneficial effects of mobile technology: (1) improved process efficiency, and (2) improved process effectiveness (e.g. Gebauer and Shaw, 2004). Also convenience is suggested as a third potential benefit (Basole, 2004). Most existing papers derive exemplary MIBP benefits based on observations in certain industries, for certain process types or in relation to specific MIBP capabilities. Examples of benefits in specific industries can be found for the building and construction industry (Löfgren, 2006; Gump et al., 2005), healthcare (Scheepers and McKay, 2004), utilities (Nah et al., 2005), intervention forces (Bazijanec and Pousttchi, 2004) and the financial services sector (Hastreiter, 2006). Benefits of specific MIBP types have been explored, e.g., for dispatching processes (Botzenhardt and Pousttchi, 2008). Effects of specific MIBP capabilities like real-time order assignment have been analysed by Habermann (Habermann, 2005). However, it can be contended that none of the studies provides an overarching business value framework and dedicated requirements on business value measures for MIBP.
Another valuable source for examples of achieved business value may also be the various whitepapers from mobile software vendors (e.g. Aventeon, 2008, Cognito, 2008). These whitepapers certainly provide a deep insight in the process improvements and achieved business values, but naturally tend to overestimate business value potentials and underestimate the drawbacks (Van der Heiden and Valiente, 2002).

Specific research on MIBP assessment methodologies is rare up to now. Deibert and Rothlauf (Deibert and Rothlauf, 2006) suggest a set of quantitative indicators (e.g., error rate or stock size) and qualitative indicators (e.g., process and stakeholder goals) for measuring the process performance. However, these indicators remain at the level of process improvement and do not relate to any measurable business value impact. Earlier, Gumpp and Pousttchi (Gumpp and Pousttchi, 2005) provided a structural framework for the implementation of mobile and ubiquitous technologies in business processes and the assessment of its impacts (Mobility-M framework). However, the benefits assessment in this framework is relying on informational added values and thus also lacking concrete measures to capture the achieved business values.

3 METHOD

3.1 Case study research

For our analysis, we apply a case study approach (Eisenhardt, 1989; Yin, 1994). We consider case study research as especially appropriate in this situation as we are aiming to contribute to theory building in a new research area. The case study research allows us to build our requirements on existing practice and leverage the expert knowledge of experienced practitioners (Benpasat et al., 1987).

Following Eisenhardt's process steps for building theory from case study research (Eisenhardt, 1989), we defined the initial research question and selected a group of organisations that seemed suitable to contribute to the research question. In order to achieve a broad understanding of the topic, we selected four organisations for which their M(I)BP is at the heart of their business and which show a high level of maturity with respect to business value assessment. This certainly does not lead to a representative result, but provides insight in advanced practices with regard to the research question. To be most diverse, we selected the organisations from different industries, countries and company size. For data collection, we applied a combination of different methods (Yin, 1994). We reviewed internal documentation, company presentations and conducted semi-structured personal interviews, both face-to-face and telephone-based. In each organization, we spoke in person to at least one person from the IT and the business side who have been involved in the mobile technology implementation phase.

3.2 Applied Case Studies

Case Study 1: Utility Company, Ireland

For the first case study we investigated a major Irish Utility Company in the area of network management and metering. In three phases, over the last 4 years, more than 1,000 technicians have been equipped with a mobile solution. Typical tasks of the mobile technicians are the maintenance of network components and services around meter reading and exchange. Compared to the previous decentralized planning, the dispatching of work orders is now handled through a central dispatching system, which also results in significant organizational changes. The central dispatching system also proposes a preferred task order based on criteria such as priority, specialization of the technicians, and availability of required equipment.
Case Study 2: Logistics company, Netherlands

For the second case study we investigated a Dutch logistics company with an own fleet of 300 and an additional 450 associated trucks. All own trucks have been equipped with ruggedized PDAs. The associated trucks had to commit themselves to use the same mobile software, so that the dispatching can be uniquely controlled through a central tour planning solution. The software on the mobile device provides an order management application, barcode scanning and integrated navigation module. Initially, the mobile solution was implemented with the goal to reduce the vehicle cost through saved kilometres and toll costs, but the efficient real-time management of new orders, and especially the provision of information problem reports on an integrated webpage, have also proven to be a differentiator in the market and resulted in an extended customer base for the logistics company.

Case Study 3: Airport operating company, Europe

For the third case study we investigated the service unit of a company in the airport operating sector. The company provides both logistics and facility maintenance services. In total, about 30 processes have already been mobile-integrated or being currently rolled out. More than 1000 mobile workplaces are impacted. The MIBPs include the maintenance of equipment, technical asset control as well as a series of processes around loading and unloading of the aircrafts. Many technical assets have also been equipped with RFID tags to ensure efficient data exchange. While the first processes were mobilized mainly to comply with legal requirements, mobile integration has increasingly become a source of business value for the company. The mobile process engineering does not only allow for real-time coordination and seamless data transfer but also for the accurate recording of start and end-time of orders which is basic for an accurate service charging to the airlines.

Case Study 4: Sales company, Germany and Austria

For the fourth case study, we investigated a specialized sales company that is offering sales and merchandizing services for major product companies especially in the cosmetics industry. The company has about 200 mobile sales consultants equipped with a scan-enabled mobile device. Main intention of the mobile solution besides the improved coordination of customer visits is the creation of a central market intelligence system. The documented data from the field will be analysed to improve the sales process, but also the optimal product positioning and best display method. The generated knowledge will be provided as extra service. Also a new consulting arm for innovative display methods is planned, where the company can run specific campaigns and document the results directly on the mobile device including photos and recorded testimonials from the shop owners.

4 REQUIREMENTS ON MIBP BUSINESS VALUE MEASURES

In this section, we derive the requirements on business value measures of MIBP based on four specific MIBP characteristics that have been rated by both the interviewed experts from the business and IT side as major impact factors on the definition of their MIBP business value measures.

• Close relation of MIBP investments to the business side
• Dependency of MIBP impact on organizational changes
• Significance of mobile employee independence
• Consideration of the innovative nature of MIBP

In the following, we describe each characteristic and derive the associated requirements on the business value measures. To raise the theoretical level, we also compare each requirement with propositions from existing IT business value literature. The outcome is a set of six requirements that should be considered when defining a suitable measurement framework for business value of MIBP (Figure 1).
Specific characteristics of MIBP from case studies

Requirements on MIBP business value measures

Figure 1. Analysis model for deriving the requirements on MIBP business value measures

4.1 Close relation of MIBP investments to the business side

The first characteristic states the close relation of MIBP investments to the business side. In all regarded cases, the investments are to a large extent sponsored from budgets outside of the IT department. Business managers are also over proportionally involved in the MIBP planning, implementation and control processes. As a consequence of their sponsorship, the business side normally associates clear expectations on the concrete technology contribution to defined business objectives. Typical business objectives in that sense are, for example, cost savings for central coordination, reduction of rework efforts or an increase of mobile work productivity. In some cases the business objective is even more concrete in the sense that it determines a concrete value target like 10 % cost savings or achieving 100 Euro of additional revenue per customer. To prove that the technology can really fulfil its expectations, the business value measures need to be set up in direct relation to the concrete objective for which the technology is deployed. Concerning the business objectives also the expected time of the business value realization is relevant. The business side typically expects almost immediate results from the MIBP implementations. This is clearly another major impact factor on the definition of appropriate business value measures and the time horizon for actual business value measurement.

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Sponsorship and project lead</th>
<th>Business objectives and expected realization time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility Company</td>
<td>All direct project costs sponsored from business side. Project lead by business side with IT support.</td>
<td>Improving dispatching capabilities with expected savings of at least 300k Euro per year. Reintegration of outsourced services saving a min. of 100k Euro. Realization of business value for back office expected for first year, total project payback within 24 to 30 months.</td>
</tr>
<tr>
<td>Logistics Company</td>
<td>Specific MIBP project outside normal IT budget. Project lead by IT side, high involvement of drivers during design phase.</td>
<td>Reducing kilometres by 5%, resulting in vehicle cost savings of 800 Euro per truck per year. Not planned, but achieved has been an extent ion of the customer base and orders by 18%.&quot;Immediate&quot; impact of mobile technology investment expected, payback within 18 month.</td>
</tr>
<tr>
<td>Airport Operating Company</td>
<td>Partly sponsored from business side. Infrastructure updates within IT budget. Project lead by IT with close relation to business side.</td>
<td>Fulfil legal requirements, reduce overall process expenses by defined percentage, improve service level to airlines to justify agreed price increase. Single projects embedded in MIBP, concrete payback horizon based on strategic plan.</td>
</tr>
<tr>
<td>Sales Company</td>
<td>Specific MIBP project outside normal IT budget. Dedicated project team from business and IT side.</td>
<td>Increasing sales by 20% through optimized product selection and improved product display Realization of business value within two years, also depending on implementation of business intelligence system</td>
</tr>
</tbody>
</table>

Table 1. Overview of case study findings in relation to the first MIBP characteristic

Based on the close relation of MIBP investments to the business side, we derive the first two requirements on MIBP business value measures.

Requirement 1: MIBP business value measures should be defined in relation to the business objectives for which the technology is deployed

Proceedings ECIS 2009
Requirement 2: MIBP business value measures should be considering the time horizon of the expected realization

Within the IT business value literature, the relation of measures to concrete goals is a central demand for realizing business value outcomes. Kohli and Hoadly state that the technology investments are typically initiated in the context of certain strategic objectives. Measures should therefore be defined in close relation to the strategic goals in order to show if the original goal has been achieved to the extent expected (Tallon et al., 2000; Kohli and Hoadly, 2006). Hitt and Brynjolfsson contend that the focus of an IT initiative, i.e. productivity, customer value or profitability, determines which measures are attended to (Hitt and Brynjolfsson, 1996). Regarding the measurement of business value outcomes, the literature typically mentions that it has to happen over a longer period of time (Wagner, 2003), especially in dynamic environments (Chan, 2000). The specific requirement of a short time horizon is mentioned by Svard, particularly in situations where there is also high pressure on the business side to deliver fast results (Svard, 2006).

4.2 Dependency of MIBP impact on organizational changes

The second MIBP characteristic derives from the awareness that the business value of MIBP can only be realized when companies also possess the flexibility to incorporate the required organizational changes. Only in very few cases, for example the logistics company's vehicle cost reductions, the mobile technology implementation leads to immediate business value effects. In most case, especially when the MIBP is intended to transform the existing process, the organisational changes are condition to the realization of business value outcomes. The utility company for example realized business value only after completely restructuring the dispatching function from a decentralized to a centralized structure. The business value measures consequently need to consider this complementarities. They also need to comprise the complete effect that is generated through the technology and the organizational changes which in turn means that also all related costs, including expenses for process reengineering and organizational changes, need to be considered in ROI calculations. Clear aim of the companies is to ensure traceability and credibility of the business value assessment. Therefore they typically define their business value measures at the internal level. This avoids the dilution of the results through external effects, which often cannot be influenced.

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Required organizational changes for value realization</th>
<th>Organizational level of MIBP business value measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility Company</td>
<td>Major organizational changes: Centralization of dispatching function across all maintenance and repair processes, option for mobile employees to retrieve additional work orders from central work order pool</td>
<td>Process level, in relation to business case assumptions. Considering BPR and change management cost in ROI calculation.</td>
</tr>
<tr>
<td>Logistics Company</td>
<td>No major changes in process flow, business value mainly from real-time dispatching</td>
<td>Direct measurement based on kilometre account per truck</td>
</tr>
<tr>
<td>Airport Operating Company</td>
<td>Major organizational changes, including repositioning of work staff, relocation of responsibilities to mobile employee</td>
<td>Measures at process performance level (process time, error rate) and recalculation in monetary terms</td>
</tr>
<tr>
<td>Sales Company</td>
<td>Job enrichment for mobile employees through extension of work tasks in combined sales and merchandising function</td>
<td>Measuring the weekly product sales depending on new product assortments and re-positioning in the stores</td>
</tr>
</tbody>
</table>

Table 2. Overview of case study findings in relation to the second MIBP characteristic

Based on these cognitions, we derive the third requirement on MIBP business value measures.

Requirement 3: MIBP business value measures should be defined at the internal organizational level, assessing the complete business effect of the MIBP

In the IT business value literature, the complementary of IT investments with organizational changes is central aspect in all process and BPR-related models. Kohli and Hoadly for example state that the IT
effect in IT-enabled BPRs is typically not to measure in isolation (Kohli and Hoadly, 2006). The IT implementation often creates business value only in combination with coordinated structural changes (Brynjolfsson, 1993) and must be changed in a coordinated manner to improve business performance (Barua et al., 1996). For the business value measurement, the process models typically suggest to set up separate measures at each organization level (Soh and Markus, 1995; Weill and Broadbent, 1998), however it is widely suggested to use the intermediate level, i.e. the internal process or organizational level, for assessing business value outcomes (e.g. Kohli and Hoadly, 2006).

4.3 Importance of mobile employee independence

As the cases show, it is crucial to respect the fact that the mobile employees are often self-employed or at least a high share of their income depends on performance. So it should be in their natural interest to improve their performance through mobile technology support. However, in reality we can also see many cases where the productivity of the mobile employee is even decreasing. The reason is to be found in the on-site work process which is sometimes loaded with additional verification and documentation needs resulting in significant business value gains at the back office level. To reach acceptance in these cases, a potential negative business value for the mobile employees needs to be considered in their performance schemes. This necessarily requires the separate measurement of business value outcomes for the mobile work process and the back office. The separate measurement becomes even more crucial in cases where the mobile process is outsourced. If the additional effort is not reflected in the price per work task, the external service provider is not likely to adopt the new technology, thus the associated business value impact at the back office is inhibited. In an outsourcing scenario even a small productivity increase for the mobile work process may often not be enough as the expenses for mobile devices and process changes are often to be carried by the external service provider. A distinct assessment of the added value is also necessary for business partners or customers. It is for example contingent for setting accurate prices for new service offerings. Furthermore, from the experiences in the case studies, it can be concluded that the measures should be agreed upon by all stakeholders to ensure that objectives are well understood by all parties involved and the measurements receive broad long-term support.

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Degree of organizational independence of mobile work force</th>
<th>Agreement process on measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility Company</td>
<td>70% of work orders currently handled by internal employees (25% performance bonus), 30% outsourced</td>
<td>Internal work force involved from the beginning of the project. Measured lower productivity in some areas.</td>
</tr>
<tr>
<td>Logistics Company</td>
<td>45% of all tours through own fleet, 55% outsourced. External drivers paid based on structure of tour.</td>
<td>Measures mainly provided from business side, aligned with major external transport organizations.</td>
</tr>
<tr>
<td>Airport Operating Company</td>
<td>Central solution provided to different business departments. Outsourcing in certain areas, e.g. cleaning.</td>
<td>Measures aligned within project team, service level measures defined in coordination with airlines.</td>
</tr>
<tr>
<td>Sales Company</td>
<td>Mobile sales personnel typically self-employed, but strongly connected to company. Average performance bonus between 20% and 40%.</td>
<td>Measures defined within project team and in accordance with mobile sales personnel. Extra payment scheme for additional work tasks that are supported by the new MIBP.</td>
</tr>
</tbody>
</table>

Table 3. Overview of case study findings in relation to the third MIBP characteristic

Based on these finding, we derive the next two requirements for MIBP business value measures.

Requirement 4: MIBP business value measures need to assess the outcomes separately for different stakeholders in the process

Requirement 5: MIBP business value measures need to be agreed with all stakeholders prior to the invest to ensure later acceptance
The separate business value measurement for different stakeholders is not a typical requirement in IT value frameworks. It is especially important in MIBP due to the high share of self-employment and outsourced work tasks. The stakeholder agreement in turn is a central topic in various concepts (e.g. Svard, 2006; Ward and Daniel, 2003). For implementation of the stakeholder agreement prior to the investment, Ward and Daniel for instance suggest to create a dedicated benefits realization plan that a priori determines the business value measures and also defines the measurement points (Ward and Daniel, 2003).

4.4 Consideration of the innovative nature of MIBP

Throughout the cases, it can also be observed that business value measures for MIBP have a specific importance in the phase of assessing the business value potential prior to the investment, i.e. when building the business case. This is due to the innovative nature of MIBP for most companies. While the companies are able to estimate the impact of automating existing processes quite accurately, they face problems when the MIBP really leads to a new process design or effectiveness increases. Depending on the intended degree of innovation, also the consideration of realization risks will become a central element for assessing the business value potential.

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Degree of intended process innovation</th>
<th>Business Value realization risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility Company</td>
<td>High. Strong transformational character of new process. New role for many back office employees, reduction of staff in call center.</td>
<td>High realization risk as centralization of dispatching is imposing a role change for many coordinators. Risk is incorporated in time horizon of business value measures.</td>
</tr>
<tr>
<td>Logistics Company</td>
<td>Low. Basically automation of previous paper-based process. New is option to control tour in real-time.</td>
<td>Low realization risk internally as stricter tour control is supported by most drivers. Higher risk of integrating external drivers.</td>
</tr>
<tr>
<td>Airport Operating Company</td>
<td>High. Often new roles with more responsibility for mobile workers.</td>
<td>Business Value realization risk is reduced by earlier involvement of staff in MIBP definition.</td>
</tr>
<tr>
<td>Sales Company</td>
<td>Medium degree of innovation. Base process of sales people is not changed, but extended with new work tasks.</td>
<td>Medium realization risk. To ensure acceptance, the mobile workers get paid extra for the new work tasks.</td>
</tr>
</tbody>
</table>

Table 4. Overview of case study findings in relation to the forth MIBP characteristic

To cope with the specifics that arise from the innovative nature of MIBP, we derive the sixth requirement on MIBP business value measures.

Requirement 6: MIBP business value measures should regard the degree of intended innovation and consider the corresponding realization risks especially when the MIBP has a transformational character.

In the IT business value literature, the importance of measures for the assessment phase prior to the investment is for example highlighted by Davern and Kauffman (Davern and Kaufman, 2000) who also conclude the measures need to be suitable for the estimation of business value and the subsequent tracking of achieved results. The business value dimension model by Mooney, Gurbaxani and Kraemer (Mooney et al., 1996) also helps in the assessment phase. Based on the degree of intended change, automational, information and transformational, the business value expectation can be set. The consideration of business value realization risk in the measurement process is considered for example by Peppard, Ward and Daniel in his distinction of problem-based and innovation-based interventions (Peppard et al., 2007).

4.5 Resulting set of requirements on MIBP business value measures

Based on the analysis of the specific MIBP characteristics, we derive a set of six distinct requirements that need to be considered when defining appropriate MIBP business value measures, which is shown in Figure 2.
5 CONCLUSION

The research objective of this paper was to develop a set of requirements on appropriate business value measures for MIBP. Based on the case studies of four different companies with distinct experiences in the implementation of mobile and ubiquitous technologies, we identified four central characteristics of MIBP and derived six subsequent requirements that need to be considered when defining business value measures of mobile-integrated business processes.

Comparing these requirements with existing IT business value frameworks, it can be concluded that especially the short time horizon and the separate assessment of the business value outcomes for different stakeholders stand out and can be seen as very particular for MIBP. Furthermore, it can be concluded that the complementarity of mobile technology and organizational changes is central for the definition of MIBP business value measures. The complementarity points to the fact that mobile and ubiquitous technology can only unfold its potential in connection with associated business process changes and organization flexibility. The innovative nature of MIBP also demands for the development of a business value concept that especially supports companies in the phase of evaluating the business value potential and designing MIBP that have a true transformational character. This will be an essential issue for realizing the true potential that MIBP offer.

This paper only represents a first step in the development of a business value framework for MIBP. In order to ensure the applicability for all types of MIBP, the research needs to include a wider range of cases, especially in areas such as warehouse optimization, facility management, construction or healthcare solutions. The eventual framework will also require an extension with regard to the public service sector where business value is rather created through improved offerings to the people rather than cost savings or revenue increases. The result is especially intended to be used ex ante and to provide an improved third quadrant to the Mobility-M, enabling better investment decisions within the process model for mobile business process reengineering.
References


## INFLUENCES OF SIX SIGMA EMBRACEMENT AND ABONDENMENT (TEACHING CASE)

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0476.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Teaching Case</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Case, Business Process Management, Business Process Improvement, Case Study</td>
</tr>
</tbody>
</table>
OPEN PROCESS INNOVATION: A MULTI-METHOD STUDY ON THE INVOLVEMENT OF CUSTOMERS AND CONSULTANTS IN PUBLIC SECTOR BPM

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0629.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Open Innovation, Business Process Management, Empirical study, Governance</td>
</tr>
</tbody>
</table>
OPEN PROCESS INNOVATION: 
A MULTI-METHOD STUDY ON THE INVOLVEMENT OF 
CUSTOMERS AND CONSULTANTS IN PUBLIC SECTOR BPM 

Niehaves, Björn, European Research Center for Information Systems, University of Münster, Leonardo-Campus 3, 48149, Germany, bjoern.niehaves@ercis.uni-muenster.de 

Abstract 

Open Process Innovation – drawing from the literature on Open Innovation and Business Process Management (BPM) – promotes the study of how to systematically make use of knowledge that lies outside of an organization’s boundaries for process innovation initiatives. Open Innovation has been heavily studied for product innovation, however, process innovation has not yet been researched from such perspective. Against this background, we seek to investigate into variables that impact on the qualities of Open Process Innovation taking the example of the public sector domain. This paper examines how personnel resource scarcity exerts influence on the involvement of i) customers (here citizens and local companies) and ii) consultants (here management and software consultants) in public sector BPM. Our multi-method analysis shows that personnel resource scarcity has consequences for BPM-related collaboration schema as it restricts the involvement of customers. Based on our findings, implications for theory and practice are discussed, including implications for studies on BPM maturity or on business process design. We call for a governance-theory perspective on process innovation as a fundamental basis for understanding and designing the institutions that shape collaboration in open process innovation. 

Keywords: Open Process Innovation, Business Process Management, Governance, Public Sector
1 INTRODUCTION

External actors play an increasingly important role in public sector reforms. Various approaches seek to modernize, improve or restructure public administrations, be it practices in context of New Public Management (Pollitt and Bouckaert 2004), Electronic Government (Lenk 2004), or business process management (BPM). Despite the distinct phrasing, all of these approaches overlap in several elements. The central goals of local government reforms are on the one hand concerned with cost reduction (Asgarkhani 2005) and the increase of efficiency (Ridder et al. 2005) and, on the other hand, with customer orientation (Reichard 2003). Due to the high pressure, the diversity of demands, and new areas of responsibility, local governments increasingly rely on innovation networks (Rethemeyer 2007). Various external actors are involved in reform processes, such as software and management consultancies (Pratchett 1998) or individual citizens (Wollmann 2000). Here, Open Innovation can be regarded as a management paradigm addressing related challenges (Chesbrough 2003a, 2003b) as it studies the role of external actors in extending the pool of capabilities of a (government) organization (Feller et al. 2008).

Open Process Innovation can be regarded as a management perspective on process innovation which promotes the study of how BPM activities could be more successful when making use of BPM knowledge which lies outside of the organizational boundaries. Research on process management maturity discusses two major types of external actors for BPM collaboration, customers and BPM experts (for instance, Rosemann et al. 2006). Against this background, we seek to investigate into variables that impact on the qualities of Open Process Innovation taking the example of the public sector domain. This paper addresses the following research question:

*Does personnel resource scarcity exert influence on the involvement of i) customers (here citizens and local companies) and ii) consultants (here management and software consultants) in public sector Business Process Management?*

In order to achieve this research aim, we undertook a comprehensive multi-method study. First, within a 22-months time frame, 16 interviews were conducted with experts in local government BPM in Germany. As a second step, a quantitative analysis of BPM-collaboration with customers and consultants was carried out, taking into account the answers of 357 local governments.

The paper is structured as follows: Section 2 builds the theoretical foundation of our analysis. Subsequently, the discussion of the research methodology applied takes into account issues of method selection, interviewee selection, and data collection and analysis. Following the result discussion, we seek to reflect on the implications for theory and try to open up new vistas for BPM practice. The final section contains conclusions.

2 THEORETICAL FOUNDATION

2.1 Policy Networks and Collaboration

Against the background of a mounting policy and system complexity in the public sector, governments have to increasingly rely on resources of knowledge, capacity, and expertise outside of their own organizational boundaries (Rhodes 1997). As Rethemeyer (2007: 201) puts it, “[c]ontemporary government cannot function unless it can leverage the knowledge and expertise of vast networks of public and private sector organizations.” Here, policy networks can be regarded as structures in which stakeholders of public policy carry out decision-making. Policy networks might be understood as “sets of public agencies, legislative offices, and private sector organizations (including interest groups,
corporations, and non-profits) that have an interest in public policy within a particular domain” (Rethemeyer 2007: 201). Despite its importance, little research has yet elaborated the role of networks in the area of BPM (a related study on collaborative e-Government is Gil-Garcia 2007). While only few public sector reform studies take an actor-oriented spin (for instance, Sarker et al. 2006; Kuhlmann et al. 2008), an explicit BPM network approach is not to be found in current literature.

2.2 Open Innovation

The paradigm of ‘Open Innovation’, first described by Chesbrough (2003a, 2003b), tries to address the high demands of innovation processes. Companies find themselves exposed to constantly rising pressure due to higher competition, increase of acceleration and rising customer demands. Research and development divisions are often dysfunctional in coping with such increased pressure. Hence, in contrast to ‘Closed Innovation’, companies focus on acquisition of external knowledge, e.g. by know-how buy-in or the support of universities. This results in blurring enterprise boundaries, in particular the boundaries of processes in product and service development. In each development phase, external knowledge can be integrated as well as knowledge can be extracted and brought to market as independent products. The outside-in process extends the knowledge base of a company, whereas the inside-out process aims at commercialisation of ideas and sale of intellectual property. The coupled process describes work in alliances of complementary partners, where give and take is crucial to success (Gassmann and Enkel 2004). While Chesbrough (2003a, 2003b) originally examines private sector product innovation, the approach can be applied to the public sector as well. Here, various forms of collaboration can already be found, including collaboration with consulting companies (Pratchett 1998) or individual citizens (Wollmann 2000).

2.3 BPM and Process Innovation

BPM describes the efforts of an organisation to manage its processes, for instance, to monitor, analyse and optimize them. BPM can be considered to subsume fields like process innovation, optimisation, improvement or reengineering (Hung 2006, Hammer & Champy 1993, Davenport 1997, Zairi 1997, Breyfogle 2003). BPM habitually includes methods to automate tasks, define processes as a sequence of work steps, and to define responsibilities (for a comprehensive overview of IS and process innovation see Tarafdar & Gordon 2007). Major characteristics of public administrations, in contrast to private companies, are a high density of legal rules and a larger variety of goals: guarantee of proper legislation and jurisdiction, promotion of economic development, defence of public rights or environmental protection are only some of them (Lenk et al. 2001). For BPM in local governments, these issues imply more complex processes that contain a multitude of decision points and that are rarely well structured.

External actors play an integral role in BPM (de Vreede 1998; Dean et al. 1995; den Hengst & de Vreede 2004, Magdaleno et al. 2008) and especially in models of BPM maturity. For instance, Fisher (2004), Rosemann & deBruin (2005) and Rosemann et al. (2006) develop BPM maturity models with the intent to assess and evaluate BPM activities in organizations. Habitually, five stages are differentiated from one another, ranging from an initial state with uncoordinated and unstructured attempts to optimized BPM being core part of strategic and operational management and incorporating customers, suppliers, distributors and other stakeholders. Following these frameworks, openness – in terms of systematically involving stakeholders in BPM activities – is a major characteristic of high BPM maturity, while the heavy reliance on external expertise – here consultants – is a characteristic of low BPM maturity (Rosemann et al. 2006; see also Fisher 2004, Rosemann & deBruin 2005). Against this background, we seek to investigate into the question of other variables than that of BPM maturity – in the context of this paper personnel resource scarcity – exert influence on BPM-related
collaboration with consultants and customers (in our public sector study: citizens and local companies).

2.4 Open Process Innovation

Accordingly, Open Process Innovation can be regarded as a management perspective on process innovation which promotes the study of how BPM activities could make use of BPM knowledge that lies outside of the organizational boundaries. While such understanding opens up for an analysis of a variety of potential BPM-collaborators (for an overview, for instance, Niehaves & Kobayashi 2009), we focus in our analysis on BPM-related collaboration with consultants (management and software consultants) and customers (citizens and local companies) on the level of an individual process.

3 RESEARCH DESIGN

3.1 Qualitative Method

Method selection. A series of 16 semi-structured expert interviews was conducted within a 22 months-timeframe. Here, qualitative expert interviews allowed for gathering rich data and for building up a deeper understanding of the phenomenon under investigation. The involvement of external actors in public sector BPM initiatives has not yet been studied intensively – neither with qualitative nor quantitative means. Against this background, we sought to reflect insights derived from the literature analysis as a first step for a subsequent quantitative study.

<table>
<thead>
<tr>
<th>Number of interviews</th>
<th>16</th>
</tr>
</thead>
</table>
| Interview partners   | _Local government officials responsible for BPM-related reforms, including mayors, department heads, CIOs, project managers_  
|                      | _Representative range of public administrations regarding the size of the organizations_ |
| Period of data collection | June 2006 to March 2008 |
| Duration per interview | Average of 65 Minutes |
| Transcribed pages | ~ 150 pages |
| Transcribed words | ~ 95,000 words |

Table 1. Data Collection Fact Sheet

Interviewee selection. Interviewees include public officials that are responsible for BPM-related reforms in German local governments. The researcher’s professional network and recommendations by other study participants have been the basis for potential interviewee identification (snowball sampling). Regarding the size of the administration represented in this study, the set is representative (regarding the German setting) as it covers small(est), medium-sized and large(st) organizations. Moreover, it covers data from public organizations from all parts of Germany. Thus, reflections on a representative organization size and geographical distribution have been major criteria for the interviewee selection.

Data collection and analysis. The interviews were semi-structured, meaning that a part of the interview was shaped by questions derived from the literature analysis, while an additional free part allowed for an open discussion of other relevant aspects brought up in the interviews. All interviews were tape-recorded and transcribed afterwards resulting in 150 pages of transcript, comprising about 95,000 words (for details see Table 1). The transcripts were then analyzed and coded against the background of the variables/constructs in the research model.
3.2 Quantitative Method

Data Collection. The quantitative part of our study was carried out in 2008 and covers 357 cases of German local governments in the sample. A questionnaire was posted online and a random sample of 8,000 government officials, each responsible for (BPM) reforms in a single local administration, were invited to participate in the study. Out of about 12,250 local governments in Germany, thus, ~65% have been contacted and ~3% answered the questionnaire. The answers comprise local governments in 13 out of 13 large-area federal states – plus Berlin. An additional non-response analysis did not reveal any biases in the study participants.

Data Analysis Technique. In order to study the impact of personnel scarcity on BPM collaboration schema, the two samples test of significance (also known as independent samples t-test) was chosen. Such statistical test examines the difference between two sample means and here indicates if an independent variable (grouping variable for the samples) exerts influence on a specific dependent variable (for details regard this statistical test see, for instance, Argyrous 2005).

As for this study, the independent variable (grouping variable) is the personnel resource scarcity of an organization while the dependent variable represents the importance of a particular actor in BPM networks. Regarding the independent variable, the degree of agreement on the statement of “There is a lack of personnel resources to conduct desirable reforms in our public administration” led to two different groups (not agree; agree). Regarding the dependent variable, the study participants were asked “How important was ACTOR_X for business process management in your local public administration in the last five years?”. The questionnaire allowed for an answer by five degrees of importance, ranging from 5 (="Very important") over 4, 3, 2 to 1 (="Not important at all"). Consequently, a code 5 means that the actor is at the core, while a code of 1 indicates that the actor is in the periphery of a network. Following the results from our literature review and our qualitative part of the study, here, four external actors were analyzed for their importance in BPM networks: citizens (customers), local companies (customers), management consultancies, software consultancies. Accordingly, for example, the answer to the question of “How important were citizens for business process management in your local public administration in the last five years?” was included with a 5-point likert scale in the group-comparative independent samples t-test. Therefore, four dependent variables (each actor) were analyzed for their dependence on personnel resource scarcity.
4 RESULTS

4.1 Qualitative Study Results

External Experts in Public Sector BPM. Business Process Management (BPM) activities cover a broad variety of issues ranging from selecting process modeling methods, defining process documentation standards, implementing IT support for BPM activities, establishing a process-oriented culture, BPM training, BPM governance to concrete process analysis, optimization, and design (Rosemann et al. 2006). Such topics might be very challenging as they require comprehensive, but very specific BPM know-how. Against the background, that BPM activities are comparably new to the public sector, the majority of public organizations do not feature this know-how and their staff is most frequently better educated in other areas, such as law, regulations, or task-specific aspects, rather than in BPM. Here, a public official states:

Due to the size of our organization, we often don’t have the necessary in-depth know-how [for BPM reform issues]. Our staff has to deal with a broad range of topics rather than being a very expert [...] in BPM. Therefore, we might sometimes run into situations in which we have to seek for support by professionals, by external experts.

Thus, the need for BPM specific know-how often requires that local governments seek external BPM expertise. To give an example, a public official identifies BPM-related communication action as a potential field which requires support from outside of the organization:

Process design and implementation heavily relies on effective internal communication. Against this background, several departments in our organization have already carried out communication analyses, tried to identify potential to improve BPM-related communication, and provided specific training for the department heads and staff members. Here, we did receive support by an external expert.

Citizens in Public Sector BPM. When it comes to innovating those processes which feature points of interaction with organization-external entities citizens are seen as a potential cooperation partner. A public official argues:

20 years ago, we could observe an attitude among our colleagues like ‘The government is in charge and has to define what is to be done’. However, this attitude changed pretty much. Today, we perceive our administration as a public service provider. And as we provide services to our customers, we don’t want to be disconnected, but to stay in close contact with them. This also means that we do want to exchange ideas with the citizens.

The administration understands itself as a service provider to its customers, the citizens. In that role, innovation of external business processes, here service innovation, is considered as a mean to enact citizen-oriented service design. Another argument to involve citizens in process management activities is concerned with the acceptance of BPM-related change efforts:

[Process innovations […] need to be accepted by the stakeholders. I’m not sure, if it’s efficient to involve that many people, but I do know that if you want to implement successfully, you have to have the people on board. My experience shows that it’s way easier to achieve acceptance for certain change processes in case you’ve involved people beforehand than in case you’re just ordering to do something.

Proceedings ECIS 2009
To sum up, the interviewees identify the necessity to involve citizens in BPM activities. However, the arguments provided—customer-orientation in service design and acceptance of BPM-related change effort—do apply mainly in the context of innovation external rather than internal business processes.

Personnel resource scarcity and the involvement of consultants. The involvement of external experts is seen from a critical perspective, especially with regard to employee motivation:

In general, we prefer to involve internal know-how and internal ideas. Employees have a totally different identification with those ideas they have created themselves than to those ideas that came from others, for instance, external experts.

However, the interviewees identify a multitude of factors, related to a lack of resources, which necessitate the involvement of external experts. A major argument is that of a lack of BPM-specific know-how among the local government employees. This argument becomes even stronger in case of a lack of personnel resources as the public administration employees are already working long hours and do not have the time to educate themselves in BPM. A public official formulates:

Many of our employees, including myself, are engineers. We’re experts on different content than [BPM] reforms and we, thus, have to familiarize ourselves with specific reform know-how … but that’s besides everything else we do.

Additionally, arguments are found that BPM projects often exceed the personnel capacities available in the local government. Such capacity deficit necessitates the involvement of external expert, here consultants:

In my opinion, it is very important to involve know-how and to get capacity support by external experts. Regarding the in fact little personnel resources we have available, it wouldn’t be possible to conduct such large [BPM] projects on our own.

We’ve just completed a large organizational redesign project where we had wide-ranging support by a consulting company. That project would have been far too capacity-intensive for the personnel resources available within our own organization.

When considering different ways to achieve necessary BPM know-how, public administration might have two basic option: first, to hire new people or to educate existing staff (internal solution), second, to hire external experts, consultants. Here, arguments are found that economic necessities might support that external experts are involved instead of hiring additional staff: BPM know-how is often too specific and too expensive to have experts working as employees in the local government. A public official argues:

We look for cooperation with external experts in case we aren’t able to solve a specific problem on our own … or if it’s just cheaper to involve external expertise.

For the future, I do believe that cooperation [in process-related reform activities] will increase. This is due to pure economic necessity and we won’t have too many other options.

To sum up, the qualitative data indicates that a limitation of personnel resources makes it necessary to involve external experts, meaning to hire consultants. To local governments, such cooperative solution to BPM efforts tends to be more cost-efficient than hiring new employees with such BPM expertise or to comprehensive educate existing staff in BPM.
Personnel resource scarcity and the involvement of citizens. Regarding the effect of a lack of personnel resources on citizens cooperation in service innovation, a public official states:

*One of the main reasons to not involve citizens [in our BPM activities] is that we’re not able to manage too many projects at the same time. We have 80 employees working in our organization [...] The big project we’re working on at the moment is concerned with shared service structures. Before that, it was the implementation of a managerial accounting system and this involved the work of all the employees of our administration. Thus, we’re only able to operate with the limited personnel resources available and this in fact means doing one project after the other.*

Here, it becomes clear that the involvement of citizens requires personnel capacity and that such cooperation effort has to be seen in the context of increasing legitimacy and transparency of public services and the administration.

### 4.2 Quantitative Study Results

The independent samples t-test and additional descriptive analyses were processed applying the SPSS 16.0 software package. Table 2 shows the Independent Samples t-Test Group Statistics, Table 3 provides the Independent Samples t-Test Results.

<table>
<thead>
<tr>
<th>Personnel Resource Scarcity?</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citizens (Customer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Agree</td>
<td>109</td>
<td>2.38</td>
<td>1.095</td>
<td>.105</td>
</tr>
<tr>
<td>Agree</td>
<td>238</td>
<td>1.98</td>
<td>1.006</td>
<td>.065</td>
</tr>
<tr>
<td>Local Companies (Customer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Agree</td>
<td>110</td>
<td>2.20</td>
<td>1.107</td>
<td>.106</td>
</tr>
<tr>
<td>Agree</td>
<td>237</td>
<td>1.93</td>
<td>.954</td>
<td>.062</td>
</tr>
<tr>
<td>Management Consultancies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Agree</td>
<td>110</td>
<td>2.62</td>
<td>1.196</td>
<td>.114</td>
</tr>
<tr>
<td>Agree</td>
<td>238</td>
<td>2.48</td>
<td>1.165</td>
<td>.075</td>
</tr>
<tr>
<td>Software Consultancies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Agree</td>
<td>111</td>
<td>3.05</td>
<td>1.205</td>
<td>.114</td>
</tr>
<tr>
<td>Agree</td>
<td>235</td>
<td>3.02</td>
<td>1.136</td>
<td>.074</td>
</tr>
</tbody>
</table>

**Table 2. Group Statistics**

The results show that the influence of personnel resource scarcity, here the group-difference, is significant for both customer groups (citizens and local companies) at 0.95 confidence level! However, group-differences are not significant for consultants (neither management nor software consultancies). The lower and upper limits of the 95% confidence interval (last two columns, Table 3) span the range in which the difference between the population means is to be expected. If such interval does not include the value of 0 (here: applies to both customer groups), we can reject the hypothesis that the population means were equal. This reads as: With a confidence of 95%, we can assume that the importance of customers (both citizens and local companies) is different in organizations depending on personnel resource scarcity.
### Levene's Test for Equality of Variances

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>Sig.</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Citizens (Customer)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>3.033</td>
<td>.082</td>
<td>3.283</td>
<td>345</td>
<td>.001</td>
<td>.393</td>
<td>.120</td>
<td>.158 to .628</td>
</tr>
<tr>
<td>.. not assumed</td>
<td>3.181</td>
<td>.002</td>
<td>3.368</td>
<td>345</td>
<td>.001</td>
<td>.393</td>
<td>.124</td>
<td>.149 to .637</td>
</tr>
<tr>
<td><strong>Local Companies (Customer)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>7.458</td>
<td>.007</td>
<td>2.307</td>
<td>345</td>
<td>.022</td>
<td>.268</td>
<td>.116</td>
<td>.039 to .496</td>
</tr>
<tr>
<td>.. not assumed</td>
<td>2.185</td>
<td>.100</td>
<td>2.185</td>
<td>186,875</td>
<td>.030</td>
<td>.268</td>
<td>.122</td>
<td>.026 to .509</td>
</tr>
<tr>
<td><strong>Management Consultancies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>.142</td>
<td>.706</td>
<td>.997</td>
<td>346</td>
<td>.320</td>
<td>.135</td>
<td>.135</td>
<td>-.131 to .401</td>
</tr>
<tr>
<td>.. not assumed</td>
<td>9.87</td>
<td>.002</td>
<td>207,139</td>
<td>.325</td>
<td>.030</td>
<td>.135</td>
<td>.137</td>
<td>-.135 to .405</td>
</tr>
<tr>
<td><strong>Software Consultancies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>.537</td>
<td>.464</td>
<td>.278</td>
<td>344</td>
<td>.782</td>
<td>.037</td>
<td>.133</td>
<td>-.225 to .299</td>
</tr>
<tr>
<td>.. not assumed</td>
<td>2.72</td>
<td>.100</td>
<td>204,789</td>
<td>.786</td>
<td>.037</td>
<td>.133</td>
<td>.136</td>
<td>-.232 to .306</td>
</tr>
</tbody>
</table>

*Table 3. Independent Samples Test*

### 5 DISCUSSION

The independent samples t-test reveals that there is a significant group difference in the importance of customer (both citizens and local companies) in public sector BPM depending on the scarcity of personnel resources in the organization. Customers play a less important role in BPM networks of organizations in which personnel resources are perceived scarce. Against the background of the qualitative study results, we understand that, in general, the involvement of customers in process innovation is regarded as a desirable action. Such insights confirm the BPM maturation of an organization towards and intelligent operating network where processes are innovated on the basis of an extensive involvement of stakeholders (for instance, Rosemann et al. 2006). Personnel resource scarcity, however, is a (new) variable to be considered in the BPM maturity and governance discussion. A limitation of personnel resources has an impact on the cooperation schema in BPM initiatives. Literature on BPM maturity (for instance, Rosemann et al. 2006; see also Fisher 2004, Rosemann u. de Bruin 2005) discusses the importance of external actors in BPM. Here, the heavy reliance on external expertise is identified as a characteristic of early maturity stages while the involvement of stakeholders, especially customers, is regarded as a characteristic of high maturity. However, personnel resources do have an impact on this picture. Here, we can interpret that the involvement of customers in process innovation, and thus the maturation of BPM initiatives, is hindered by personnel resource scarcity. This opens up for follow up questions of how customer collaboration in BPM can be designed more resource-efficiently.

Taking a governance-theory perspective (Williamson 1975, see also Rowley et al. 2000), potentially fruitful avenues for future research could include the study of how to design the institutions that govern the involvement of customers (and other actors) in process innovation. How could institutions of open process innovations be designed in order to allow for a more effective and efficient involvement of customers? For instance, such studies could include the design of business process modelling methods or toolsets – in terms of design science (Hevner et al. 2004; Niehaves 2007a) – in order to reduce transaction costs of collaboration. Such might not primarily focus on the optimization of business processes on a very detailed level, but rather provide a more general understanding of the
processes, their aims, and their context. Here, modelling could take place on a higher degree of abstraction, containing modules or building blocks instead of detailed description of each minor step. This approach ensures an increase of transparency and, hence, legitimacy. In addition, management-oriented studies could examine methods for improving quality and costs of customer-collaborative BPM, for instance, drawing from methods of open innovation (Chesbrough 2003a).

Moreover, we consider it potentially fruitful future research to examine the effect of personnel resource scarcity – and other variables – on the involvement of additional external actors (for instance other local governments (Becker et al. 2003; Algermissen et al. 2005) or superordinate organizations such as central governments (Niehaves 2007b)) and, thus, to provide a more holistic picture on BPM networks and Open Process Innovation. Too, a comparison of Open Process Innovation in the public and private sector could be regarded prolific future research (for network structures and innovation in the private sector see, for instance, Capaldo 2006).

6 SUMMARY

BPM requires specific expertise and knowledge, assets which might not be available inside of an organization, here local governments. These organizations may seek to involve external actors in their BPM activities. Here, Open Process Innovation – drawing from the literature on Open Innovation and Business Process Management (BPM) – promotes the study of how to systematically make use of knowledge that lies outside of an organization’s boundaries for process innovation initiatives. Open Innovation has been heavily studied for product innovation, however, process innovation has not yet been researched from such perspective. Against this background, we investigated into variables that impact on the qualities of Open Process Innovation taking the example of the public sector domain. This paper specifically examined how personnel resource scarcity exerts influence on the involvement of i) customers (here citizens and local companies) and ii) consultants (here management and software consultants) in public sector BPM. Our multi-method analysis shows that personnel resource scarcity has consequences for BPM-related collaboration schema as it restricts the involvement of customers. Such insight opens up for future research. Here, we call for a governance-theory perspective on process innovation as a fundamental basis for understanding and designing the institutions that shape collaboration in Open Process Innovation. How can organizational BPM benefit from knowledge outside of the organizational boundaries without immense increase of transaction costs? Determinants/Institutions could be addresses, for instance, by management-oriented studies as well as by design science studies (e.g., on the design of business process modeling methods or tools). Too, the analysis of the influence of other variables on BPM collaboration, also taking into account other actors than customers and consultants, can be considered potentially fruitful future research.

ACKNOWLEDGEMENT

The author thanks Michael Räckers, Burkhard Weiß, and Jörg Becker for helpful comments and support as well as the ECIS editors and three anonymous reviewers for insights and helpful recommendations. This work would not have been possible without the support of our graduate students Shaoxian Zhou, Andreas Krause, Robin Trenkner, and Robert Malsch.
REFERENCES


ONLINE, SET, GO - DESIGN AND EMPIRICAL TEST OF AN IT-BASED PHYSICAL ACTIVITY INTERVENTION

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0193.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Telemedicine, Behaviour change, User participation, Action research</td>
</tr>
</tbody>
</table>
ONLINE, SET, GO - DESIGN AND EMPIRICAL TEST OF AN IT-BASED PHYSICAL ACTIVITY INTERVENTION

Knebel, Uta, Technische Universitaet Muenchen, Boltzmannstr. 3, 85748 Garching, Germany, knebel@in.tum.de
Leimeister, Jan Marco, Universitaet Kassel, Nora-Platiel-Str. 4, 34127 Kassel, Germany, leimeister@uni-kassel.de
Esch, Sebastian, Technische Universitaet Muenchen, Boltzmannstr. 3, 85748 Garching, Germany, esch@in.tum.de
Pressler, Axel, Technische Universitaet Muenchen, Conollystr. 32, 80809 Muenchen, Germany, pressler@sport.med.tum.de
Krcmar, Helmut, Technische Universitaet Muenchen, Boltzmannstr. 3, 85748 Garching, Germany, krcmar@in.tum.de

Abstract

Inactivity is the most widespread health risk factor in modern societies today, causing not only individual health problems but also immense costs for the healthcare systems. This emphasizes the need for improving population-wide impact of activity interventions, with particular attention to cost-effectiveness, scalability, and delivery channels. In this paper, we present the theory-motivated design (drawing on the transtheoretical model) and empirical test of an IT-based physical activity programme (Personal Health Manager, PHM). In order to be as cost-effective as possible, the PHM was designed to have only few face-to-face contacts and to deliver supervision through the internet.

Our design and implementation proved to be successful in a pilot test with 88 employees of an automotive company. The PHM increased participants’ activity, motivational readiness for change, functional capacity and transported the feeling of being well taken care of. Enhanced supervision did not increase performance. The results are first evidence that internet-mediated supervision can be successful in promoting physical activity and provide a starting point for investigating the role of face-to-face-contact and service levels in physical activity programs. The PHM and similar designs are also relevant to practice as the semi-automation makes them eligible for large-scale corporate or public health programs.

Keywords: telemedicine, behaviour change, user participation, action research
1 INTRODUCTION

Physical activity and exercise are widely recognized today as key factors in maintaining and restoring health (Wagner & Brehm 2008). Many health problems and diseases are associated with certain habits, among them lack of physical activity. For the US, it is estimated that about 12% of yearly deaths can be ascribed to lack of regular physical activity (Pate et al. 1995b). Furthermore, it is assumed that about one quarter of the population of the western industrial nations are suffering from metabolic syndrome (Whaley et al. 1999). The metabolic syndrome as a term for the simultaneous occurrence of multiple risk factors (all linked to physical inactivity) is presumed to be a warning signal for the development of coronary heart diseases. However, only about 10 to 20 percent of the adult population in western industrial nations can be described as active enough to achieve health protecting effects. Inactivity therefore is the most widespread risk factor in the population, causing not only individual health problems but also immense costs for the healthcare systems (Wagner & Brehm 2008).

The prevalence of inactivity and its economic consequences emphasize the need for improving population-wide impact of activity interventions, with particular attention to cost-effectiveness, scalability, and delivery channels. As a conclusion of a review of physical activity interventions, Eakin (Eakin 2000) points out the need for research on non-physician delivery of activity intervention and on cost-effective ways of activity consulting and the conduction of follow-ups. In a recent review, Marcus et al. (2006) confirm and detail these research demands: “Further research should pursue better understanding of the minimal amount of face-to-face contact necessary for behaviour change and related cost-effectiveness issues. Questions concerning the most effective channel or combination of channels (e.g. print, telephone, or internet) for intervention delivery must be answered […]”. The development or internet- and IT-based programmes seems even more appropriate as the diffusion of internet access across all social classes and ages has been rising constantly over the past years (Statistisches Bundesamt 2008).

This research takes up these issues from an information systems point of view. Drawing on the transtheoretical model, we designed an IT-based physical activity programme for currently inactive people, combining both automatic and face-to-face services. The programme was implemented and tested in a three month pilot test in cooperation with the health promotion department of a German automotive company. In this article, we will focus on the participants’ side and perception of the programme. Key questions are if a physical activity intervention with supervision mainly delivered through internet-mediated channels is effective in increasing physical activity behavior, and if participants perceive this supervision as appropriate and satisfying.

2 THEORETICAL BACKGROUND: THE TRANSTHEORETICAL MODEL APPLIED TO PHYSICAL ACTIVITY

The transtheoretical model (TTM) is an integrative framework for understanding how individuals progress toward adopting and maintaining health behaviour change (Prochaska et al. 1998). The TTM uses stages of change to integrate processes and principles from various theories of intervention and psychotherapy, hence the name “transtheoretical”. Originally developed in the context of addictive behaviour, it has been widely applied to physical activity (e.g. Marcus & Forsyth 2003, Marcus et al. 1998b, Marcus et al. 2000, Velicer et al. 2006, Prochaska et al. 2008, Sarkin et al. 2001).

The TTM describes intentional behaviour change as a temporal process that unfolds through a series of five sequential stages. Progress through the stages is impacted by 10 processes of change. Other constructs affecting progress through the stages are the decisional balance, reflecting the individual’s relative weighting of pros and cons of changing, as well as self-efficacy, reflecting the situation-specific confidence in practicing the new behaviour as intended (Prochaska & Velicer 1997).
2.1 Stages of Change

The stage construct represents a temporal dimension, a characteristic that sets it apart from any other construct used in many social-cognitive theories. According to DiClemente et al. (DiClemente et al. 1991), the construct is in between temporally stable personality traits and dynamically changing personality states – the stages can thus be both stable and dynamic. The original concept assumed a unidirectional progress through the stages. This has been replaced by a cyclical or spiral approach: The stages can be passed through repeatedly, allowing for both remaining on a certain stage as well as relapsing to a prior stage (Prochaska et al. 1992).

The following paragraph describes the five stages of change (Keller et al. 1999) and their adaption to physical activity (Marcus et al. 1992).

1) Precontemplation: Individuals in precontemplation have no intention to change. This can be due to ignorance about the consequences of their current behaviour, but can also be a consequence of resignation after having failed to change repeatedly. For physical activity: Individuals do no physical activity and do not intend to start in the next six months (“not thinking about change”).

2) Contemplation: Individuals in contemplation intend to change. They think about a certain behaviour and weigh pros and cons. Ambivalences arise. For physical activity: Individuals do no physical activity but intend to start in the next six months (“thinking about change”).

3) Preparation: Individuals in preparation have decided to change. They have designed an action plan and have taken first steps to change. For physical activity: Individuals do some physical activity but not at levels meeting the recommended level1 (“doing some physical activity”) and they may or may not intend to become more physically active.

4) Action: Individuals in action have been practicing the new behaviour for a short time (up to six months). The risk for relapses is especially high on this stage. For physical activity: Individuals participate in recommended amounts of physical activity but have done so for less than six months and may or may not maintain this level of physical activity (“doing enough physical activity”).

5) Maintenance: Individuals in maintenance have been practicing the new behaviour for more than six months and have reached a relative stability. The risk for relapses decreases. For physical activity: Individuals have participated in recommended amounts of physical activity for six months or longer (“making physical activity a habit”).

2.2 Processes of change

The processes of change are covert or overt activities that people use to progress through the stages. Prochaska et al. have identified ten processes of change reflecting two main categories: cognitive and behavioural processes (Prochaska et al. 1988). For many behaviours, cognitive processes seem to be especially relevant in early stages (contemplation, preparation), whereas behavioral processes are mainly applied in later stages (action, maintenance) (Prochaska et al. 1994). Specifically for exercise behaviour, Marcus, Rossi, Selby et al. (Marcus et al. 1992) have identified the cognitive and behavioral processes displayed in Table 1. The question if some processes of change are of more or less importance on different stages of change in exercise behaviour has not been answered yet.

---

Comparative analyses indicate that both cognitive and behavioral processes are relevant in all stages (Marcus et al. 1992), being applied more often the higher the stage is (Marshall & Biddle 2001).

<table>
<thead>
<tr>
<th>Cognitive processes</th>
<th>Behavioural processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increasing knowledge</td>
<td>Substituting alternatives</td>
</tr>
<tr>
<td>Being aware of risks</td>
<td>Enlisting social support</td>
</tr>
<tr>
<td>Caring about consequences to others</td>
<td>Rewarding yourself</td>
</tr>
<tr>
<td>Comprehending benefits</td>
<td>Committting yourself</td>
</tr>
<tr>
<td>Increasing healthy opportunities</td>
<td>Reminding yourself</td>
</tr>
</tbody>
</table>

*Table 1: Processes of change in exercise behaviour (Marcus et al. 1992)*

### 2.3 Further constructs in TTM: self-efficacy and decisional balance

Self-efficacy (Bandura 1997) is the individual situation-specific confidence a person has that he or she can cope with high-risk situations without relapsing to their unhealthy behaviour. For exercise, it reflects the confidence in doing physical activity even if confronted with certain barriers. Self-efficacy is assumed to be lower in early stages than in higher stages (Lippke & Plotnikoff 2006), (Marshall & Biddle 2001).

The decisional balance reflects the individual’s relative weighing of the pros and cons of changing. The perceived pros increase in the first three stages and then remain on an almost equally high level throughout the last two stages. In contrast, the cons decrease throughout the last three stages (Keller et al. 1999), (Marshall & Biddle 2001).

### 2.4 Implications for the design of the Personal Health Manager

We use the theoretical foundations of the TTM for designing the PHM following a theory-based design approach (Briggs 2006, Leimeister et al. 2009a). Following the TTM, interventions must be adapted to the specific stages of change the targeted persons are in (Loughlan & Mutrie 1997). Helping people to progress one stage in the course of the program seems a reasonable goal (Prochaska et al. 1998). Past research suggests that movement from the early to the more advanced stages of motivational readiness for physical activity adoption is also significantly associated with changes in functional capacity (maximum oxygen uptake, VO2 max) (Marcus et al. 1998a).

In the case of the PHM, we target people engaging in physical activity less than once a week or not at all. According to the definitions above, those people would be in the early stages of change, namely precontemplation, contemplation and preparation (I1). Processes of change act as independent variables that people need to apply to move from stage to stage (Prochaska et al. 1998). They therefore provide important guides for interventions for every stage that should be addressed in the design of the PHM (I1.1). Self-efficacy in our targeted stages is still low, and should be increased (I1.3). The communication in these stages should emphasize the pros of physical activity (I1.4). Another effective way of adapting interventions to participants apart from targeting them according to a certain variable (e.g. stage of change), is tailoring them to the individual. A combination of both approaches has been found to increase physical activity (Marcus & Forsyth 2003) and should be addressed in the PHM (I2). Figure 1 shows an overview of the described relationships. All implications are presented in Table 2.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I1.</td>
<td>The PHM should target people in the stages of precontemplation, contemplation and preparation.</td>
</tr>
<tr>
<td>I1.1.</td>
<td>The PHM should increase cognitive processes of change.</td>
</tr>
<tr>
<td>I1.2.</td>
<td>The PHM should increase behavioral processes of change.</td>
</tr>
<tr>
<td>I1.3.</td>
<td>The PHM should increase self-efficacy.</td>
</tr>
<tr>
<td>I1.4.</td>
<td>The PHM should emphasize the pros of physical activity.</td>
</tr>
</tbody>
</table>
3 DESIGN OF THE PHM

3.1 Procedure

As most physical activity interventions, the PHM program will be subdivided into a start phase (SP), an activity phase (AP) and an end phase (EP). The start and end phase consist of mainly singular activities, such as informing the participants on the procedure of the intervention or the end of the intervention, training them on the topic or collecting evaluation data. The collection of evaluation data (e.g. medical data) mostly requires face-to-face contact. We will therefore design most activities of the start and end phase as face-to-face contacts.

In our opinion, the most interesting phase for testing alternative delivery channels is the activity phase. We have two reasons for this: First, supervision in the activity phase is very time and resource-consuming. Supervision of the participant in the activity phase requires the supervisor to perform mainly repeated activities, such as checking the participant’s performance, giving feedback, answering questions, motivating and reminding over a longer period of time. Second, supervision in the activity phase is very important to help participants stick to physical activity. Partial automation or IT-support for supervising processes and their delivery through the internet or other media therefore hold great potentials for reducing cost and increasing the scalability of activity interventions.

As a preparation for the design, we first analyzed and documented the basic supervision activities delivered in offline physical activity programs. We then identified potentials for automation or IT-support in these activities using a four step approach assessing priority (frequent, time- and cost-intensive activities with high possible savings are considered first), quality standard and customer experience (customer still should perceive individual supervision), technical feasibility, and required parties in the activity (for instance, registration for the PHM could be performed by the participant interacting with a system in absence of a supervisor). With this background information, we started designing elements of the physical activity programme to meet the requirements derived from TTM. The design was done in multiple iterative steps, and was reviewed and refined by an interdisciplinary team with backgrounds in information systems, sports medicine, business and communication (Knebel et al. 2007). A prior, much simpler prototype had previously delivered promising results in field tests (Leimeister et al. 2009b). The following sections provide an overview of the design elements addressing the requirements from TTM.
3.2 Addressing precontemplation, contemplation and preparation (I1)

In the early stages, people usually have little knowledge or experience with physical activity. The content of all information should therefore be basic knowledge about physical activity presented in a simple way (D1). How this content should be communicated is detailed in the following sections.

3.3 Addressing processes of change, self-efficacy and pros of physical activity (I1.1-I2)

For physical activity, all of the processes are important in all of the stages of change (see section 2.2). However, it is hardly feasible to emphasize all processes at each contact with the individual, and thus key processes are usually selected based on a person’s stage (Marcus et al. 1998a). In our design, we focused on the cognitive processes and on 4 of the behavioural processes. Some elements seem appropriate to address more than one process, or address self-efficacy and the pros of physical activity as well. Table 3 shows the variables from TTM (implications from 2.4; I1.1-I2) and the PHM design elements addressing them (D1.1-D2).

<table>
<thead>
<tr>
<th>TTM Implications (I)</th>
<th>PHM Design element (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D1.1: Support cognitive processes (I1.1)</strong></td>
<td></td>
</tr>
</tbody>
</table>
| Increasing knowledge | • training unit (importance of physical activity, advice on how to become active) (SP)  
• print information on how much and which activities to do (SP)  
• training unit (how to use PHM online system) (SP)  
• online exercise plan with instructions (AP)  
• e-Mail information on diet and exercise (2 per week) (AP) |
| Being aware of risks | • training unit (risk factors for metabolic syndrome) (SP)  
• print information on metabolic syndrome (SP) |
| Caring about consequences to others | (implicit) |
| Comprehending benefits | • training unit (benefits for physical and psychological health) (SP)  
• physical fitness test (oxygen uptake) before and after the program (SP, EP) |
| Increasing healthy opportunities | • training unit (everyday activity as taking stairs etc) (SP)  
• time- and place-independent access to online system (AP)  
• flexibility (individual exercise at home possible, choice from a variety of endurance sports possible) (AP)  
• organize inexpensive, time-limited test membership in corporate gym (AP) |
| **D1.2: Support behavioural processes (I1.2)** |
| Substituting alternatives | (not addressed) |
| Enlisting social support | • training lecture organized in groups (SP)  
• invitation to participation sent out by employer (SP)  
• individual support by trainers possible (mail, phone) (AP) |
| Rewarding yourself | (not addressed) |
| Commiting yourself | • online documentation of activity (AP)  
• online exercise plan with instructions (AP)  
• weekly activity goal (AP) |
| Reminding yourself | • online exercise plan displayed as calendar (AP)  
• documented exercise units displayed in calendar (AP)  
• colour-coding for planned, completed, cancelled and missed exercise units (AP) |
| **D1.3: Support self-efficacy (I1.3)** |
| | • encourage people in all personal meetings (SP, EP, AP) |
| **D1.4: Emphasize the pros of physical activity (I1.4)** |

Proceedings ECIS 2009
• training unit, personal and e-mail communication present positive aspects of physical activity (SP, EP, AP)

Table 3: Design elements addressing processes of change, self-efficacy and pros. SP=start phase; AP=activity phase; EP=end phase (own illustration)

3.4 Addressing the tailoring of the intervention (I2)

In addition to the stage-targeted design elements presented in the previous section, we integrated elements that provide individually tailored feedback based on the participants’ input (Table 4).

<table>
<thead>
<tr>
<th>Requirement</th>
<th>PHM design element</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2: Tailoring of the intervention (I2)</td>
<td></td>
</tr>
<tr>
<td>• system calculates individually recommendable heart frequency for exercise (AP)</td>
<td></td>
</tr>
<tr>
<td>• systems checks goal achievement and adjusts or retains exercise plan (AP)</td>
<td></td>
</tr>
<tr>
<td>• trainer checks documentation and contacts participants actively every month (AP)</td>
<td></td>
</tr>
<tr>
<td>• individual supervisor feedback based on exercise documentation (monthly) (AP)</td>
<td></td>
</tr>
<tr>
<td>• message system in PHM system to contact trainer (AP)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Design elements addressing the tailoring of the PHM. SP=start phase; AP=activity phase; EP=end phase (own illustration)

3.5 Schedule

Start phase: Participants start the PHM activity program with a 90 min training, covering importance of physical activity, benefits, risks and advice on how to become more active, as well as how to use the PHM internet platform. The training is accompanied by print information material. All participants take a fitness test.

Activity phase: After the fitness test, the participants get access to an online exercise plan containing endurance and strength units. The plan contains instructions and also a documentation tool. To reach the weekly activity goal, participants must have performed and documented all planned exercise units. They perform their exercise following the plan autonomously, using a heart rate monitor to check the intensity. Exercise can take place at home, outdoors, or at the gym. In specific situations the participants are contacted by a supervisor: First, in a monthly phone call, second, if the supervisor notices major irregularities in the participant’s activity documentation (for example not reaching the weekly goal multiple times, documentation about not being satisfied etc), or third, if the participant himself actively requests information. This enables the supervisor to interfere where it is most necessary and motivate potential dropouts to continue with their activity. E-Mail information on exercise and diet are sent out to all participants twice a week.

End phase: At the end of the programme, the participants repeat the fitness test and take part in a closing event.

4 PILOT TEST

4.1 Evaluation goals

We have two major evaluation goals. First, to evaluate the overall effectiveness of our design. This will be answered by comparing the output variables described in Figure 1 (motivational readiness, activity level, functional capacity) before and after the program. We will also consider the subjective
impressions of the participants concerning the semi-automatic nature of the PHM. Second, we aim to evaluate the effectiveness of those design elements supporting active supervision in the activity phase. This will be answered by comparing two test groups, one receiving active supervision in the activity phase (premium group), and one not receiving active supervision in the activity phase (basic group).

Our data collection is therefore guided by the following research questions:

- Do the PHM participants progress in their motivational readiness for change? Is this effect stronger for the premium group?
- Are the PHM participants more active? Is this effect stronger for the premium group?
- Does the functional capacity of the PHM participants increase? Is this effect stronger for the premium group?
- Do the participants feel positive about the PHM? Is this effect stronger for the premium group?

### 4.2 Test setting

The pilot test was conducted in cooperation with the worksite health promotion of a major German automotive company. To make participation as easy as possible for the employees, all fitness tests and all administrative issues took place at the corporate gym, the initial trainings and closing events at a university site at a small distance from one of the major company buildings. The elements of the program were implemented as described in the schedule in the previous section (total duration: 4 months, April-July 2008). The participants were randomized into two groups. The basic group received all designed services in the start and end phase, but only passive support in the activity phase (e.g. by providing them with an activity documentation tool). The premium group received all the services the basic group received and in addition to that active support in the activity phase (Figure 2).

![Figure 2: PHM implementation for basic and premium test group](image)

### 4.3 Data collection and measures

**Motivational readiness** for physical activity adoption was assessed using the five-item motivational stage measure for physical activity by Marcus & Rossi (1992) both at the beginning and at the end of the PHM programme. To capture **functional capacity**, we measured the maximum oxygen uptake (VO2 max) in the fitness tests at the beginning and the end of the program, on the same day as or very few days from the stage measure.

**Physical activity level** at the beginning and the end of the program was assessed using a self-report physical activity questionnaire (Freiburg questionnaire of physical activity, Frey et al. 1999), subdivided in everyday activity, sport activity and total activity. People reaching a total score of min. 30 points or min. 14 sport activity points are classified as doing enough physical activity, a total score of 15-29 as fulfilling the minimum, and below as being inactive.
Perception of the PHM was assessed by single items on a 5-point scale at the end of the program. Items included statements to evaluate if participants felt healthier, gained knowledge and liked using the PHM system.

4.4 Participants

We decided on an active recruitment approach, which has proved to be more successful than reactive approaches in other health promotion contexts (Prochaska et al. 1998). 140 inactive, overweight men and women with 2 or more metabolic syndrome risk factors were recruited through an in-house mailing sent by the worksite health promotion department and randomized into the basic (BG, n=55) and the premium group (PG, n=85). 107 took the fitness test in the start phase, 88 completed all questionnaires (baseline and 3 month questionnaire) in the pilot test (n basic group (nb)=33, n premium group (np)=55). Their age ranged from 25 to 60, averaging at 45.4 years. The remainder of this paper focuses on the participants who completed both questionnaires.

5 RESULTS

Motivational readiness for change: As expected, the vast majority of participants in both groups was in precontemplation, contemplation or preparation at baseline (PG: 78.2%, BG: 97.0%), 21.8% of the PG and 3% of the BG surprisingly already on a higher stage. After three months, 65.5% of the PG and 81.9% of the BG had progressed at least one stage. At the end of the programme, 72.4% of the PG and 81.8% of the BG had reached at least the action stage, meaning they were physically active on a regular basis (Table 5). These results indicate a very positive effect of the PHM, but must be interpreted with caution. According to the definition of the stages, a progress from preparation to maintenance (active for more than six months) is not possible within three months. Apparently, some participants did not understand or not answer the questions correctly. The average overall functional capacity (VO2 max) increased (p=0.03; np=46, nb=26), the difference between PG (+ 4.4%) and BG (+2.8%) was not significant (p=0.66).

<table>
<thead>
<tr>
<th>Stage of change</th>
<th>Premium group</th>
<th>Basic group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base-line</td>
<td>3 months</td>
</tr>
<tr>
<td>Precontemplation</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Contemplation</td>
<td>21</td>
<td>6</td>
</tr>
<tr>
<td>Preparation</td>
<td>21</td>
<td>7</td>
</tr>
<tr>
<td>Action</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Maintenance</td>
<td>9</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 5: Number of individuals in each stage of change and progress through the stages (np=55; nb=33; own illustration)

Physical activity level: Everyday, sport and total activity have increased significantly in both groups at the end of the program. 49.1% of the PG and 42.4% of the BG have reached the recommended minimal activity level of 15 total activity points (Figure 3). There is no significant difference between the test groups (p>0.2).

Perception: The PHM was well received by the participants. On average, the participants felt well taken care of, liked using the system, found it easy to use, and have gained knowledge about physical activity and health. They felt healthier and more active (Table 6). The statements “I feel healthier” and “I have become more active” are strongly correlated (p<0.01). Moreover, “I feel healthier” is weakly correlated with functional capacity (p<0.05). Surprisingly, neither of the two is significantly correlated with the increased scores of the physical activity questionnaire. There is no significant difference between the test groups (p>0.2).
Figure 3: Physical activity level (np=55, nb=33; own illustration)

Table 6: Participants' perception of the PHM program (np=55, nb=33; own illustration)

<table>
<thead>
<tr>
<th>Statement (5=strongly agree, 1=do not agree at all)</th>
<th>Premium group</th>
<th>Basic group</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have become more active.</td>
<td>4.05</td>
<td>4.18</td>
</tr>
<tr>
<td>I have gained knowledge about physical activity and health.</td>
<td>4.22</td>
<td>4.15</td>
</tr>
<tr>
<td>I feel healthier.</td>
<td>4.02</td>
<td>3.82</td>
</tr>
<tr>
<td>I like using the PHM system.</td>
<td>3.98</td>
<td>3.76</td>
</tr>
<tr>
<td>Using the PHM system is easy for me.</td>
<td>4.09</td>
<td>4.24</td>
</tr>
<tr>
<td>I was well taken care of and supervised.</td>
<td>4.16</td>
<td>3.85</td>
</tr>
</tbody>
</table>

6 DISCUSSION

6.1 Summary and contributions

In this paper we have presented the design and empirical test of an IT-based physical activity programme, the PHM. Apart from face-to-face contacts in the beginning and in the end of the program, supervision was mainly delivered internet-mediated. Our design decisions were motivated by the TTM. This approach has proved to be successful in the pilot test. The PHM increased participants’ activity, functional capacity and motivational readiness for change. The supervision via internet was perceived positively, and the participants had the feeling of being well taken care of. However, enhanced active supervision in the activity phase did not produce stronger effects.

Our research has successfully implemented the approach of theory-driven design as proposed by Briggs (Briggs 2006), applying a theory from psychology to the design of an information system. The results are first evidence that internet-mediated supervision can be successful in promoting physical activity and provides a starting point for investigating the role of face-to-face-contact and other forms of supervision in physical activity programs. The multi-method evaluation data allows a comparison of different measurement results and therefore better interpretation. A major advantage is that the PHM allows supervision of and situational reactions to user behavior, but at the same time, the computer-mediated delivery and supervisor support functions allow a fast and efficient supervision and administration. Therefore, practice will find the PHM and similar designs eligible and useful for large-scale corporate or public health programs, providing an alternative to conventional interventions requiring personal presence.

6.2 Limitations and outlook

The duration of our pilot test was only three months, which might be too short to induce long-term behavior change. We are currently conducting a follow-up survey to shed light on the long-term effects. We are aware that self-report measures as we applied among others can always be biased by the participants’ intended or unintended false answers. Where possible, we try to validate self-report
data with other measurement methods, e.g. stages of change and functional capacity. The fact that the premium group did not perform better than the basic group might be due to some limitations resulting from the implementation of the study. There might have been a sense of competition between the groups biasing the results. Technical problems at the beginning of the test did not allow the full delivery of all supervision services to all members of the PG in the first month. An analysis of the qualitative interviews we conducted with all participants at the end of the test will reveal insights into this. Another reason might be that the importance of enhanced supervision increases with the duration of the activity phase, and the effect would only be visible in a longer test.

Further research should therefore examine the effectiveness of IT-supported physical interventions over longer periods of time and test different service levels of supervision. In addition to theories on exercise behavior, it should further draw on knowledge from the computer-human-interaction and technology acceptance and refine the intervention design. It is also necessary to inquire on appropriate easy and reliable methods to accurately capture and supervise participants’ activity levels.

References


HEALTH INFORMATION PORTALS: HOW CAN WE IMPROVE THE USER’S SEARCH EXPERIENCE?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0148.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Health information systems / medical record systems / care delivery /, Human computer interaction (HCI), E-health, Usability</td>
</tr>
</tbody>
</table>
HEALTH INFORMATION PORTALS: HOW CAN WE IMPROVE THE USER’S SEARCH EXPERIENCE?

Fisher, Julie, Monash University, Julie.Fisher@infotech.monash.edu.au
Burstein, Frada, Monash University Frada.Burstein@infotech.monash.edu.au
Rosetta Manaszewicz, Monash University rmanasz@optusnet.com.au
Lazarenko, Kate Monash University Kate.Lazarenko@infotech.monash.edu.au
Monash University PO Box 197, Caulfield East, 3145, Australia

Abstract

Health information portals (HIP) are now common place. Governments and other health agencies are using HIPs extensively to reduce costs and distribute information more effectively. Generally, HIPs are not very technically sophisticated specifically in terms of options for improving searching. There are many ways in which searching and retrieving relevant information can be improved. This paper presents an exploratory study which investigated five HIPs. Each HIP offered a range of features and functionality to assist with searching. Our research explored the features and functionality of each HIP. Through usability evaluations we compared the response of users to each HIP and identified users’ preferences for improved searching. We found that HIPs with improved search functionality and other features that assisted searching were better received by the users. Users regarded these portals as easier to understand, easier to use, required fewer steps in retrieving information and were more likely to say they would return. Comments from users are provided to illustrate further the importance of providing effective functionality. The paper concludes with recommendations for Health Information Portal builders on what is needed to improve the user search experience.

Keywords: Health information portals, functionality, usability.

1 INTRODUCTION

Governments worldwide are grappling with increased demands for health information and pressures on health systems. Coupled with this is an explosion in the quantity of health information available particularly through the Internet. Many governments and health agencies are disseminating health information via the Internet. Reasons for doing so include mitigating against the problems of poor quality information and helping defray costs in information distribution. Research findings by Kennedy (2003) suggest a strong correlation between health outcomes and the level to which patients are informed, that is, when patients have more information relating to their health this leads to improved health outcomes. Despite the best efforts of the various agencies distributing information, health consumers frequently report that the Internet based health information they find is not useful and are not meeting user needs (Vermaas and Wijngaert 2005). Information relevance is most important to health consumers (HON 2006) yet finding appropriate, relevant and timely information is often difficult (Zeng et al. 2004). Information is often incomplete (HON 2006) and difficult to read (Sillence et al. 2004). Further, most health Websites and portals swamp users with information (Burstein et al. 2005). There are however, a variety of ways in which users can be assisted in overcoming the identified problems in searching and retrieving relevant information.

To date there has been limited research exploring users’ perceptions of HIPs and what encourages use, this research sought to fill this gap. The exploratory study reported in this paper investigated the features and functionality of five non commercial Health Information Portals (HIP). Through usability
evaluations in a laboratory setting we explored users’ responses to each of the HIPs, their expectations when searching for health information, in particular how they responded to their search experience and the extent to which this was influenced by the features and functionality of the HIP and their willingness to use the same source in the future. The research found that HIPs with improved search features/ functionality, were better received by the users. Users also rated these portals as easy to use and as a result, were more likely to return to them.

2 HEALTH INFORMATION PORTALS

This section discusses mechanisms for improving searching and information retrieval from health websites or HIPs and factors that determine success. For the purposes of this paper a HIP is defined as an internet based system which assists users find relevant internet based health information resources. A HIP would generally facilitate information retrieval from other websites through a range of features and functionality but not store or provide information specifically.

Health websites should enable health consumers to easily find relevant information. A number of features a health information website or portal might have to support searching were identified by Luo and Najdawi (2004): “a catalogue of health information” - a mechanism for organizing edited health information; a search engine; “a personalization system” described as allowing health consumers to customize the interface based on their preferences and; “a network of communities” where health consumers can exchange information with others (Luo and Najdawi 2004). Other technologically possible features include 1. Providing a level of information differentiation to improve searching such as information organized based on a specific disease, drugs, for children, personal stories. 2. Spellchecking, many medical terms are unfamiliar to health consumers and difficult to spell. 3. Parsing where users can enter a question or sentence and the search examines both the key word/s and the context. 4. An ontology and/or thesaurus helps users identify specific medical search terms and narrow searching.

Other factors that will determine how successful a HIP will be include the users’ ability to find relevant information (Josefsson 2006; Pew/Internet 2006). One study reported that 37% of users were unable to find relevant health information (Zeng et al. 2004). Through observations, Zeng et al (2004) found a major reason was “the consumers’ use of simple search strategies (browsing or short text queries) that did not characterize their information needs well.” Zeng et al (2004) conclude from their research that there are three ways searching can be improved; health consumers making queries more specific, improved search functionality and limiting information retrieved. Research by Pew/Internet (2006) found 25% of users were overwhelmed by the amount of information retrieved and 22% were frustrated by their inability to find relevant information. Easy to use websites are also critical to success (Klein 2007). The 2006 HON survey (2006) found ‘Easy to use’ was ranked highest as helpful for searching. Factors such as the quality of the navigation, how long it takes to retrieve information and how easy the text is to read on the screen all contribute to ease of use (HON 2006). Research also suggests that these elements will impact on the overall user experience, influencing whether or not users will return to a Website (Fisher et al. 2007).

There are many examples of health information portals. Non-commercial providers are the focus of this research as previous research has found health consumers trust these sites are more (HON Survey 2006). Non-commercial providers include hospitals, governments and patient/advocacy groups. The choice for the average health consumer is therefore vast. The question this poses is how does the health consumer choose between these sources? This is no trivial decision if the patient is facing a life threatening illness and information and answers in relation to treatment are critical (Josefsson 2006).

The purpose of the research reported in this paper was to examine the features and functionality offered by HIPs in relation to searching, the impact this has on the user’s experience and to identify what users want. Whilst the authors recognise the importance of information quality, quality is a complex issue and was outside the scope of this research. The questions we sought to answer were:
What different features / functionality are provided for users of HIPs to improve searching?
What searching features / functionality do users want?
What impact do different features / functionality have on users’ overall experience, including searching and their intention to return to a HIP?

3 RESEARCH APPROACH

Five generic non-commercial, HIPs were identified. Portals from Europe (Health on the Net), North America (Medline Plus and Mayo Clinic) and Australia (Better Health and HealthInsight) where the research was conducted, were identified. The sites were selected based on the following criteria:

- Government sponsored. The HON Survey (2006) identified that 79% of health consumers preferred a government agency to be responsible for online health information provision.
- The health information provided is generic that is not specific to a disease, age group etc. This was to ensure the relevance of the task to all users.
- HIPs were selected based on the level of features / functionality available. We looked for HIPs with a greater number of features / functionality to test their effectiveness. Note only two Australian HIPs met the criteria.

The Australian portals are the two key government sponsored portals, Medline Plus claims to be the largest medical library in the world (Medline Plus 2009), the Mayo clinic is the largest not for profit practice in the world (Mayo 2009) and HON describes itself as one of the most respected HIPs in the world (HON 2006). It should be noted that in 1999 (not in subsequent years) the HON survey (1999) asked users which Website most closely met their needs. The results found Medline Plus was rated as the highest, Mayo Health System (now Mayo Clinic) was listed number eight and HON number 10.

From previous work we identified a number of features/functionality that can assist users searching for information (Fisher et al. 2007). We examined the home page of each HIP for evidence of features/functionality. Users are most likely to use features/functionality visible from the home page. Table 1 details the features/functionality and the method used to assess each. All had a search engine.

<table>
<thead>
<tr>
<th>Feature/functionality</th>
<th>How the feature or functionality was determined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personalisation</td>
<td>Examined if the portal asked for information to help in retrieving information relevant to the user. Did the HIP start an initial dialog to identify user’s personal needs eg ask the gender of the user, age or information type?</td>
</tr>
<tr>
<td>Differentiated Information access</td>
<td>Analysed information differentiation offered. Were users given topics to search within? Could the user retrieve other types of information eg personal stories?</td>
</tr>
<tr>
<td>Spell check and “Sounds like” index</td>
<td>Spell checking and ‘sounds like’ indexing tested using misspelt words.</td>
</tr>
<tr>
<td>Parsing</td>
<td>The sentence “What are the side effects of Ventolin?” was used to search each HIP to test if returned results contained both the phrase ‘side effect’ and ‘Ventolin’</td>
</tr>
<tr>
<td>Ontology and Thesaurus</td>
<td>Checked for an ontology and/or thesaurus. For example was a list of possible search words or terms available for users to access?</td>
</tr>
<tr>
<td>Other features</td>
<td>Were newsletters, forums or individual feedback available?</td>
</tr>
</tbody>
</table>

Table 1 – Features and functionality for HIPs

Usability is “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” (Bevan 2001). The research examined users’ response to the five HIPs, through a usability evaluation - in particular aspects of the HIPs that related to searching and retrieving information. The usability instrument was based on an instrument developed and tested previously by one of the authors (Fisher et al. 2004). A full copy of the instrument can be found at http://www.sims.monash.edu.au/staff/jfisher/.

The instrument is based on key items identified in the literature as important in websites designed for information distribution. The key items are how easy information is to read including terminology and...
instructions on using the Website (Becker and Mottay 2001; Cox and Dale 2002; Kunst 2002). The quantity of information presented (Bruce 1999; Zeng et al. 2004). The design and presentation of the text (Becker and Mottay 2001; Sutcliffe 2002) and how easy it is to use including locating information and navigation (Nielsen 1999; Fu and Salvendy 2002; Hargittai 2003).

The usability evaluation involved 223 users. The instrument consisted of Likert scale type statements, questions requiring a response from 1 scored lowest to 5 highest, questions requiring a Yes/No response (scored Yes 2, No 1) and questions requiring a free text response. Numerous usability instruments exist however each evaluation conducted needs to be designed specifically for the goals of the research and adjusted to meet those goals (Dumas and Redish 1994, 185). The instrument contains questions and statements similar to those in other studies (Zhang et al 2000; Nel et al. 1999).

The users were students, studying Human Computer Interaction, as part of their studies and during a class they participated in the usability test. 223 users evaluated the portals resulting in 411 usable evaluations. 33% of participants were female and 67% male and 91% described themselves as very experienced with Internet searching. The use of students for usability evaluations such as this is in line with other studies (Nel et al. 1999; Zhang et al. 2000). Abdinnour-Helm et al. (2005) argue that students can be appropriate providing they are similar to Web users generally and are likely to perform the tasks being investigated. The majority of users (91%) were aged between 21 and 30 years. This age group are likely to have searched for health information (Fox 2006). The users were asked to ‘Think of a health issue that is important to you, a friend or a family member. Using each of the health websites, search for information on that topic’. Users were provided with examples of information they might search for (exercise, diet, medication, a specific health condition). Each user conducted a search on two HIPs; one was an Australian portal and one international portal. Once the search was completed users completed the questionnaire and reviewed the next HIP. Each user spent about 15 minutes searching each HIP.

The quantitative data were analysed using SPSS. A Factor Analysis was conducted to explore the relationship between the identified elements and to assess the degree to which factors were measuring the same or a similar concept (Bryman and Cramer 1992) and descriptive statistics to compare results for individual HIPs. The qualitative data was analysed for the research themes relating to searching, using a meta-matrix approach which is described by Miles and Huberman (1994, pg 93) as “essentially the 'crossing' of two lists, set up as rows and columns.”

4 RESULTS

Users were asked if they had searched for information on the internet, 10% said they had never used the internet to search for health information, 70% indicated they occasionally search, 18% search often. Users were also asked how they searched. Users could tick more than one option. Most users, 185 (97%) had used a search engine, 7% used an Australian Website and 2% used an international Website. Next each portal is described briefly and the features and functionality available.

4.1 Features and functionality

**Medline Plus** [http://medlineplus.gov/](http://medlineplus.gov/) is the National Library of Medicine (NLM) located in Bethesda, Maryland, USA. It describes itself as bringing “together authoritative information from NLM, the National Institutes of Health (NIH), and other government agencies and health-related organizations.” Table 2 presents an analysis of features and functionality of Medline Plus.

<table>
<thead>
<tr>
<th>Feature / Functionality</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personalisation</td>
<td>Not available</td>
</tr>
<tr>
<td>Differentiated Information access</td>
<td>750 available health topics organised according to diseases, diagnoses, demographics and wellness. Searchable topics on the different parts of the body, diagrams are provided also.</td>
</tr>
<tr>
<td>Spell check, “Sounds like” index</td>
<td>Yes. Dictionary is also provided to help with spelling and definitions.</td>
</tr>
</tbody>
</table>
### Table 2: Features and functionality available on Medline Plus

**Mayo Clinic** (www.mayoclinic.org) Website describes its website as providing “information and services from the world's first and largest integrated, not-for-profit group medical practice.” Table 3 presents the analysis of features and functionality of the Mayo Clinic.

<table>
<thead>
<tr>
<th>Feature / Functionality</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personalisation</td>
<td>Not available</td>
</tr>
<tr>
<td>Differentiated Information access</td>
<td>Diseases / Treatments listed alphabetically. Users can search by clicking on a letter. Patient stories are provided for some topics.</td>
</tr>
<tr>
<td>Spell check, “Sounds like” index</td>
<td>Yes</td>
</tr>
<tr>
<td>Parsing</td>
<td>No</td>
</tr>
<tr>
<td>Ontology and Thesaurus</td>
<td>No</td>
</tr>
<tr>
<td>Other features</td>
<td>Subscription to an e-mail newsletter is available</td>
</tr>
</tbody>
</table>

### Table 3: Features and functionality available on Mayo Clinic

**Health on the Net Foundation (HON)** (www.hon.ch) Swiss based says it is “the leading organization promoting and guiding the deployment of useful and reliable online medical and health information, and its appropriate and efficient use.” Table 4 presents the analysis of features and functionality.

<table>
<thead>
<tr>
<th>Feature / Functionality</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personalisation</td>
<td>Not available</td>
</tr>
<tr>
<td>Differentiated Information access</td>
<td>Differentiates between information for patients/individual from medical persons. Searching was possible based on age and gender.</td>
</tr>
<tr>
<td>Spell check, “Sounds like” index</td>
<td>Yes</td>
</tr>
<tr>
<td>Parsing</td>
<td>Yes</td>
</tr>
<tr>
<td>Ontology and Thesaurus</td>
<td>No</td>
</tr>
<tr>
<td>Other features</td>
<td>None</td>
</tr>
</tbody>
</table>

### Table 4: Features and functionality available on HON

**HealthInsite** (www.healthinsite.gov.au), established by the Australian Commonwealth Government and aims “to improve the health of Australians by providing easy access to quality information about human health.” Table 5 presents the analysis of HealthInsite features/functionality.

<table>
<thead>
<tr>
<th>Feature / Functionality</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personalisation</td>
<td>The advanced search allows users to indicate if the information is for a ‘child, youth, adult easy, adult medium or professional person’. Preference for other document types can also be indicated for example data, images, document, multimedia, statistics.</td>
</tr>
<tr>
<td>Differentiated Information access</td>
<td>Wide range of health topics organised under different headings such as diseases, wellbeing and stages of life. A-Z is search available.</td>
</tr>
<tr>
<td>Spell check, “Sounds like” index</td>
<td>No</td>
</tr>
<tr>
<td>Parsing</td>
<td>No</td>
</tr>
<tr>
<td>Ontology and Thesaurus</td>
<td>Yes but only if the user used ‘Advanced Search’</td>
</tr>
<tr>
<td>Other features</td>
<td>Monthly newsletter available on subscription</td>
</tr>
</tbody>
</table>

### Table 5: Features and functionality available on HealthInsite
Better Health ([www.betterhealth.vic.gov.au](http://www.betterhealth.vic.gov.au)) established by the Victorian Department of Human Services, Australia. Described as “a consumer health information Website for the Victorian community.” Table 6 presents the analysis of features/functionality for better Health.

<table>
<thead>
<tr>
<th>Feature / Functionality</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personalisation</td>
<td>Not available</td>
</tr>
<tr>
<td>Differentiated Information access</td>
<td>Has personal stories and limited number of health topics. An additional category of ‘healthy eating’ is provided. Some information is organised according to gender.</td>
</tr>
<tr>
<td>Spell check, “Sounds like” index</td>
<td>Yes</td>
</tr>
<tr>
<td>Parsing</td>
<td>No</td>
</tr>
<tr>
<td>Ontology and Thesaurus</td>
<td>No</td>
</tr>
<tr>
<td>Other features</td>
<td>Fact sheets organised by category or A-Z, hot topic of the week and some podcasts. Latest updates are available for those who subscribe.</td>
</tr>
</tbody>
</table>

Table 6 Features and functionality available on Better Health

Users were asked “Thinking about what you need to help you find the most relevant information please indicate on a scale of 1 (not important) to 5 (very important) how important each of the following features are”. This question was asked after users had searched using both portals. 200 users answered the question. Table 7 describes how each feature/functionality was described to participants and responses.

<table>
<thead>
<tr>
<th>Feature/ functionality</th>
<th>Description provided to participants</th>
<th>User rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ontology and Thesaurus</td>
<td>A list of terms to help you refine your search</td>
<td>4.08</td>
</tr>
<tr>
<td>Spell check and “Sounds like” index</td>
<td>Spell check: for example offering a list of options if a word is spelt incorrectly</td>
<td>3.87</td>
</tr>
<tr>
<td>Differentiated Information access</td>
<td>Choice of information type for example being able to choose medical /scientific information or information in simple language or personal stories.</td>
<td>3.77</td>
</tr>
<tr>
<td>Other features</td>
<td>Provided newsletters, e-mail, feedback from a professional.</td>
<td>3.33</td>
</tr>
<tr>
<td>Personalisation</td>
<td>Personalisation for example asking your preference for information, your age, gender any other information about your search that relates to you</td>
<td>3.28</td>
</tr>
</tbody>
</table>

Table 7 User views of features and functionality

4.2 Usability evaluation results

The usability evaluations explored more items than those presented in this paper. The results are confined to reporting on aspects influencing the users’ search experience. This includes factors such as ease of use, information quantity, terminology, design and presentation of text, navigation and information on searching. Each user evaluated two HIPs and all users evaluated one of the Australian HIPs. 76 users evaluated Medline, 66 evaluated Mayo Clinic, HON was evaluated by 65 users, 82 users evaluated Better Health and 122 evaluated HealthInsite. The HIPs were randomly assigned hence the difference in the number of users who evaluated each of the different portals.

4.2.1 Quantitative user responses to individual Health Information Portals

To understand the relationship between the different elements a factor analysis on nine of the variables was undertaken. Tables 8 and 9 present the questions/statements put to users and the results.

Communalities

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Extraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>The size of the text was easy to read</td>
<td>1.000</td>
<td>.833</td>
</tr>
<tr>
<td>The text was displayed in a way that was easy to read</td>
<td>1.000</td>
<td>.833</td>
</tr>
<tr>
<td>The language used was easy to understand</td>
<td>1.000</td>
<td>.765</td>
</tr>
</tbody>
</table>
All the information I required to complete the task was on the Website 1.000 .623
The number of steps required to get to the information I wanted was acceptable 1.000 .595
It was easy to find information on the topic that was relevant for me 1.000 .765
I understood the terminology used on the Website 1.000 .712
I found the search function useful in helping me locate relevant information 1.000 .589
The Website was easy to use 1.000 .496

Extraction Method: Principal Component Analysis.

Table 8 Factor Analysis Communalities Extraction Method: Principal Component Analysis.

Rotated Component Matrix

<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>It was easy to find information on the topic that was relevant for me</td>
<td>.846</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All the information I required to complete the task was on the Website</td>
<td>.776</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The number of steps required to get to the information I wanted was acceptable</td>
<td>.770</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I found the search function useful in helping me locate relevant information</td>
<td>.759</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Website was easy to use</td>
<td>.575</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The size of the text was easy to read</td>
<td>.908</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The text was displayed in a way that was easy to read</td>
<td>.888</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The language used was easy to understand</td>
<td>.863</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I understood the terminology used on the Website</td>
<td>.803</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Table 9 Factor Analysis Extraction Method: Principal Component Analysis.

The Kaiser-Meyer-Olkin value, measuring sampling adequacy was .783. This should be greater than .6 for a satisfactory factor analysis (Pallant 2001). Bartlett’s test of sphericity was .000. The factor analysis indicated 69% of the variance can be explained by three components: Factor 1 relates to how easily users were able to find relevant information, Factor 2 relates to the presentation of the text and Factor 3 is the language used.

Questions requiring a Yes/No response were asked of the users. Table 10 presents the results, both the percentage of the Yes / No responses and the average response is provided. Bold indicates a Yes result higher than the overall average (questions 1-3) or No (questions 4 and 5) lower than the average.

<table>
<thead>
<tr>
<th>Question</th>
<th>Medline</th>
<th>Mayo Clinic</th>
<th>Health On the Net</th>
<th>Health Insite</th>
<th>Better Health</th>
<th>All portals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Would you use this Website again to search for other health information?</td>
<td>No: 1.88</td>
<td>87%</td>
<td>25%</td>
<td>1.71</td>
<td>29%</td>
<td>82%</td>
</tr>
<tr>
<td></td>
<td>Yes: 1.30</td>
<td>89%</td>
<td>72%</td>
<td>1.71</td>
<td>29%</td>
<td>83%</td>
</tr>
<tr>
<td>2. Were you able to find information on the topic you wanted information on?</td>
<td>No: 1.92</td>
<td>92%</td>
<td>71%</td>
<td>1.80</td>
<td>1.30</td>
<td>46%</td>
</tr>
<tr>
<td></td>
<td>Yes: 1.42</td>
<td>92%</td>
<td>71%</td>
<td>1.80</td>
<td>1.30</td>
<td>46%</td>
</tr>
<tr>
<td>3. Was enough information for your needs or question you had on the topic provided?</td>
<td>No: 1.17</td>
<td>83%</td>
<td>69%</td>
<td>1.30</td>
<td>46%</td>
<td>1.42</td>
</tr>
<tr>
<td></td>
<td>Yes: 1.42</td>
<td>92%</td>
<td>71%</td>
<td>1.80</td>
<td>1.30</td>
<td>46%</td>
</tr>
<tr>
<td>4. Was there anything else you wanted to know but could not find out from the site?</td>
<td>No: 1.17</td>
<td>83%</td>
<td>69%</td>
<td>1.30</td>
<td>46%</td>
<td>1.42</td>
</tr>
<tr>
<td></td>
<td>Yes: 1.42</td>
<td>92%</td>
<td>71%</td>
<td>1.80</td>
<td>1.30</td>
<td>46%</td>
</tr>
<tr>
<td>5. Were you at any stage frustrated using the site?</td>
<td>No: 1.17</td>
<td>83%</td>
<td>69%</td>
<td>1.30</td>
<td>46%</td>
<td>1.42</td>
</tr>
<tr>
<td></td>
<td>Yes: 1.42</td>
<td>92%</td>
<td>71%</td>
<td>1.80</td>
<td>1.30</td>
<td>46%</td>
</tr>
</tbody>
</table>
Table 10 Responses to each HIP

Table 11 details the responses to the Likert scale statements and questions. The average for each HIP is presented and the average for all HIPs. Bold is again used to indicate better than average results.

<table>
<thead>
<tr>
<th>Overall how would you describe your experience using this site?</th>
<th>Medline</th>
<th>Mayo Clinic</th>
<th>Health On the Net</th>
<th>Health Insite</th>
<th>Better Health</th>
<th>All portals</th>
</tr>
</thead>
<tbody>
<tr>
<td>The site was easy to use</td>
<td>4.23</td>
<td>3.82</td>
<td>3.26</td>
<td>3.66</td>
<td>3.59</td>
<td>3.71</td>
</tr>
<tr>
<td>All the information I required to complete the task was on the Website</td>
<td>3.89</td>
<td>3.43</td>
<td>3.48</td>
<td>3.36</td>
<td>3.11</td>
<td>3.44</td>
</tr>
<tr>
<td>I found the search function useful in helping me locate relevant information</td>
<td>4.17</td>
<td>3.86</td>
<td>3.86</td>
<td>3.84</td>
<td>3.48</td>
<td>3.83</td>
</tr>
<tr>
<td>The size of the text was easy to read</td>
<td>3.87</td>
<td>3.65</td>
<td>3.20</td>
<td>3.39</td>
<td>3.77</td>
<td>3.56</td>
</tr>
<tr>
<td>The text was displayed in a way that was easy to read</td>
<td>3.92</td>
<td>3.71</td>
<td>3.29</td>
<td>3.46</td>
<td>3.70</td>
<td>3.60</td>
</tr>
<tr>
<td>The language used was easy to understand</td>
<td>4.07</td>
<td>3.73</td>
<td>3.73</td>
<td>3.92</td>
<td>3.68</td>
<td>3.84</td>
</tr>
<tr>
<td>I understood the terminology used on the Website</td>
<td>3.89</td>
<td>3.64</td>
<td>3.15</td>
<td>3.69</td>
<td>3.39</td>
<td>3.57</td>
</tr>
<tr>
<td>It was easy to find information on the topic that was relevant for me</td>
<td>4.12</td>
<td>3.62</td>
<td>3.64</td>
<td>3.46</td>
<td>2.93</td>
<td>3.53</td>
</tr>
<tr>
<td>The number of steps required to get to the information I wanted was acceptable</td>
<td>4.04</td>
<td>3.82</td>
<td>3.74</td>
<td>3.27</td>
<td>3.16</td>
<td>3.55</td>
</tr>
</tbody>
</table>

Table 11 Response to scale questions for individual HIPs

4.2.2 Qualitative user responses to the individual Health Information Portals

To better understand the users’ responses to the questions in Tables 11 and 12, open ended questions were asked. Qualitative comments relating to searching and finding information were analysed for two questions: “What was the best feature or part of the Website?” and “What was the worst feature or part of the Website?” Table 12 contains some responses where the comments.

<table>
<thead>
<tr>
<th>Best Feature</th>
<th>Worst Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medline Plus</td>
<td>Topics not categorised enough</td>
</tr>
<tr>
<td>Easy to find information by using the search field</td>
<td>Helpful search function</td>
</tr>
<tr>
<td>The health topics broke down well</td>
<td>Too many links to a specific page</td>
</tr>
<tr>
<td>It categorises information well such his symptoms, treatment etc (very well actually)</td>
<td>Some time more detailed information is there. You just have to export more</td>
</tr>
<tr>
<td>Variety of features/links available</td>
<td>Too much information is external</td>
</tr>
</tbody>
</table>

| Mayo Clinic | |
| Website is well structured and information for non-medical people is presented in quite general terms | There is no ‘back to search results’ link |
| Multiple ways of searching, list of problems, easy to understand headings and select them by the letter | Some diseases can’t be found by looking based on the first character |
| Topics can be accessed by alphabetical links | Limited information available on certain topics |
| Alphabetical list of diseases, treatments and services. Good search engine. | Information not enough |

| Health on the Net (HON) | |
| Types of searches. Search also checks for similar words in case you misspelt something | Some search results don’t match the category it is under |
| Search results categorised into groups. Search is possible on a variety of categories. | Hard to find particular health issue if you are not sure of the issue name |
| Search function | Customised search, confusing |
| Simple search function | Too much medical terminology |

<p>| Better Health | |
| There is a search function. There are headings and subheadings | Too many clicks. Not all relevant information at the same time. |</p>
<table>
<thead>
<tr>
<th>Provided relevant information</th>
<th>Search was chaotic returns irrelevant information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy to navigate and find information</td>
<td>Search engine didn't provide information if people entered the wrong spelling.</td>
</tr>
<tr>
<td>The menu provided information by health topic, saved lots of time</td>
<td>All. Don't know where to go other than the search when looking at the topic.</td>
</tr>
<tr>
<td><strong>Health Insite</strong></td>
<td><strong>Health Insite</strong></td>
</tr>
<tr>
<td>It provided an advanced search so users can easily find information they want.</td>
<td>Hard to read the results, confused between ‘Related HealthInsite topics’ and ‘Resources found’</td>
</tr>
<tr>
<td>Search and A-Z health topics is really easy to use and allows fast access to information</td>
<td>Too many links to click on before searching – the ultimate goal of this page.</td>
</tr>
<tr>
<td>The ‘Conditions and Diseases’ part has useful information to help me understand the diseases</td>
<td>No content of its own, I can just use Google instead of this site, unless I want Australian sites.</td>
</tr>
<tr>
<td>Searching by keywords</td>
<td>Too crude, not useful information for a health Website</td>
</tr>
</tbody>
</table>

*Table 12 Qualitative comments from users*

5 DISCUSSION

Fox and Rainie (2002) report that 53% of 18-29 year olds had searched the internet for health information. By 2006 this had grown to 77% (Fox 2006). We found 90% of our 18-30 year old users, have searched for health information, 97% using a search engine not a HIP. This is an internet savvy generation and the Internet is an important source of health information for this demographic. Our research demonstrates a link between the features users want, the features available on the portals we explored and the usability evaluation results. The two areas we examined and found impact on the user’s search experience are discussed next.

Features and functionality available to improve searching: Of the five HIPs analysed we found Medline Plus had the most extensive range of features and functionality. Providing an ontology or thesaurus to help users refine their search was the feature users wanted most followed by spell checking and differentiated information access. Only Medline Plus included an ontology and thesaurus. All, except HealthInsite had spellchecking however, only two assisted searching through parsing (HON and Medline Plus). All offered some form of differentiated information access. Only one HIP (HealthInsite) offered a personalised search but this was not available from the home page.

Users qualitative responses highlight further the importance of features to assist searching. It was evident in the responses users gave when asked about the best and worst features; many mentioned how useful the ‘dictionary’ was or a list of search terms or topics organised alphabetically. Users commented on the lack of spell checking, terminology and what they needed to improve their search. A number of users mentioned that the information was too broad suggesting that because some of the portals did not have well differentiated information that they could not refine their searches adequately enough. Even if a portal only provided an alphabetical search users commented that it was useful.

HIP design: The quality of the design includes how easy a HIP is to use and search, the way information is displayed and how frustrated users were in using the HIP. Tables 10 and 11 detailed users’ responses to a range of questions and statements relating to design. Medline Plus was ahead of the overall average on all items. Mayo Clinic was ahead on ten, HON, seven, HealthInsite five and Better Health, two. Users’ responses suggest a strong link between the quality of the HIP design, the features and functionality to assist searching and the overall reaction a user had to the HIP. The more features and functionality the HIP had the better the user experience. Medline Plus was the portal most users preferred, fewer users were frustrated and almost all said they were able to find information on the topic. Users also found it was the easiest to use, the search function was the most useful, the text display was regarded as the best and users said they were able to retrieve relevant information most easily from this HIP. Mayo Clinic was rated second by the users on most items; it also provides a high level of information differentiation and spell checking but does not offer an ontology or thesaurus. By
contrast the most poorly rated portal was Better Health, it provided minimum assistance to users for searching in terms of features and functionality. It had some information differentiated but this was limited to the information organised around a small number of topics. Better Health rated most poorly in terms of users’ ability to find information and to find enough information, 65% of users were frustrated at some point. Users indicated they did not like Better Health, they did not have enough information to complete the task, could not find relevant information and there were too many steps needed to retrieve information.

The number of negative comments users made can be an indication of dissatisfaction. User frustration was evident when presented with poor search results and lack of information. The number of users who said there was a ‘worst feature’ have been counted and expressed as a percentage of the total number of comments made for that portal. Note not all users provided a best or worst feature. There were very few negative comments from users on Medline Plus. 42 (68%) users commented on the worst feature eight users actually said there were no ‘worst features’. This portal attracted the fewest negative comments on the question of worst feature. A number of users mentioned specifically that the Mayo Clinic had a variety of ways of searching. In particular users mentioned the alphabetical search and how easy it was to search for a particular health issue. 77% (52 users) commented on the worst feature and of these none said there was no worst feature. Many of the users commented negatively on the quality of the search function of the HON Portal and many generally did not find it easy to use. 89% (58 users) made a comment on the worst feature, none said there was no worst feature. Fourteen users commented on the poor quality of the search function. 65% (72) users of the Better Health site commented on the worst feature, of those five said there was no worst feature. Most of the negative comments from users focused on the poor quality of the search function and the overall usability of the portal. For HealthInsite, 90% (110) of users comment on the worst feature. Only one user said there was no worst feature.

6 LIMITATIONS AND CONCLUSION

We acknowledge the fact that, most of the users were under the age of 30 and maybe less concerned about health matters than older users. It would therefore be useful to compare these results with users over 30. However, little previous research has reported on the extent to which people under 30 have searched for health information. It is not surprising that within our group more than 97% had searched for health information using a search engine. As it is likely this demographic will turn to the internet first for health information they are an important group to consider when designing a HIP. More statistical testing is needed to provide a deeper understanding of the impacts of the factors.

7 IMPROVING THE USER SEARCH EXPERIENCE

A key indicator of success is whether a user would use the HIP again. Most users (87%) would use Medline Plus again, 75% would use Mayo Clinic, 65% HON, 70% HealthInsite and 52% Better Health. One factor, the search experience, is likely to have influenced a user’s preparedness to return this includes finding information, finding information easily, the number of steps it takes to retrieve information and finding enough information. Medline Plus was ahead of all the other HIPs on all these items. Better Health however was last on three of the four (users found HON slightly less easy to use than Better Health). From this research it suggests Medline Plus offers users a better search experience, Medline Plus also provides the widest range of features and functionality to assist users in their search.

If HIP designers want to ensure users are satisfied with their search experience and will return consideration has to be given to improving the search experience. An analysis of the data helped us to identify features/functionality that should be included in a HIP. Table 13 describes those features/functionality and are in priority order.
<table>
<thead>
<tr>
<th>Feature/functionality</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ontology and Thesaurus</td>
<td>Provide easy to access list of medical words and search terms and alternative words.</td>
</tr>
<tr>
<td>Differentiated Information access</td>
<td>Include a wide range of topics and then sub categories and provide alternative ways to access information. Users may not know the exact term or phase they are searching for therefore an alphabetical list of health topics is a useful feature one users like and want.</td>
</tr>
<tr>
<td>Spell check and &quot;Sounds like index&quot;</td>
<td>Many users cannot spell medical terms. Spell checking is essential. HIPs must avoid returning nothing at all if the term is misspelt.</td>
</tr>
<tr>
<td>Parsing</td>
<td>Providing a quality parsing facility where the search takes into account phrases rather than one word in a user’s question. This assists users with making queries more specific. Users should be able to ask questions as part of a search.</td>
</tr>
<tr>
<td>Other features</td>
<td>Include pictures particularly to help explain what is presented, provide the ability to search within results, support multiple languages, newsletters and discussion boards can be useful.</td>
</tr>
<tr>
<td>Personalisation</td>
<td>Personalisation can be used to limit the quantity of information retrieved and ensure a higher degree of relevance of information to individual users. This makes for better searching outcomes and reduces the quantity of retrieved information.</td>
</tr>
</tbody>
</table>

**Table 13 Recommended features and functionality for a HIP**

If the designers and sponsors of health information portals want to attract a wider audience, in particular those under 30, and draw health consumers away from search engines such as Google then attention has to be paid to the elements that improve users’ search experience. As argued by Kunst et al (2002) the Internet has “the potential to facilitate but also to jeopardise health care provision”. How HIPs are designed particularly how easily users can search and find relevant information is critical. Including key features and functionality such as an ontology or thesaurus, differentiated information access and spell checking are important for ensuring good search results and a good search experience. Until the design of HIPs incorporate what is needed to improve the search experience health consumers will continue to be dissatisfied with their searching results and the benefits if HIPs will not be fully realised.

**References**


Data stakeholders interacting with patient data

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0545.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Conceptual modelling, Health information systems / medical record systems / care delivery /, Human computer interaction (HCI), IT artifact</td>
</tr>
</tbody>
</table>
DATA STAKEHOLDERS INTERACTING WITH PATIENT DATA

De la Harpe, Retha, Cape Peninsula University of Technology, 80 Roeland str., Cape Town, South Africa, delaharper@cput.ac.za

ABSTRACT

How data stakeholders interact with data is considered in this paper by incorporating a socio-technical perspective. A data interaction enactment model is proposed to better understand how a logical structure is enacted when data stakeholders interact with data. The proposed model expands a technologies-in-practice view to also provide for the embodied structures of data objects. This extended technology-in-practice view provides designers with an approach that focuses on the interaction between data stakeholders and data objects. Although the proposed model is a generic model it is applied in the healthcare context for this study. An interpretive approach was used for a private healthcare centre as the case to investigate how healthcare professionals interact with patient records. The findings indicate the design of patient records based on the perceived anticipated use may not sufficiently provide for their actual use where different stakeholders need to interact with the patient records. An understanding of the relationships between designers, data stakeholders, the patient record structure enacted during interaction and its physical components is important. This may be one of the reasons why patient records are still mostly paper-based in spite of efforts to develop and adopt electronic patient record systems.

1 INTRODUCTION

Organizations find it difficult to determine the value of their data, an important organization resource today and the abundance of data and increased sharing of it within and between organizations have amplified the problem (Bertolazzi & Scannapieco, 2001, Loshin, 2003). A practice related study has confirmed this problem (Eckerson, 2002). The objective for this study is to determine how data stakeholders interact with data by considering how they respond to each other during the period of interaction. This paper specifically deals with the interaction between the persons using data and the technology on which it is stored and which facilitates its processing. The focus of this paper is on a patient record representing the patient who is in need of a healthcare service.

This research is in the information systems field and the interpretive approach used provide for its socio-technical nature. The problem this paper addresses is that the perceived anticipated use of data objects is an inadequate means for designing data objects ultimately leading to data that does not sufficiently support the needs of the intended users. In order to address this perceived misalignment between designed anticipated versus actual use it is necessary to better understand the interaction between the designers, users and their perceived observations of real-world objects and data objects as the resulting presentations. The data object is the result of the interaction between the designer and user based on their perceived observation of the real-world object and the anticipated use of the data object. It is necessary to understand exactly what happens when the user interacts with a data object. A data interaction enactment model is proposed to extend Orlikowski’s (2000) technology-in-practice model. This is an attempt to address the concerns of Hanseth et al. (2004) that her model does not sufficiently address how technology-in-practice is shaped by the technological artefact as well as the relationship between the two and her responding suggestion to better understand the problem of agency by also considering the status of the object. Data stakeholders are regarded as social actors and data objects as conceptual artefacts because it is not yet clear how these objects are instantiated from their design templates in practice every time data stakeholders interact with data.
The objectives of the paper is twofold, firstly to propose a data interaction enactment model based on the research findings of other similar studies that can be used as a theoretical lens to analyse the interaction of data stakeholders with data objects; and secondly to apply the proposed model to a healthcare case. The purpose of the empirical part of this paper is to illustrate how the proposed model can be used in practice. Its validation is at this stage outside the scope of the paper since a more extensive empirical study is required.

The layout of the paper is according to the two objectives and the first part represents the logic that guided the conceptualising of the thoughts, constructs, reasons and structures of the different concepts of the proposed model presented in Figure 1. The focus of the paper is on the enacted structure indicated by the circle in Figure 1. The literature review did not reveal any appropriate model that could be used and the literature findings were instead conceptualised as the proposed model. In the first part of this paper the different views of data is discussed in terms of data stakeholder as the person using it and the data object as the logical unit for a patient’s data. Next the different views of technology in information systems are discussed with a specific focus on structure, agency and inscriptions to provide for the socio-technical nature of the study. This part is rather lengthy to give justice to the complexities around how humans interact with technology. It is concluded with a proposed model of how data objects are enacted during interaction. This model incorporates technologies-in-practice as the data object is constructed during interaction of data with the anticipated use inscribed during its design and the actual use inscribed in the actual data. The next part considers a healthcare case and in the discussion the proposed model is applied. In conclusion the usefulness of this model to investigate how data stakeholders interact with data objects is discussed.

2 INTERACTION IN SOCIAL AND TECHNICAL WORLDS

Persons using data have an interest in how it is defined and how they can use it. A data stakeholder is a person who has an interest in data and examples of data stakeholders are management using data for decision-making; a salesperson; a doctor, etc. Although stakeholders are usually humans, it is possible that a stakeholder is another organization, a department, a hospital, etc. The term data stakeholder is used as a generic term to describe any party interested in data regardless of the context. An external party is an outside person or organization that shares organizational data with an organization. All the data stakeholders interact with data in the roles of data producer, consumer or custodian (Strong et al. 1997). In the healthcare context data stakeholders are healthcare professionals and other persons and parties who need to interact with patient records.

A data object can be regarded as a logical record that is constituted of one or more physical records/parts. Data objects -data representations of real-world objects -are generic templates of the different organizational objects obtaining their values from data sources when instantiated through how they are used. Data objects are identified and defined during systems development by systems developers as they perceive the real-world object as well as the needs of the users (data consumers) of the data objects. The representation of real-world objects, such as for example a patient, is not limited to its representation as part of an information system, but also includes all data generated during the interactions of the organization with the real-world object, e.g., the patient. Not only is the healthcare organization interested in the best match of the data object with its corresponding real-world object, but it is also interested in the history of that data object, i.e., snap shots of all its states. This history is complete when it includes copies of all the different states as well as copies of all the interactions that lead to the changes in the states. Wand and Wang (1996) suggested the following states when considering the mapping of a real-world object, in this case the patient, to its IS data representation (patient record): proper, incomplete and ambiguous representations and meaningless state. These states will influence how the records are used in practice in terms of how well they were designed.

The healthcare sector is typically divided into a public and private sector and in some cases a mixture of the two. Although, for example, the primary goal of medical practices, hospitals and clinics is to provide a healthcare service to their communities, they also have to operate as businesses. Patient data
is used by healthcare professionals when providing a healthcare service for their patients whilst it is also used to manage the different healthcare enterprises.

The focus of this study is on patient data which to a large extent is still paper-based. There are different types of patient records but for the purpose of this study a patient health record is considered. An electronic health record (EHR) includes the patient’s health profile, behavioural and environmental data over time to represent different episodes of care by various providers (WHO, 2006). Fitzpatrick (2000) emphasises that a health record does not only contain data about a patient and the care of patient but should rather be regarded as an integral part of the practical “doing” of healthcare and as a record “at work”. She further argues that it is a working record evolving around a complex collection of related forms, papers, documents and records embedded in the social, spatial context of a healthcare enterprise.

For the purpose of this study technology as an infrastructure is regarded as pervasive and present, but data objects as technological components need to be specifically considered. Although data objects are non-human objects, they are socially constructed by humans, namely, designers and data stakeholders and it is necessary to consider the position of data objects in terms of the technical and social worlds.

In much of information systems (IS) research the focus is unavoidably on the interaction between social and technical systems. An improved understanding of the phenomena that emerges when these systems interact will therefore enrich IS research (Rose et al., 2004). The human-artefact relation is viewed by Widjaja and Balbo (2006) through a macro-level lens as human and artefacts are both social products as well as social makers in shaping and remaking each other. Technology is social in its origins as well as in its implications in that it does not influence human agency by imposing a single and mechanical functionality but by “inviting specific courses of action” (Kallinikos, 2002). Technology is not used within a vacuum but always involves social actors where technology shapes its users and at the same time responds to how it is used. Hosein (2002) responded to the concern of technology being “pushed into the background” by suggesting that researchers should, when relevant, attempt to understand the technological within their research.

While the IS researcher is neither in the social nor in the technical worlds, but rather where these two worlds overlap, it is possible to move beyond the properties of each world and to consider the relationship between them. It is then necessary to study how the human, from the social world, interacts with the technical components of the technical world when and how these two worlds become one. The definition of IS already provides for this boundary crossing between the two worlds.

2.1 Information Technology

Technology can be regarded as practical or useful, rather being an end in itself. It includes all the tools, techniques, materials, and sources of power that humans have developed to achieve their goals (Rose & Scheepers, 2001). Technology is not only objects or technical artefacts, but also ideas (object-oriented paradigm), practice (testing), etc., and is socially constructed and should not be investigated according to the laws of natural sciences (McMaster et al., 1998). Information technology (IT) is the acquisition and processing of data to support human processes and can be regarded as a material resource that supports information practices and such information practices in turn support a wider set of social practices (Rose & Scheepers, 2001). Information technology is the product of human agency and will reflect the structures of the social system that designs and manages it as well as interpretations of its anticipated use. Information systems (IS) can be regarded as a social system of information practice that is supported by information technologies (a material resource) and representation the essence of IS according to Weber (2004).

When the emphasis is on the interpretation of technology use then this could lead to a detachment of the IT artefact (Widjaja & Balbo, 2006). They base their arguments on the practical knowledge in dissecting and constructing artefacts of human-computer interaction research (HCI) as opposed to the socio-technical research that views human-artefact relations with both human and artefact as social
products and social makers in shaping and remaking each other. Monteiro and Hanseth (1996) argued that it is necessary to move beyond the “IT enables/constrains” position and they base their arguments on the departure point of IT being the crucial factor, simultaneously enabling and amplifying the trends for restructuring organisations. They felt the necessity to focus more on the interplay between IT and organisations by reaching a deeper understanding of this interwoven relationship as to how IT shapes, enables and constrains organisational changes. It is necessary to explain why a group of diverse people, e.g., data stakeholders, reach a single accepted agreement, based on an interest resulting in an inscription into material form.

Technology artifacts are made up of multiple components and are not fixed or independent with their stability being conditional. Interrogating interests and actions provide moments within a discourse where the actors may take form. This may particularly be true for objections, resistance to interpretation or obdurancy. Kallinikos (2002) suggests that “an essential part of the conditions underlying human agency are given expression and shaped through the very organization of cognition-based artifacts and the procedural standardization of their construction, interpretation, and utilization”. He bases this understanding of the relationship between humans and their artifacts on the humans’ ability to objectify and organize their experiences by means of various systems of notation and symbolic codification. A person may initially be constrained by a technological artifact such as, e.g., a mobile phone, but with repeated use may become more comfortable using it as the device’s properties become more familiar. Although the user as an agent interprets the artifact’s enabling capabilities each time it is used resulting in an improved understanding of how it can be used, the shaping effect of the technological artifact on the human cannot be denied. The human initiated interaction with the technological artifact and during this agency process displays intentions and has the ability to interpret whilst the technological artifact’s “machine agency” has the capacity to make a difference.

2.2 Structures

The view of technology as structure with the ability to enable and/or constrain human actions was contradictory to Giddens concept of structure as “traces of the mind” and with no material existence. This created a debate about the agency ability of technological artifacts as well as different interpretations of structures (Akrich, 1992; De Vaujany, 2005; Hosein, 2002; Kallinikos, 2002; Orlikowski, 1992, 2000, 2005; Orlikowski & Iacono, 2001; Rose, 1999; Rose et al., 2004, 2005; Widjaja & Balbo, 2006). Agency for both humans and technological artifacts plus the concept of structures as these apply to this study are discussed next.

Technology-in-practice is proposed by Orlikowski (2000) to bring her original structurational perspective of technology in line with Gidden’s view of structure as being only constituted when users interact with technology. Such a structure is enacted by humans only for the duration of that interaction, thus only fragments of the mind. Hanseth et al. (2004) argue that such an approach still does not address how the technology-in-practice is shaped by the technological artifact as well as the relationship between the two. Orlikowski (2005) responded to this criticism by agreeing that “by privileging either the technology or the social, we lose sight of their intermingling”. She suggests that further research should help with a better understanding of the problem of agency and by specifically looking at the status of the object. Instead of referring to material agency they prefer to use the term “material performativity” as opposed to “human agency” and these terms may be more useful to recognize the power of both without equating them. Both terms are implicated in the other since human agency could be materially performed whilst material performances are always enacted by human agency.

When humans interact with technology, a technologies-in-practice structure is enacted for the duration of the interaction. Humans use facilities, norms and interpretive schemes in response to the technology’s enabling/constraining properties to make sense of how to use the technology. This is
done by enacting a temporary structure based on rules and resources. This however does not provide for details about the relationship between the technologies-in-practice and actual technological artifact.

Action and structure cannot be separated and structure is both the means and the constraint for action (De Vaujany, 2005). Rose et al. (2004) refer to structure as “…to the common enduring pattern in social interactions and linguistic and semiotic discourse in which those interactions primarily take place”. They use Wand and Weber’s (1995) three structure types, namely: physical as the technological implementation of the computer system; deep as reflecting the meaning or underlining rule set of how the computer system represents the real-world system; and surface as the way the system represents itself in the form of interface, inputs and outputs to the user. The deep and surface structures form the basis of the linguistic perspective.

They concluded their paper with the statement that by “understanding how the deep and surface structures of social practice and discourse are consciously and unconsciously embedded in material computer systems, by the actions of designers, and subsequently interpreted by other stakeholders and reincorporated in the production and reproduction of linguistic and social structures continues to be one of the most challenging areas of inquiry in the discipline” (Rose et al., 2004). They further quoted Mol (2002) who refers to the practices in which an object is manipulated as not a single passive thing in the middle “waiting to be seen from the point of view of seemingly endless series of perspectives. Instead objects come into being – and disappear – with the practice in which they are manipulated”.

Burton-Jones and Grant (2008) adopt positivist assumptions for their study to investigate how IS can be used effectively and build their theory on representation theory. They propose constructs such as fit-in-use and transparent interactions, amongst others, as constructs in their model for effective system usage. They consider structures as the basis as to how IS provides its representations (deep structures) via surface and physical structures.

Inscriptions are discussed next to indicate how designers inscribe their understanding of how data stakeholders will interact with data and how data stakeholders inscribe their actual use of data objects in practice.

2.3 Inscriptions

Inscriptions are attempts to inscribe pattern of use or behavior in for example artifacts, work routines, legal documents, standards, procedures and other institutional or organizational arrangements (Monteiro & Hanseth, 1996). They further state that inscriptions may have different forms and it may increase understanding of how these artifacts, etc., are used when investigating how and where “patterns of use” are inscribed and the first step would be to study how the users’ interests are translated and inscribed. Callon (1986) describes inscription as a “process of creating technical artifacts that would ensure the protection of the actor’s interest”. It is important to note that an artifact never begins as a blank slate and in fact the inscriptions of IT artifacts represent the design, construction and use of different stakeholders, all with their own interests, values and assumptions in variety of communities of designers, users, investors, etc.(Faraj et al., 2004). The technology artifact can be modified through use with patterns of use inscribed and such inscriptions can then be viewed as properties of the artifact.

Inscriptions as the “way technical artifacts embody patterns of use” do not imply that actions are “inscribed, grafted or hard-wired into an artifact”, but rather that artifacts are “interpreted and appropriated flexibly”. During software development the “programs of action” for the users are inscribed in the data object based on the perception of how the data object will be used and this then anticipates how the data stakeholders interact with the data object in their different roles. Assumptions about the required competencies and roles of the data stakeholders are made and these roles and competencies are delegated to the components of the socio-technical network. Technology, having programs of actions inscribed, is an actor imposing its inscribed program of action on the data stakeholders. It can for example “resist” responding in the way expected by the data stakeholder such
as when it does not allow a value to be entered based on its validation criteria. Data stakeholders may also use data objects in unanticipated ways by, for example, bypassing validation criteria (enter 1111111 for a telephone number) or finding creative ways to enter data values in other fields such as entering a comment in an address field, etc.

The designer inscribing the perceived anticipated use for technology and as a caution Kallinikos (2002) warns that this should never become the arbiter of technological malleability. He further states that the spectrum of embodied intentions that characterize a particular technology cannot represent all possible embodied intentions that the user can enact when interacting with technology. In fact it may be important for the designers to consider the assumed user versus the actual user, real-world object versus its representation as a data object. Orlikowski (2000) states that the use of technology is situated, but not totally open to any and all possibilities. There will always be boundary conditions when using technology artifacts because they have physical properties that may limit their use. This does not imply predictable responses of human action (Widjaja & Balbo, 2006). Data stakeholders inscribe their understanding of how data can be used during their daily activities. The problem could then be that these understandings may not be recorded and that the understanding is reflected by the outcome of the action that is then inscribed, e.g., by how they produce or consume data to compensate for quality problems. Understanding is the outcome of the internalization process where different data inputs are combined with the data stakeholders own knowledge that (s)he has gained through experience as to how a similar situation was handled before. The inscriptions are not only based on the data stakeholders’ own competencies and assumptions, but may be influenced by factors of the environment in which the data stakeholder operates, as well as by other actors of the network. Data stakeholders may influence each other by convincing each other of their interests and how that may benefit them. All of these actions may be inscribed as patterns of use in the data object.

Next the design and use of data objects to provide for the data stakeholders technologies-in-practice is discussed.

2.4 Design and use of patient records as data objects

The designer, in this case the software developer, inscribes the beliefs, social and economic relations, previous patterns of use, legal limit and assumptions of what the artifact is about when defining the data object (Akrich, 1992) quoted in Faraj et al. (2004). Once the data object is created with its inscriptions its structure becomes material and these “embodies” structures represent various social rules (Orlikowski, 2000). Data in its different forms, e.g. records, database tables, reports, emails, etc., can be regarded as artifacts and as an artifact it has an embodied structure (refer to the physical structure) that includes the patterns, rules and resources that are embedded in the data object during its design as it may be used by data stakeholders (Widjaja & Balbo, 2006). Design is also regarded as the social construction of the data object’s embodied structure that emerges from their perceived understanding of the real-world object and through communication with the data stakeholders. The embodied structure could also contain inscriptions providing for data quality such as “rules” that are coded as software programs doing data validation or the use of triggers on databases. The design output (technology artifact) of the data object is referred to as a data object template and has inscriptions of the data object’s anticipated use. It is important to note that a data object does not refer to a single artifact as Jensen (2004) refers to an electronic patient record as “a multiplicity of things which forms a whole only sometimes or for some purpose”.

It is important to note that an embodied structure, i.e., the data object as an artifact, is a passive object that at the most has the potential to enable or constrain action but only when it is used in practice. By itself the data object cannot initiate any action or displays agency and an embodied structure can therefore be regarded as potential structuring elements emerging through the recurrent use of the data objects by humans. This in-use structure that is constituted during the process of agency, when data stakeholders interact with data though the data object structure, refers to the enacted structure of data objects. The same data object may be enacted differently by different data stakeholders and therefore
the appropriation of the data object can be viewed as “a weave of embodiment in-design and enactment in-use” (Widjaja & Balbo, 2006).

Data objects represent not only its physical structure, the template(s) defined by software developers during design, but when a data object is instantiated for each customer it is populated with the data representing the details of that customer. This is when data stakeholders capture or modify data. One data object template is instantiated by several data object instantiations when populated with data values to represent, for example, different customers and this multiplicity aspect requires a different view of structures and inscriptions as opposed to other technological artifacts. The result of this is that there are multiple instances of the data object, no longer necessarily being processed together but each could be enacted separately during its use. The data values can also be regarded as inscriptions because such values are captured or modified according to the data stakeholders’ perceived understanding of the data values during interaction with the data object.

The logical data object is constituted from the different physical parts during the process of agency and is, in that form, only available during the duration of that process.

The above discussion is illustrated next by considering patient records as data objects.

2.5 Patient records in practice

Patient records are used in healthcare for coordination and cooperation of a large collection of healthcare stakeholders as large and complex networks (Berg & Bowker, 1997; Hanseth & Monteiro, 1997; Fitzpatrick, 2000). Healthcare records, in this case patient records, are configured according to the needs and practices of the organisations they serve. The patient health record evolves into different views that are both independent and interdependent as the details about the patient’s care are recorded. In a medical practice the patient record is used by individuals rather than large groups as in hospitals.
Government regulations and Industry standards (e.g., code systems) are incorporated in how these records are used. The nature of what and how data is captured and used in patient records reflects the highly subjective judgments of their users, e.g., medical doctors, nurses, etc.

There is an interaction between the patient and the healthcare professional at the different points of healthcare therefore representing a network between the patient, healthcare professional and the corresponding data object of the patient. This interaction results in the exchange of data and information between the different data stakeholders.

The healthcare professional interacts with both the patient and patient record, contributing values to the patient data record based on his/her perception of the patient’s condition. This means that the medical practitioner is engaging with the patient’s data, integrating data from both the patient and patient data object, internalising these facts together with his/her observations, perceptions and professional knowledge. Data obtained from the patient are the facts provided by the patient and observation of the patient’s condition. Data from the patient data object are the facts previously recorded and stored, i.e., inscribed. This information is internalised and communicated by the medical practitioner to the patient and captured and stored, i.e. inscribed, in the patient data object in the form of clinical data. The patient receives the information which has to be processed and internalised by the patient in order to understand its meaning. Once the consultation is over the link between the patient and medical practitioner network is broken with the healthcare practitioner, patient and patient data object becoming inactive nodes of the network in terms of caring for that patient.

The logical patient record refers to the enacted structure temporarily constructed by the healthcare professional when interacting with both the patient and physical patient record at the point of care. The outcome of this interaction is inscribed in the actual patient record as the data recorded by the healthcare practitioner to represent the observations, diagnosis, treatment and other details. The physical patient record can be the patient file, notes, results, etc. One example is the prescription with the medication details that the patient takes to the pharmacy for dispensing of the medicine. All these physical components constitute the patient record even though it is possible that the individual parts may be physically in different places. This combination of components represents the real-world object – the patient. The temporary structure, the logical record, contains only the data necessary for the interaction between the healthcare professional and the patient’s data. This can be constrained by what data is available as the result of insufficient design where the anticipated use was not sufficiently perceived and or translated. It is also possible that the healthcare professional does not record all the interpreted data, such as diagnosis, observations, treatment, etc., resulting in incomplete inscription of data values. As a shared data object this means that other healthcare data stakeholders may have insufficient data to treat the patient and this can be regarded as a data quality problem.

Designers have a smaller role in the case of paper-based patient records because the structure and layout of the record evolves through its use rather than being designed first in terms of its anticipated use. Paper-based records usually provides for some structure combined with free text and are typically already in the format required for the type of document, e.g., X-rays, ultra-sound pictures, EEG, etc. The situation is different for electronic patient records because the systems supporting electronic patient records require a proper investigation, planning and design in terms of its anticipated use before such records can be used successfully. This is a complex task and research dealing with issues around electronic patient records should not only concentrate on identifying obstacles and barriers but also on exactly how such a record will be used in practice. In the case of electronic patient records the patient record has a pre-defined template to provide for all the possible data required to constitute the required patient record. This template can and probably will have more than one template components to provide for all the different parts of a patient record. This requires the designer sufficiently to understand the anticipated use of different healthcare data stakeholders to integrate it with the mapping of the real-world object. These templates guide the healthcare practitioner during the interaction to capture and use the data according to the specifications.
Even though patient records should not be regarded as passive repositories of data, it is true that the record is only active during the interaction between the healthcare practitioners and patient and then only for the duration of the interaction. Thereafter it is passive and cannot contribute anything to the patient’s care unless it is used by a healthcare practitioner. It is possible that the arrival of another part of the patient record, e.g., lab results, could result in the patient record becoming active again, but then only if a human attends to the arrival of such data. The patient record cannot alert the healthcare practitioner to something that may be unnoticed by the healthcare professional when it could in fact be significant for the patient’s care, e.g., unrelated observations on different parts that could in fact point to a specific condition or development. The patient data is as good as it is interpreted and used by the healthcare professionals.

Although the different parts of the patient’s record can be used independently, the significance of the interdependencies on each other has to rely on the healthcare professional’s ability to notice these interdependencies – the more experienced they are, the better they are equipped to interpret the patient’s data sufficiently to provide effective care. The data at this stage does not actively contribute to these interdependencies because it does not have the ability to initiate an action unless a procedure has been incorporated to process the interdependencies of data from time to time.

3 HEALTHCARE CASE

HealthCenter is a healthcare enterprise located in Cape Town, one of the fifteen such centres planned to operate in the private healthcare sector in South Africa by the end of 2008. They believe that an informed society is a healthy society and aim their services at accessibility and convenience. It provides a number of healthcare services under the same roof that cover the whole spectrum from sickness to wellness and primarily targets upper class citizens. Their patients either belong to a medical aid or can afford to pay for their healthcare services. The sub-acute hospital bridges the gap between hospital services and home-care for those patients who still require medical care on their way to recovery and who can no longer be treated in hospitals.

In the sub-acute and rehabilitation hospital a separate patient record is used where additional information is captured. This information supports the care period during the time that the patient has been admitted to this hospital until the patient is discharged. Most of these forms were designed by the current sister in charge based on her previous experiences at other hospitals. The main components of the ward patient record, also kept in a patient folder, are: discharge planning ward round; the FIMS (functional independence mobility score) form; admission form; medical admission record; nursing progress report; referral letter; medication administration chart; and vital signs chart.

3.1 Research methodology

This is an interpretive study in the information systems field with a relativist ontological stance and a subjectivist understanding of the research objects’ behaviour. A literature study was conducted to establish the different issues, themes etc. of the research topics. An exploratory case study was conducted with the group of data stakeholders and data objects as the units of analysis. Empirical data was collected through unstructured interviews where probing questions were asked and documentation and forms studied. Representatives of the following stakeholder groups were consulted: patient; management (centre manager, patient care manager and financial manager); administration (reception); healthcare professionals (doctors, nurses); and developers. Patient records were considered as the data objects. The approach was not to unpack all the possible uses of the different stakeholders but to rather understand how the different groups interact with patient records. The

1 Pseudonym used for the healthcare case
empirical data was analysed using the proposed data interaction enactment model as the theoretical lens as well as by identifying common themes and categorising data. Questions were derived from the proposed model and translated into suitable questions that were used during the interviews. Responses and observations were recorded and organised according to identified themes. The findings were further interpreted by establishing the meaning of the findings based on hermeneutics. The size limitation of this paper restricts the reporting of the case study in extensive detail and only the parts relevant to the interaction between data stakeholders and data objects are discussed. The data interaction enactment model is part of a more comprehensive research framework that also consider details about the influence of the context, profiling of the data stakeholders and the approaches they use during the interaction, roles, available resources and attributes. These details were considered but are not reported in this paper due to size limitations.

3.2 Analysis

HealthCenter’s patient data is still mostly paper-based and although electronic options could be made available the practice doctors and nurses in the sub-acute hospital the healthcare professionals still prefer paper records. This is in spite of patient folders that are often miss-placed. The business side of HealthCenter uses an electronic IS with effective appointment booking, medical aid claims, financial and other systems that support their business processes. Management has access to summary reports that they can use for decision-making but need easier access to be able to manipulate data for improved decision-making. The doctors indicated that their relationships with their patients are more important and their knowledge of their patients or family dynamics (not captured in patient records) may influence their diagnosis and treatment plans. The case notes based on the consultations are used differently by different doctors, e.g., for one doctor comprehensive and detailed notes are very important whereas another doctor feels that a template could deal with most of the standard observations. Data stakeholders adapted their use of patient records by inventing communication mechanisms, e.g., adding a red dot to folders of patients with outstanding accounts; blue message slip with phone messages to folders for a doctor’s attention; and using pink A4 paper sheets for female patients. There is limited sharing of patient data and this is even the case between the different care facilities of HealthCenter and doctors use referral letters to refer patients to other doctors or facilities. A new record, for example, is created for a patient admitted to the sub-acute hospital where all the patient details are again captured even though the patient is an existing patient of HealthCenter with a patient file at the doctors’ practices. There is still much duplication and disparate data sources at the different facilities.

When a patient is admitted in the sub-acute and rehabilitation hospital the nurse’s observations become important to manage the care of the patient. The nurse is the primary care giver and has direct contact with patients over a period of time. Patient records are stored and updated and then kept for a period of five years. Only the following patient details are captured on the computer: medical aid; payment; and medical condition.

The difference between patient records used by doctors and nurses reflects the purpose for which these records are used. The doctor interacts with the patient and/or patient record only for a short period during the visit in order to observe the patient’s condition and/or progress. The doctor may then decide to intervene by describing or changing medication, treatment; request for tests, e.g., blood, x-rays, etc. Nurses interact with both the patient and patient record over a longer period while the nurse cares for the patient. Nurses are therefore in a better position to become more familiar with the patient and the data that reflects their observations during the care period, in terms of the patient record, should be a more realistic representation of the patient. Typically, nurses data is not captured electronically and the value “hidden” in the data could become lost in the treatment unless the nurse and doctor specifically respond to observations recorded that may be relevant to the patient’s condition and treatment.

Following are the findings derived from the empirical data collected from Healthcenter:
• Different data instances of the same patient are located at different healthcare centres resulting in a lot of duplication and disparate sources of patient data.
• Patient data is important to all data stakeholders and although business data stakeholders interact with patient data to support an effective healthcare service from a business perspective, healthcare professionals still have the final say over how they will interact with patient data which at this stage is still secondary to their person-to-person interactions with their patients.
• Patient data used by healthcare professionals is still mostly paper-based and they are not specifically concerned about data quality except to provide for a comprehensive medical history for each patient in order to provide an appropriate healthcare service according to the legal requirements of their profession. Sharing of patient records is currently limited with referral letters the only mechanism to share patient data.
• Data inscriptions depend on how the data stakeholders interpret the use of patient data and although standard formats are used, e.g., patient observation cards, the actual values may be recorded and/or captured differently.
• Data is designed and used for two separate purposes, namely, business and healthcare. The data stakeholders are interested in using patient data for their purpose and may not recognise the benefit that they can derive from a design and use of patient data for a single purpose that supports both healthcare and business activities.

3.3 Discussion

When designing patient records, paper-based or electronic, designers need to consider how data stakeholders interact with patient data. This cannot simply consist of a list of the perceived user requirements but rather should be based on a sufficient understanding of the complexities associated with how the data stakeholders interact with data. The user interfaces to aid data stakeholders with interacting with patient records should be flexible enough to provide for a range of activities; for the different levels of expertise of the different data stakeholders; for the different purposes and ways that healthcare professionals interact with patients; to accommodate new, different requirements or enhancements. It is essential for designers and developers to understand how the different logical patient records are enacted for the different healthcare purposes according to the different healthcare professionals’ needs. Only then can the different patient record components be designed to support the different enacted logical patient records in practice. At this stage it seems as if most effort is spent on designing for anticipated use by defining the physical patient record (template) rather than focusing on the different enacted structures (logical patient record) to better provide for the different uses in practice.

The importance of the HealthCenter findings indicates that patient data is designed for two different purposes, namely to operate HealthCenter as a business and to provide a healthcare service but not for different uses in practice. Data stakeholders are only interested in using data for their own purpose and may not recognise the benefit they can derive from a design that can support more diverse activities. Design for use in practice should take into account the complexities of human interacting with technological artefacts by providing how different stakeholders interact with data at different times, at different locations and for different purposes. This cannot simply consist of a list of the perceived user requirements but rather should be based on sufficient understanding of the complexities associated with how data stakeholders interact with data. User interfaces, templates, records, reports, are all mechanisms designed for data stakeholders to interact with data and these should be flexible enough to provide for a range of activities; different level of expertise; and different purposes.

The socio-technical nature of IS should specifically be considered when designing for actual use based on representing the real-world in a way that effective facilitate actual use in practice – specifically its human-agency and Orlikowski’s (2005) machine-performativity as well as how data objects are enacted as “moments-of-interest” (Hosein, 2002). Patient records are shared between different data stakeholders, but in different ways as large and complex networks (Berg & Bowker, 1997; Hanseth &
Monteiro, 1997; Fitzpatrick, 2000). This is confirmed by the empirical study that emphasized the complexities around the interaction between data stakeholders and data objects where the actual use differs from the anticipated use. The reason why this is still the case could be the lack of understanding of how logical data objects are enacted as logical data objects and how these differ from the design templates (physical data objects). The empirical results of the healthcare case confirmed that a patient record currently still constitutes of many different physical components that are stored in different places and the concept of a logical record is not clear since users enact the required temporary logical structure with those components that they can locate with different levels of success and frustration. This enacted structure is also only available during the moment of interaction and needs to be enacted from scratch each time it is needed. Since these are still mostly paper-based careful consideration is required to provide for the same facility in any electronic patient record systems.

There is an important difference between the physical and logical structures of a record. Different temporary logical records are enacted to support different needs, i.e., it is necessary to provide for flexible logical records to enact the logical structures from the different physical parts as and when needed. In order for HealthCenter to compete as a business in an ever-changing global environment, they need to consider external factors that have an impact on the environment in which they operate. Better utilisation of their patient data will enable HealthCenter to protect themselves against legal actions and to improve their Healthcare service delivery. HealthCenter will be in a better position to interact with their partners, both internal and external, by recognising the importance of sharing patient data and this can be best achieved when their systems and patient record designs provide for their stakeholders to operate in a global world.

4 CONCLUSION

The purpose if this paper was to consider the complexities associated with humans using technology and to propose a model to consider how data stakeholders interact with data objects. This model illustrates the temporary enactment of the logical data object during the interaction with data objects. The problem seems to be the difference between the design for anticipated use and the actual use of these data objects. Data stakeholders inscribe their understanding of how the data object is used and these understandings are then inscribed into the actual data objects. The proposed data interaction enactment model can be used to investigate how logical patient records are enacted during interaction by considering which components are required to constitute the logical patient record for the different purposes. The proposed model indicates the extension of the technologies-in-practice model to provide for the embodied structures. The usefulness of the proposed data interaction enactment model still needs to be validated in practice to establish if this extended view sufficiently provide for improved design of the patient record components. Only once the data interaction enactment model is operationalized will it be possible to evaluate its usefulness as a practical model to support better design for actual use in practice.

The model built by Burton-Jones and Grange (2008) for effective usage provides for several constructs to consider for effective system usage, e.g., fit-in-use, transparent interaction, etc. Although their model also considers the three structures of Wand and Weber (1995), they focus more on to what extend the structures in terms of representing the users’ world, support effective usage as opposed to the data interaction enactment model that may provide more insights on the actual interaction between users and data as part of an IS. On the other hand it may be necessary to specifically consider representation theory for the proposed model to focus more on the representation aspect of the model, i.e., how the designer inscribe the perceived observations of the real-world object and anticipated use as the data object template and how that is instantiated to represent the real-world object (refer to the parts outside the circle in Figure 1). In this paper the focus is on the enacted structure indicated by the circle in Figure 1. It may be possible to combine some aspects of both models in which case further
research is required. Both models are at a formative stage and need to be operationalized and validated.

At this stage the contribution is more theoretical where the data interaction enactment model provides an explanation of the difference between a temporary enacted structure (logical data object) versus an embodied structure (data object template). The conceptual model was derived from the different explanations and descriptions to present a generic representation of the concepts on an abstract level. It is an attempt to visually represent the thoughts, constraints, events, reasons, behaviour or structures of the specific concepts. The interpretive research approach used contributes towards the improved understanding of the interaction between social actors and technological artefacts as the latter are designed to represent real-world objects.

The findings confirmed the notion that patient records are complex and that a consideration of both social and technical considerations are necessary, especially when these overlap when data stakeholders interact with data objects. Designers should focus more on the use of patient records in practice rather than only on their anticipated use. The contribution of this paper is more conceptual although the empirical results confirmed the issues identified by the literature review. Further research is required to translate these findings into practice design recommendations.

REFERENCES


http://www.dis.uniroma1.it/~monsan/ResearchActivity/Articoli/IQ2001_BS.pdf [7 November 2002]


http://www.tdan.com [7 August 2007].
Rose, J. (1999) Towards a structurational theory of IS, theory development and case study illustrations  
40(5):103-110.
27(2):iii-xi.
“WHO IS IN CHARGE AND WHOSE RULES ARE FOLLOWED..?”: POWER IN A INTER-ORGANISATIONAL IS PROJECT

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0253.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>interorganizational systems, interpretivist research, Power, Grounded Theory</td>
</tr>
</tbody>
</table>
“WHO IS IN CHARGE, AND WHOSE RULES ARE FOLLOWED…?” POWER IN AN INTER-ORGANISATIONAL IS PROJECT

Hekkala, Riitta, University of Oulu, P.O. Box 3000, FIN-90014 University of Oulu, Finland, riitta.hekkala@oulu.fi
Urquhart, Cathy, University of Auckland, Owen G Glenn Building, 12 Grafton Road 1142 Auckland, New Zealand, c.urquhart@auckland.ac.nz
Iivari, Netta, University of Oulu, P.O. Box 3000, FIN-90014 University of Oulu, Finland, netta.iivari@oulu.fi

Abstract

This interpretive grounded theory study describes and analyses major power issues in a Scandinavian inter-organisational IS project which spanned four organisations, two suppliers, one national organisation and a research organisation. The paper explores different dimensions of power that emerged during the project – sources of power, power as resistance, reasons for the power struggle, and power as exercised. We conclude that, while power issues in ISD projects are by no means a new phenomenon, these power issues were exacerbated in this IOIS project due to the project governance structures necessary to coordinate several organisations.

Keywords: Power, inter-organisational IS project, Glaserian grounded theory

1 INTRODUCTION


It has been argued that the nature of IS development is always conflictual and political, and that researchers, instead of defining better methodologies and accepting manageralist agendas of IS development, should carefully analyse this conflictual and political context (Howcroft & Wilson 2003, Silva 2007). Silva (2007) has highlighted a lack of research on the politics of IS projects, and the challenges inherent in the belief that politics are ‘dark’ or illegitimate and somehow unsuitable for study. It has also been argued that academics should focus on dominance, power, marginality, and exclusions that take place both in IS development, adoption, and use (Beck 2002). Silva (2007) has argued in favour of an interpretivist approach for studying power and politics in IS, and this research fits into that category.

Conflicts and power struggle between users and IS professionals, and between users and managers, have been studied extensively over the years. In those studies, managers’ power over users has often also been criticised (Alvarez 2002, Hirschheim & Newman 1991, Howcroft & Wilson 2003, Kirsch & Beath 1996, Markus & Bjørn-Andersen 1987, Sarkkinen & Karsten 2005, Symon 1998, Yeh & Tsai 2001).
In the twenty-first century, inter-organisational projects have become much more common as a consequence of globalisation and standardisation in information systems. So far, inter-organisational information systems (IOIS), and especially their implementation with several stakeholders, have received only minor attention in the IS research (Evaristo, Scudder, Desouza & Sato 2004, Salmivalli 2008). Because there is even a larger number of stakeholders involved in inter-organisational IS projects, there is an even greater potential for conflicts (e.g. Kumar & van Diesel 1996, Levina 2005). In contrast to traditional problems in power between developers and users, our research considers the power issues between many more actors and parties in an IOIS project.

This paper discusses the power issues that arose in a public sector Scandinavian IOIS development and implementation project.

The research problem addressed by the paper is as follows:

What were the major power issues in a Scandinavian public sector IOIS project?

The paper is organised as follows. In the next section we present a summary of the relevant literature to this study. The third section outlines the research methodology. The fourth section gives some of the complex project case background of the study. The fifth section presents the findings of our grounded theory analysis. The sixth section discusses the implications of our findings, then we conclude our study with a brief summary of our contributions.

2 THEORETICAL FRAMEWORK

It is acknowledged that power is a very ambiguous and intangible concept and therefore exact definitions of it are difficult to give. Power is a multidimensional concept, and many definitions, interpretations and theories about it abound (Bourdieu 1998, Foucault 1980, Giddens 1984, Hardy & Leiba-O’Sullivan 1998, Jasperson, Cart, Saunders, Butler, Croes & Zheng 2002, Markus & Bjørn-Andersen 1987).

Many influential social theorists - such as Foucault, Giddens and Bourdieu - have conceptualised power. According to Foucault (e.g. 1980), power must be analysed something which circulates, or rather as something which functions in the form of a chain. Foucault was interested to study power in its external visage, which means that power installs itself and produces its effects. According to him, power is inescapable. In Giddens (1984) structuration theory, power has two different perspectives: the perspective of an action of the actor and the perspective of the structural aspect. Power is, then, the ability to make changes to behaviour, and control or dominate from an institutional perspective. Bourdieu (1998), on the other hand, is interested in power from the perspective of individual strategies. Bourdieu’s practice theory discusses sources of power (economic, cultural/ knowledge, social) as a particular kind of relational resource. Agents can influence their own and other agents actions in a particular context using these resources.

These social theories have been widely utilized in IS. Foucauldian analyses of power have been quite popular during recent years (e.g. Doolin 1999, Sayer & Harvey 1997, Wynn, Whitley, Myers & DeGross 2002). These studies analyse disciplinary power in different IS contexts. Many studies using structuration theory have argued that IT conditions and shapes human action, but also that human action conditions and shapes IT (see e.g. Majchrzak, Rice, King, Malhotra & Ba 2000, Orlikowski & Robey 1991). Levina (2005), among others, has adopted Bourdieu’s practice theory and focused on what people do and how their actions shape, and are shaped by, diverse sources of power resources.

Silva (2007) suggests that none of the three epistemologies used commonly by IS researchers – phenomenology, critical theory and structuration theory – are sufficient to engage with the ‘dark side’ of power and politics in organisations, as opposed to researching only the ‘legitimate’ face of power.

Jasperson et al. (2002) have pointed out that researchers have had problems defining and measuring the theoretical construct of power in IS area. They have identified common themes in power
conceptualisations: 1) authority 2) centralisation, decision rights, participation in decision making 3) influence 4) politics and 5) power. According to Silva (2007) authority is always contested, as formal rules are open to interpretation and that is the source of politics.

A very well known and extensive categorization of power is provided by Hardy and Leiba-O’Sullivan (1998). The first dimension shows that power is wielded by using various resources to affect the outcome of decision-making processes. In the second dimension, power is wielded by supervising access to those processes. In the third dimension power is wielded through legitimation, where power is embedded in the fabric of the system. The first two dimensions lean on the assumption that power is introduced only in the face of conflict (and opposition), whereas the third dimension acknowledges that power can be used to ensure that conflict never arises. The fourth dimension (‘limits of power’) enables the investigation of aspects of power which do not normally appear in the mainstream literature of power – for instance, while some actors may receive advantages from power relations, they can not control or escape them.

To Hardy and Leiba-O’Sullivan, power is integral to empowerment. To managers and mainstream management researchers, power is legitimate and functional. Power can be thus shared. In this case, empowerment can be as a tool to motivate employees to achieve organisational goals. For critical theorists, on the other hand, power is domination, and empowerment provides the means to combat the sources of domination.

<table>
<thead>
<tr>
<th>Power of A over B</th>
<th>First dimension</th>
<th>Second dimension</th>
<th>Third dimension</th>
<th>Fourth dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interaction between A and B</td>
<td>Management of resource dependencies</td>
<td>Management of decision-making processes</td>
<td>Management of meaning</td>
<td>None, power is embedded in the system</td>
</tr>
<tr>
<td>Reason for B’s failure to influence outcomes</td>
<td>Overt conflict</td>
<td>Overt or covert conflict</td>
<td>Apparent cooperation</td>
<td>Local struggles</td>
</tr>
<tr>
<td>Empowerment of B’s requires</td>
<td>B is aware of the issue and able to get it to the decision arena, but is unable to use power effectively to influence outcomes</td>
<td>B is aware of the issue but unable to get it to the decision arena</td>
<td>B is unaware of the issue and, so, has no will resist</td>
<td>Both A and B are prisoners of the prevailing discourses of power although A may derive greater advantage from them</td>
</tr>
<tr>
<td></td>
<td>Acquisition of resources and ability to mobilize them</td>
<td>Ability to gain access to the decision arena</td>
<td>Consciousness - raising and “delegitimation” strategies to create will to resist</td>
<td>Empowerment in the sense of freedom from power effects is not possible although local struggles may produce more positive experiences</td>
</tr>
</tbody>
</table>

Table 1. Empowerment and the Dimensions of Power by Hardy and Leiba-O’Sullivan, p. 462.

As stated in the introduction, many studies (e.g. Kirsch & Beath 1996, Sarkkinen & Karsten 2005, Symon 1998, Yeh & Tsai 2001) have shown conflicts between different user groups and between IS professionals and user groups as widespread. It has been argued that user involvement has been used only as a buzzword or a weapon for achieving management goals (e.g., Hirschheim & Newman 1991, Howcroft & Wilson 2003, Kirsch & Beath 1996, Symon 1998). Gärtner & Wagner (1996) have analyzed the political frameworks of IS design and participation, and state that agenda setting related to the IS design and participation is important, as well as the legitimation of certain agendas over the others. Conflicts between different actors - workers, managers, consultants, unions and IS professionals - are evident in this process.

Markus (1983) has highlighted that the strength of resistance in an IS project is likely to be affected by the organisational position of the person to whom one loses power. According to Markus (1983) the
explanations of resistance are important because, however informal or implicit, they guide the behaviour and influence the actions taken by managers.

In distributed and multi-party IS projects, there is even a larger number of stakeholders involved, and empirical studies have revealed that there is a great potential for conflicts in this context, and that power relations between the multitude of stakeholders (e.g. IS professionals, users, graphical designers, strategists, different participating organisations) should all be acknowledged (e.g. Kumar & van Diesel 1996, Levina 2005).

It should be pointed out that, because this is a grounded theory study, although we reviewed the literature before embarking on analysis, we didn’t use labels from the theoretical framework for to code the data. Glaser (1992) directs researchers to avoid forcing the data down preconceived theoretical avenues. The idea is that the emergent theory of the study determines the relevance or otherwise of the literature review. Thus we proceeded with an ‘open mind rather than an empty head’ (Dey 1999). Once the theory has emerged, it is then the duty of the grounded theorist to engage their emergent theory with the existing literature. The next section considers our methodology.

3 METHODOLOGY

This study is an interpretative study using Glaserian grounded theory (Glaser 1978, Glaser 1998) for data analysis and theory building. Grounded theory method is very suitable for research areas where there is little existing theory. In this case, grounded theory method was very useful because there is little existing theory in inter-organisational IS (IOIS) projects and especially their implementation area.

This research studied 8 organisational project teams and 2 inter-organisational project teams, in a large, three years long IOIS development and implementation project. The IOIS project studied, ViWo, was a Scandinavian public sector organisation collaboration. This research tracked the whole IS project and it had a unique approach – no framing questions were used, the focus was entirely on the experience of the project member.

Data collected in the project ranged from in depth interviews (250 pages of transcripts), to observations of project meetings (20), diaries (80 pages of notes), 48 memorandums of project and steering group meetings, and e-mails (over 700) containing what project members sent to each other during these years and other secondary data (the data of previous project) were also analysed.

Over the three year timeframe of the project, 36 different people were involved. Some people were involved only once or twice in project meetings. There were 20 active project members in the project, 14 of whom were willing to be interviewed. Among the interviewees were managers from the steering group, representatives of suppliers, members of the research organisation (Rhoo) associated with the project, and users active in the project. The interviews lasted from 45 minutes to two and a half hours. The open nature of the interviews enabled the interviewees to explain their deep feelings about the project that would not have surfaced otherwise. The interviewers told their own story about the project and its progress. It is said that through narrative stories we are able to get close to people’s experiences (Clandinin & Connely 1994).

In this study ‘Glaserian’ grounded theory technique was used as the method of analysis. Since 1990, grounded theory has evolved into two distinct versions (Urquhart 2001, 2007, Urquhart & Fernández 2006). This occurred on the publication of Strauss and Corbin’s (1990) book which is a distinct departure from the classic “discovering of theory from data” in the seminal book of Glaser and Strauss (1967) which introduced grounded theory. The 1990 book helped popularise grounded theory and is widely used; however, it has also been described as rather formulaic and overburdened with rules Kendall (1999). From our perspective then, the Glaserian version has the twin advantages of being closer to the original, classic version of grounded theory, and of being much more flexible.
Glaser recommends that the researcher takes a very open approach in order to ensure that concepts genuinely arise from the data as opposed to preconceived questions, categories and hypotheses (1992). We allowed the data to suggest categories to us, rather than using preconceived categories. It is also recommended by grounded theorist that researchers collect the data over many phases of research - when the same concepts occur over and over again, the saturation of concepts can be considered reliable. This was our experience that the concepts recurred over different phases. Glaser also (1992, 1998) recommends the collection of rich, versatile data in the form of different interviews, observations, and diaries, and our data collection followed this directive.

We followed the Glaserian (and classic grounded theory) coding stages – open coding, selective coding and theoretical coding. According to Glaser (1978), the open coding is the most important building block of GTM. At the open coding stage, the interview data, field notes and e-mails were analysed line by line, and the project memorandums were analysed paragraph by paragraph. Urquhart (2001, 2007) has pointed out that line by line coding is recommended by both Strauss and Glaser and is demonstrably fruitful. However, as the project memorandums were secondary data, it was appropriate to code at a paragraph or page level (Urquhart 2007). The coding process was done by one researcher (first author of this article) and then discussed with second author of this paper. After discussions, some code names were changed and open codes reallocated to different categories.

During selective coding and through an iterative process, we discovered our emergent categories. We then considered the relationships between categories during theoretical coding, and analytic memos (Glaser 1992) assisted with this process. Glaser (1978) emphasis that the bedrock of theory generation is the writing of theoretical memos. One of our emergent categories was power, and it is this concept that we concentrate on in this paper. In this study we can indicate what are the main strengths of Glaserian grounded theory method. A detailed and systematic analysis of data allows the discovery of new concepts.

4 PROJECT BACKGROUND

Here we give some of the complex background of the IOIS project, to help with interpretation of the findings.

4.1 History of the project

ViWo was preceded by a pilot project called PreViWo. PreViWo was implemented in three steps (specification, interface pilot and planning) in the years 2002-2003. The aim of the PreViWo project was defined as “to specify and implement a pilot IS to support a process, its actors and task performed by them” (Project card, March 8, 2002). Table 2 contains the actors in the pilot project. The pilot project was influential in framing the organisation of the larger project we studied (ViWo), and it could also be seen that the history of the pilot project influenced the perceptions of the participants. Alpha was the leading organisation for the pilot project as the organisation who applied and received funding for the project.

<table>
<thead>
<tr>
<th>Organisation</th>
<th>Role of Organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ministry</td>
<td>Ministry responsible for funding the pilot project</td>
</tr>
<tr>
<td>Nofco</td>
<td>Consortium of user organisations in charge of the project (a virtual organisation)</td>
</tr>
<tr>
<td>Opti</td>
<td>Consortium of user organisations (an organ of cooperation) that used a similar IOIS</td>
</tr>
<tr>
<td>Nuovo, Eino</td>
<td>Suppliers of the software</td>
</tr>
<tr>
<td>Cumma</td>
<td>Expert consultants</td>
</tr>
<tr>
<td>Alpha</td>
<td>User organisation that was a member of Nofco and Opti and initiated the project</td>
</tr>
</tbody>
</table>

*Table 2. Organisations involved in PreViWo*
4.2 Main players – ViWo project

The goal of the IS project was that an Inter Organisational IS (IOIS), named ViWo, would be designed and taken into use by several organisations of the same type. The project aimed to carry out a pilot test of the IS in these organisations before establishing the system at the national level. The development of ViWo involved electronification of a work process to facilitate office work, consolidate information across organisations, and manage key activities.

In the ViWo project, Nofco was no longer in charge of the project - a project management organisation, Rhoo, was brought in. They also managed some research objectives around the project. The key user organisations now consisted of Alpha, the original lead user organisation, plus user organisations Beta, Gamma and Delta who came from Nofco and Opti. Nofco now consisted of 21 organisations, and it would be these organisations that would eventually use ViWo. The organisations collaborated with the relevant Ministry, suppliers and consultants.

<table>
<thead>
<tr>
<th>Organisations</th>
<th>Role of Organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ministry</td>
<td>Ministry responsible for funding the IOIS project</td>
</tr>
</tbody>
</table>
| Nofco             | Consortium of 21 user organisations (Virtual organisation)  
                   • The basic function of Nofco was to promote and develop locally, regionally, and nationally the utilisation of IT and to enhance inter-organisational collaboration in multiple research-related issues and administrative practices |
| Alpha, Beta, Gamma, Delta | Lead user organisations in the project  
                   • Alpha was also the fund holder for the project |
| Rhoo              | Organisation responsible for project management and research objectives                                                                           |
| Socca             | Software company that supplies the software solutions for the project                                                                               |
| Cumma             |  
                   • Part of the national research network that develop research and IT based services for the needs of research and education, and the supporting IT administration  
                   • Acted as an expert advisor. Withdrew from the project before it ended                                                                          |

Table 3. Organisations involved in ViWo

Cumma, eventually, withdrew from the project: “We withdrew from so many occasions that we realized that we could not continue in this way. This was probably because we received a role that was more demanding than the one we pursued in the initial discussions and negotiations…”(Jack, Supplier Cumma)

4.3 Organisational project members in ViWo

The table below names members of each organisation and their roles in the project. As can be seen below, there were a large number of people involved, and some had experience of the previous project.

<table>
<thead>
<tr>
<th>Organisations</th>
<th>Members and their roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ministry, Financier</td>
<td>Marie – Govt Minister, steering group member</td>
</tr>
</tbody>
</table>
| Nofco, Consortium of user organisations | Sarah; also previous member of PreViWo  
                   • Sheila; steering group member, previous project manager of PreViwo  
                   • Gabriel attended project group meetings occasionally  
                   • Hale; Paul; Steering group members |
| Alpha project team | Lucy; Project leader. Also previous member of PreViwo, Member of steering group                                                                  |
Lisa; User representative (of 11 organisational units), also previous member of PreViwo
Arthur; Expert, Opti Consortium, previous member of PreViwo
Esther, Lauren and Thod; Opti Consortium people, attended project group meetings occasionally
Sam, user representative, attended project group meetings occasionally

Kathy; Opti Consortium person, attended project group meetings occasionally. Steering group member, also previous member of PreViwo
Heather, Tom; User representatives
Katie; User representative, present in some steering group meetings

Ellen, User representative, present in some steering group meetings
Martha; User representative
Pamela; Steering group member, previous member of PreViwo
Alice; Steering group member, previous member of PreViwo

Tim; Expert, Steering group member
Sophie, Ann; User Representatives
Susan; Steering group member

Matthew; Project leader (also previous member of PreViwo and Opti Consortium)
Ruut; Project Manager, Steering group member
Rachel; assistant project manager, Member of Quality Group
Thomas, Simon; Members of Quality Group

Walter
Tom. Attended project group meetings occasionally

John, previous member of PreViwo
Peter, Jack, attended project group meetings occasionally. Previous member of PreViwo
Daniel, attended project group meetings occasionally
Ellie, member of project group and also present in some steering group meetings
Mark

Table 4. Project group organisations and their members related to ViWo

A difficult question was who would be the ViWo project manager. ViWo was perceived to be a demanding project, and an experienced manager would be needed. Matthew, the Project Leader of Rhoo, suggested to Lucy and her colleagues from Beta and Gamma that Rhoo could take the responsibility of leading the project, Ruut being the project manager. This suggestion was approved, and so the project manager changed: in PreViWo it was Sheila from Nofco but in ViWo it was Ruut from Rhoo. Ruut had extensive experience in practical software development. It was envisaged that, in addition to Matthew and Ruut, Rhoo could provide a three person quality assurance group for ViWo development. When it came to the choice of software vendors, Matthew’s argument was that Socca would deliver a useful system even in the situation of unclear client requirements.

5 THE FINDINGS

Power was one core category which emerged through the grounded theory analysis, and this section discusses the category in detail. We identified Sources of Power, Power as Resistance, Reasons for the

Proceedings ECIS 2009
**Power Struggle** and **Power as Exercised** as important selective codes of the category. Table 5 presents the open codes and selective codes that make up the category.

<table>
<thead>
<tr>
<th>Category</th>
<th>Selective Codes</th>
<th>Open Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sources of power</td>
<td>Legitimate power, Expert power, Politic power</td>
<td></td>
</tr>
<tr>
<td>Power as resistance</td>
<td>Control of decision making, Tensions between Old and New, Seeming Acceptance, Insecurity</td>
<td></td>
</tr>
<tr>
<td>Reasons for the power struggle</td>
<td>Previous project, Positions in project, Jargon, Time pressure, Unclear responsibilities</td>
<td></td>
</tr>
<tr>
<td>Power as exercised</td>
<td>Final authority, Veto power, ‘Forged power’</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Construction of Power Category

5.1 **Sources of power**

This selective code illustrates the different sources of power in the project. For instance, Ruut (Project Manager, Rhoo) prohibited some project members attending the project meetings by using her legitimate power. She was pulling strings by not inviting all former people (Nofco, Consortium of User Organisations) to the project meetings. Some members guessed that she did it that was able to avoid competition between her and the previous project manager of PreViWo. She also sent an email to Simon (Rhoo) that his presence in project meetings is not necessary. Simon was amazed and asked if some other project management presence was necessary, if his presence was not. It was speculated that for some reason they did not get on well with each other.

Expert power was also in evidence and conflicted at times with the project managers legitimate power. For instance, there was a ‘tug-of-war’ between the suppliers and the project manager around various issues. The discussions were “little bit hostile” (Thomas, Rhoo). Supplier Cumma felt that disagreements were frequent and faults were dealt with by “tattling” to the project manager. So Supplier Cumma sought background support for their work from other project members on the basis of their expert power. At that time, Cumma had a good reputation and there was discussion among the project management people that it was not easy to disagree with Cumma because of the skill and know-how owned by the company. Later, however, the confidence in Cumma started to wane.

Thomas (Rhoo) pondered how the steering group should regard the matter, since nothing was happening. Thomas thought that the roles of “generals” and “officers” were not defined and consequently attempts were made “to transfer war leadership onto wrong shoulders” Thomas thought that because legitimate power were not defined in project it caused that people ‘took’ power and there were not abilities to manage it. This caused ‘anarchistic’ behaviour in his opinion.

In addition to legitimate and expertise power there was political power used in the project, as shown in how the project was represented as a success to those outside the project. At the end stage of the project, Nofco announced that a journal article had been published on the ViWo project. The announcement incorporated a message requesting receivers to notify their international partners of the publication of the article. At that stage, this raised criticism among the project members, because they thought that Nofco had wrongly collected merit from work that it had neither planned nor implemented alone. The issue came up among the employees of the other supplier as well as among the project management. Thus, the representative of the other supplier, Walter, posed the question: “…what was it that Cumma had planned and Nofco implemented? And noted that Socca’s name had not been mentioned at all in that connection … “ (Walter, Supplier Socca, Email June 30, 2005).
5.2 Power as resistance

In the experiences of project members, the notion of ‘power as resistance’ also emerged. There were situations where project members combated or at least wanted to combat domination by other project members.

Some members wanted for example to take control of decision-making in the project. “Who decides and on what? It would be good to know so that the matters do not need to hashed over unnecessarily at meetings...” (Walter Supplier Socca). Control of decision-making was a central problem in the project which caused resistance. The decision-making process was seen as “yeh-naw discussion” (Thomas, Rhoo) and as a “competition” (Ruut, Project Manager, Rhoo). There were tensions between the new and old project members. Both project management and supplier Cumma felt that it was often necessary to return to decisions due to questions or critique presented by Nofco (Consortium of User Organisations). Ruut (Project Manager, Rhoo) complained that it was difficult to see whose rules should be followed, although decisions were made at project meetings. Both Suppliers and Project Management felt that the representatives of Nofco (Consortium of User Organisations) inhibited decision-making. Jack’s (Supplier Cumma) quotation reveals that problem was that “too often problems that emerged from practical work or were brought up on discussions were ignored by pointing out that the process had already been defined...” Sheila (Nofco,), for her part, saw that really big problem was that suppliers were given the power to decide on matters in the project group.

*Tensions between old and new* were evident because of the previous project. Ruut, Project manager felt that she was an “outsider” when Sarah and Sheila (Members of Nofco) felt that maintaining an artificial separation between these two IS projects caused problems for organisational memory. “We assumed then that since Cumma was chosen as the second supplier, it would ensure the contiuance...but the old information had not been passed on, that gatekeeper’s task did not continue...” (Sheila) Sheila was, for example, surprised that Socca had begun to design a user interface even though one was already available that had been done in PreViWo.

There were also situations, where some people were aware of controversial issues, but were unable to use power effectively to influence outcomes or unable to get it to the decision making arena. In many cases project members just agreed to accept proposals (*seeming* acceptance), despite disagreeing with the decision.

Unclear plans caused *insecurity* among the project members. It was difficult to plan project schedules and estimate future workloads. So the members of Nofco demanded that some kind of long term plans should be made. "In other words, matters have some up kind of unexpectedly, or is that typical in IT projects and IS projects that it is so? I have pondered even from the standpoint of my own work that is it so... (related to) project planning and project management and these types of things..." (Sheila)

5.3 Reasons for the power struggle

When interviewing project members about their experiences, the topic ‘reasons for the power struggle’ also emerged. There were various open codes which made up the selective code of reasons for the power struggle; *previous project, positions in project, jargon, time pressure, and unclear responsibilities.*

The *previous project* affected the power struggle in many ways, for example leading questions could be asked by those who had knowledge of the previous project to those who did not. Ruut the project manager felt for example that it was difficult to see whose rules should be followed: “Naturally the previous project has caused pressures especially because the former people are there. I have sometimes sensed an air of competition concerning who is in charge and whose rules are followed...” (Ruut).
The previous project (PreViWo) managers were not invited to take charge of the ViWo project and the suppliers were replaced. Matthew (Rhoo) pointed out that discontent with PreViWo (schedule and specifications problems) had resulted in an effort to change actors in the new project.

It also became evident that the representative of the supplier (Cumma) who had been involved as an expert in (PreViWo) did not support the use of the material in the further project: “John described the specifications in his colourful style as suitable to be thrown into a waste basket…” (Matthew)

The positions in the project were very different to the PreViWo project. The quality of the specifications and the poor success of PreViWo were the reasons for the adoption of a different pattern of organisation in the new project. The project manager from Nofco was changed in summer 2003 because of project management issues. The suppliers were replaced in early 2004 because of the poor quality of the specifications.

Interestingly, Nofco’s members felt that too much power was given to suppliers to decide on matters. Suppliers, on the other hand, said that Nofco’s representatives and project management ignored many problems by saying that the process had already been defined.

The use of jargon was a problem in the project. John (Supplier Cumma) thought that users should take more part in decision-making, but the problem was, according to users, that it was difficult to understand technical matters. Lisa (User Representative, Alpha) felt that it was difficult to form opinions because she didn’t understand what was discussed. “If someone mentions the word interface once more, I’ll jump out the window…! Let’s speak about matter without technology…” (Lisa). Eventually, the users demanded that the project manager and supplier use language which they could understand.

Time pressure was also something which contributed to the power struggle. Project members had very different perceptions about the time pressure in the project. One project member pondered on how it was possible that people felt that there was no time to have a lunch or even to go to the bathroom. But she encountered a different attitude from other project members “Amazing that we wait at a meeting for an hour while someone (user representative) feeds her dog”.

Unclear responsibilities become also apparent in an e-mail message sent to the researcher by Cumma’s representative: “Interesting definition of policy, that because it is related to a [technical matter], it belongs to Cumma! In my opinion the application form belongs to Socca, but Cumma has to take part in ensuring the implementation of the form by specifying necessary interfaces…” (Peter, e-mail Aug. 31, 2004).

5.4 Power as exercised

How power was actually exercised in the project was very interesting. We found three open codes for the selective code power as exercised.

Final authority: The other project leader, Lucy (Alpha) said that she make a lot of decisions trusting to others views, using intuition and feelings, because she thought that she was layman in these things. She said that, for example, that when project manager pointed out something in a plausible way, she gave the necessary final authority. Final authority did not always rest with the same person. A good example was the situation where the project manager and Sheila (Nofco’s representative) battled about the appearance of the display, and where project manager finally climbed down. Some decisions were achieved asking project manager to use her final authority.

Veto power: Nofco had the ability to veto decisions, albeit informally. It was often necessary to revisit to decisions due to questions or critique from user organisations. Ruut (Project Manager, Rhoo) complained that it was difficult to see whose rules should be followed, although decisions were made at project meetings. Both Suppliers and Project Management felt that the representatives of Nofco (Consortium of User Organisations) inhibited decision-making.
“Forged” power. This was a positive experience as a result of organising and of social relationships. Things were done at short notice at the request of the project manager because she wielded this type of power. The members of Nofco felt that this was how the project manager got people to do things she wanted.

6 DISCUSSION

We have shown in the previous section the complex power issues that arose in a Scandinavian IOIS project. We illustrated different selective codes of power that occurred – sources of power, power as resistance, reasons for the power struggle, and power as exercised. While sources of power were easy to identify, it was also easy to see how resistance occurred. Reasons for the power struggle seemed to revolve around both the history of the project and unclear responsibilities. Power as exercised in the project came down to who could actually have the final authority in the project organisation structure, but this final authority was often contested and switched between people. Resistance was indeed more likely if the individual was a peer of the individual trying to impose the decision (Markus 1983).

The governance of the project was particularly challenging because of the number of organisations and structures involved. This is a potential problem for all IOIS projects, as of course governance does need to be defined between organisations in such projects. In this particular instance that the project structure, as set out, was a very complex one, with undefined governance and responsibilities. In particular, the authority of Nofco, the user consortium and its relationship to the lead organisations was poorly defined. What makes it even more complicated is that the background of the project was very ambiguous and unclear to many participants. The complex project structure led to all sorts of unforeseen problems. Silva (2007) argues that authority is always contested, as formal rules are open to interpretations and that is the source of politics.

In this case, all participants (users, suppliers, project management personnel and representatives of Nofco) had power in certain situations, but were also resisting the power of somebody else in other situations. In many cases it seemed that this situation was reciprocal, each actor in turn having power and resisting domination. For example, the users used their power to change topics of discussion, suppliers required users to make decisions, and Nofco criticised the project manager for giving suppliers too much power in decisions. There were also some situations, where some people were aware of controversial issues, but were unable to use power effectively to influence outcomes or unable to get it to decision making arena. The users were not necessarily the resource weak group in the project. Users could “talk back”, and in some instances even prevent the decision making and the progress of the project. Both project management and suppliers were affected by this, but could not use their power effectively to change the situation. In all, the ‘limits of power’ from the critical viewpoint was palpable in the case.

Hardy and Leiba-O’Sullivan (1998) enable us to see that aspect of power, which does not normally appear in the mainstream literature of power - ‘limits of power’. Limits of power were clearly seen in the case. Jasperson et al. (2002) and Markus and Bjørn-Andersen (1987) also emphasise that power which inheres in an official position. Our study reveals that in some situations it is easy to use power over others in some position, but on the other hand, it is also shown that it is possible to counter that power.

We can note that in IOIS projects, when there are any number of organisations and structures involved, it is more likely that there will be ‘competition’ and ‘tug-of-war’ situations, and that legitimate power is used as ‘justice’.

Our study also illustrated how unclear roles influenced the exercise of political power. According to Silva (2007), the study of power poses challenges because of the twofold nature of power: 1) power that arises from positions of authority or 2) its informal dimension, i.e. politics. Most of the challenges in this study came from the informal dimension, which then came up against a veto in the formal authority structures.

Proceedings ECIS 2009
7 CONCLUSION

Our research raises many important issues related to research on power in the IS field. We agree with Silva (2007) that we need research on power that emphasises the interpretations of meanings, intentions and actions which are suitable for making sense of such a complex phenomenon. In contrast to traditional problems and power struggle between developers and users or managers and users, our research shows that in a multiparty IS project it is extremely difficult to say who ‘has’ power and who is in need of ‘empowerment’. We have also made a methodological contribution, we feel, by our detailed consideration of day to day issues of power, using grounded theory analysis.

In contrast to traditional problems and power struggle between IS professionals and users or managers and users, this research shows that in an inter-organisational IS project it is difficult to say who ‘has’ power and who ‘lacks’ power. In all, power is clearly a complicated matter and there is no clear cut way of defining “whose power over whom” is to be analysed, in IS setting or elsewhere.

We would also contend that, with the advent of globalisation, there are an increasing amount of IOIS projects in existence, and that there is a need to research power issues in such projects. The potential for conflicts in such projects are greater than in organisational projects, because of the need to set up agreements and governance structures between parties involved in such projects. We urge IS researchers to explore how particular governance structures might either constrain or enable conflicts in such projects.

References


## GENERATIVE MECHANISMS FOR INNOVATION IN INFORMATION INFRASTRUCTURES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0257.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Case Study, Critical Realism, Innovation Theory, Inter-organisational change</td>
</tr>
</tbody>
</table>
GENERATIVE MECHANISMS FOR INNOVATION IN INFORMATION INFRASTRUCTURES

Bendik Bygstad
Norwegian School of IT, Schweigaards gt.14, 0185 Oslo, Norway
bendik.bygstad@nith.no

Abstract

This paper investigates how innovation of ICT based services takes place within existing infrastructures, including the whole network of technology, vendors and customers. Our research question is, how can an information infrastructure provide generative mechanisms for innovation of ICT based services?

Building on a critical realist approach, our empirical evidence was a case study within an international airline, aiming to diversify its services. From our analysis we propose that there are two self-reinforcement mechanisms in information infrastructures. First, we identified the innovation reinforcement mechanism, resulting in new services. Second, there is the service reinforcement mechanism, resulting in more users and profits.

The practical implication of our framework is to show that although ICT-based innovation cannot be planned and managed in detail, the innovation mechanism may help organisations to facilitate the innovation process in a structured way.

Keywords: Innovation, information infrastructure, generative mechanism, case study
1 INTRODUCTION

During the past decade innovation of ICT based services has transformed several industries, such as financial services, telecom and IT, and media. Other sectors are following, for example the music industry and e-government. Innovation is not easy; it is hard to plan and manage (Tidd and Hull 2003), it is socio-technical and non-linear (Janszen 2000) and should be done mindfully (Swanson and Ramiller 2004). As Christensen has shown, innovation is also paradoxical; the innovative company faces the danger of becoming a victim of its own success, as it develops a mindset that hinders new innovations (Christensen 1997).

This paper explores innovation from the perspective of information infrastructures. As defined by Hanseth an information infrastructure is “a shared, evolving, open, standardized, and heterogeneous installed base” (Hanseth 2002) (p. 2). The key term in this concept is the installed base, which denotes the number of components and users in the information infrastructure. For example, the installed base of iPod consists of the iPod players, the central music database and the millions of users. The key attribute of a successful information infrastructure is the self-reinforcing mechanism, illustrated with Gindley’s figure below. An installed base attracts complimentary products. This makes the information infrastructure more attractive to users, and generates more use, which in turn increases the size of the installed base.

![Figure 1. Grindley’s standards reinforcement mechanism. (From Hanseth, 2002).](image-url)

Compared to the traditional concept of information systems, the notion of information infrastructure offers two important advantages. First, it changes the object of study from a single application within a company to the world of large socio-technical networks. Second, it offers a new perspective on how such solutions are developed. While the stand-alone system can be designed and implemented, the information infrastructure is seldom designed from scratch; rather it is growing more organically from an existing base.

In the context of innovation this creates a certain paradox with information infrastructure; on one hand they are usually not made from scratch, on the other hand a number of information infrastructures are spectacular innovations. Well known examples are Internet based supply chain networks, FaceBook and eBay. This calls for more research on how successful infrastructures are developed in more detail.

This paper investigates how innovation of ICT based services takes place within existing infrastructures, including the whole network of technology, vendors and customers. One point of departure is that information infrastructures usually represent large investments in IT architecture. Thus, one ambition is to understand the self-reinforcement mechanisms in more detail, in particular the role of the IT architecture.

The research question is:
• How can an information infrastructure provide generative mechanisms for innovation?

This paper proceeds by a review of some central contributions on information infrastructure and innovation. Then, in section 3, the research approach is briefly presented, while section 4 presents the case study. Findings are discussed in section 5 and conclusions offered in section 6.

2 A BRIEF REVIEW ON INNOVATION IN INFRASTRUCTURES

Star and Ruhleder (1996) asserted that “infrastructure is a fundamentally relational concept. It becomes infrastructure in relation to organized practices” (p. 4). They defined information infrastructure in the following terms: It is embedded into other structures, transparent in use, has reach and scope beyond a single event, is learned as part of a membership, it links with conventions of practice, embodies standards to be able to plug into other structures, is built on an installed base and, finally, it becomes visible upon breakdowns (Star and Ruhleder 1996).

Thus, innovation in information infrastructures presents a double set of challenges.

First, ICT based service innovation is in itself a complex process. The service innovation process differs from the innovation of products: Services are usually developed in close interaction with the customers, and they are more often innovated in networks rather than labs. (Tidd and Hull 2003; Abramovic and Bancel-Charenso 2004). ICT based service innovation often redefines the roles of the service provider and the users. An illustrative and very successful example is the Internet bank. The real innovation of Internet banking is not the web software, but the redefinition of roles: The bank organization provides the technological infrastructure, the technology is available 24/7, and the customers are doing the transactions themselves. The actual innovation is the interplay between the providing organization, the new technology and the users.

Second, research on information infrastructures has shown that the development trajectories of information infrastructures are hard to predict and control (Ciborra 2000; Hanseth 2007). Extensive case research showed that many of the key approaches from strategic IT management, such as top-down planning, management control and strategic alignment, do not work as intended in the context of corporate information infrastructures. Rather, the authors assert that top-down approaches rather will increase the managerial problems of large information infrastructures, and instead they point to such concepts as cultivation and care in order to foster innovation and growth. Taking these ideas further Hanseth and Lytinen pointed out that large-scale information infrastructures have a complexity that goes beyond traditional systems design. They proposed a full theory on the design of information infrastructures, focusing on how to foster the growth of an installed base, building on networks economics and complexity theory (Hanseth and Lytinen 2008).

The five principles are (p.11):

• Design initially for usefulness
• Draw upon existing installed bases
• Expand installed base by persuasive tactics
• Make it simple
• Modularize by building separately key functions of each infrastructure, use layering, and gateways

These key design principles exploit the dynamics of self-reinforcing mechanisms of growing information infrastructures. In addition, we need to consider another feature of information infrastructure innovation, namely innovation as a collective process.

Fifteen years ago van de Ven described innovation as a collective process, including not only the entrepreneur, but also includes a variety of industrial and public actors (Van de Ven 1993). The insight that innovations today more seldom take place within a single organisation has triggered a strong interest in different forms of co-operative innovation processes (Bessant and Tidd 2007). Andersson et
al. found that architectural knowledge is crucial in inter-organisational innovation, in four dimensions: technology capability awareness, use context sensitivity, business model understanding and boundary-spanning competence (Andersson et al. 2008). Other researchers have investigated the innovative capabilities of distributed and heterogeneous networks, and showed that innovations in this context may be regarded as a series of cognitive and social translations (Yoo et al. 2008). Cognitive translations include the creation of ideas into actionable artifacts, while social translations take place at the borders of different knowledge communities, where involved actors negotiate and mutually adapt a solution. While the cognitive translation process is relatively linear, the social translation process is much less predictable.

Summing-up this very brief overview: We know that the innovation process in information infrastructures is complex and non-linear, that it is essential that it exploits the dynamics of self-reinforcement of the installed base and that it is a collective endeavour which is both cognitive and social. What we know less about, is the causal structure of innovation in information infrastructures, and how this is linked to the growth of the installed base. Thus, what we are looking for is a recursive structure of mechanisms, which links innovation and growth in more detail.

3 METHOD

3.1 Research approach

The general approach for this study was a critical realist case study. The basic assumption of critical realism is the existence of a real world independent of our knowledge of it. Reality is conceived as being stratified in three domains. The real domain consists of objects, both physical and social, with capacities for behaviour called mechanisms. These mechanisms may (or may not) trigger events in the domain of the actual. In the third layer these events may be (or not) observed, in the empirical domain. Thus, structures are not deterministic; they enable and constrain events (Archer 1995; Sayer 2000).

Critical realism combines a realist ontology with an interpretive epistemology (Sayer 2000); although a real world exists, our knowledge of it is socially constructed and fallible. This does not imply an epistemological relativism; since a real world does exist critical realism holds that some theories approximate reality better than others. This process of approximization is seen as a key part of scientific enquiry. It follows from this that critical realism does not aim to uncover general laws, but to understand and explain the underlying structure and mechanisms. This is done through retroduction; we take an empirical observation and hypothesize a mechanism that might explain that particular outcome. These mechanisms are associated with the nature of the object of study, not to the regularities of events (Sayer 2000).

Our object of study is an information infrastructure. Following DeLanda we envisage an information infrastructure as an assemblage, i.e.”wholes characterized by relations of exteriorities” (p.10). The basic attributes of assemblages are (DeLanda 2006):

- An assemblage consists of various types of components, which in themselves may be assemblages. An assemblage allows for interactions between components that are emergent, i.e. mechanisms whose behaviour cannot be explained by the properties of the component.
- A component is self-subsistent and may be unplugged from one assemblage and plugged into another without losing its identity.
- A given component may play a mixture of material and expressive roles by exercising different sets of capabilities. Assemblages may increase their homogeneity by sharpening its boundaries (territorialization) or destabilizing it (determinitorialization).

In an information infrastructure context we should conceive these structures as semi-stable. They are the result of relatively stable patterns of behaviour over time, but at the same time they are changing...
continuously, as the result of growth and change. For example, the behaviour of a user community will have some stable patterns (in spite of people joining or quitting), but it may also change its patterns of behaviour in interaction with a new service, at the higher level of assemblage.

3.2 Data collection

The case company, Norwegian, was chosen for two reasons. First, it was a young and successful company, with a reputation for innovation. Second, the company was expanding its initial successful infrastructure of booking services into new ICT-based services, thus constituting a fruitful case to study innovation in infrastructures.

Data collection at Norwegian was conducted during a period of six months in 2008. Ten managers and specialists were interviewed, each circa 2 hours, some of them twice. In addition a large volume of technical documentation (business plans, project plans, contracts, technical architecture documents) was analysed. To ensure internal validity the preliminary findings were discussed with informants, and paper drafts were sent key informants for comments.

3.3 Data analysis

The inspiration for the data analysis was the following passage from de Landa (2006):

“To give a complete explanation of a social process at a given scale, we need to elucidate not only micro-macro mechanisms, those behind the emergence of the whole, but also the macro-micro mechanisms through which the whole provides its component parts with constraints and resources, placing limitations on what they can do while enabling novel performances” (p.34).

The practical search for these mechanisms was conducted in the following three steps (Pettigrew 1985). First, a time line was established, and important events were identified. Then a comprehensive analysis of organizational design and culture, technical development and business strategies was done, focusing particularly on the interplay between these dimensions. Third, to ensure internal validity the preliminary findings were discussed with informants, and paper drafts were sent to key informants for comments.

4 THE CASE STUDY

Norwegian is an international airline carrier based in Norway. Its strong growth started in 2002, when it established a national network, helped by the government deregulation of the airline industry. Today the company has 1.300 employees, 85 destinations in Europe and carried 9.1 mill passengers in 2008. More than 85 % of ticket sales are accomplished on the web (Norwegian.no). The company has pioneered the Scandinavian low price airline market, and has been quite innovative. Some important events were:

2002 : Introducing low cost airline in Norway, with print-out tickets with bar-code identification
2004 : Introducing the low-price calendar (this was internationally patented)
2005 : Dialogue with 85% of customers is electronic
2007 : Bank Norwegian is introduced
2008 : Call Norwegian (mobile telephone operator) is introduced

In 2007 the company decided to enter the banking market with Bank Norwegian. Said the CEO Bjørn Kjos at the start: “Today we have one of the most visited web pages in Norway, with 2-3 million visitors each month. We aim at coupling this traffic towards bank services.” (E24, 4th May 2007). The initiative has been quite successful; in fact so successful that Norwegian will offer a mobile service Call Norwegian, based on the same thinking.
If we look closer at the company parts of the success may be explained by a particular IT architecture. It is illustrated and simplified in figure 2. The key elements are the web page for each service, the databases and “the bus.” Each service constitutes an information infrastructure, with a number of registered customers. For the airline this is currently ca. 1 million, for the bank around 80,000, while the mobile company was starting up in the autumn 2008. In this paper the focus is on the development of Call Norwegian, as an extension to the established infrastructure.

![Figure 2: Infrastructure at Norwegian](image)

The architecture allows the company to innovate on an existing infrastructure, in much the same way as Virgin and Amazon (Cai et al. 2008) have done. The traffic on the airline website may be routed to other services at very low marketing costs. Accordingly, extended infrastructure services, such as bank system and mobile system (from external providers) may be linked to the “bus” at low costs and in time windows of opportunity. It is essential that all communication with the customers is electronic, as a combination of web pages and e-mail. Of course, this lowers transaction costs, but more importantly, it makes it much easier to enroll new customers into the infrastructure.

### 4.1 Call Norwegian

The idea of a mobile company came in 2005, after the initial success of Norwegian. Where do ideas come from?

Said the Director of business development:

“We had established a very flexible IT architecture, and we realized at the time that it would be possible to innovate new services on this. First we were just brainstorming rather freely; how could a combination of brand and technology generate new business? A GSM operator seemed to have some similarities with airline booking. Later, after the establishment of Bank Norwegian and the reward system in 2007, the idea had matured. We now focused much more on achieving synergies with the airline by providing an integrated set of mobile services for the traveler. We believe that the mobile phone will become more important as a terminal than the PC”
After a board decision in January 2008 a project was established. The main ambition of the project was to establish a number of mobile services:

- A mobile portal, with booking, payment and check-in services, plus content from many providers
- Broadband mobile access at airports and during flights
- GSM mobile services

In March a detailed Request for Information was sent to various telecom vendors, and a series of meetings were held in April. The business model was completed in May, and the contracts with the key vendors signed in August. The project was then organised with 5 sub-projects, illustrated in table 2. Norwegian organised the project as close as possible to the future operations. The project was run by Norwegian Portal Director, acting as CIO in Call, using hired expert consultants in the planning and development phase.

<table>
<thead>
<tr>
<th>Sub project</th>
<th>Technology</th>
<th>Responsible in operational phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile portal</td>
<td>Norwegian solution developed in .NET</td>
<td>Norwegian and partners’ services</td>
</tr>
<tr>
<td>Airport services</td>
<td>Norwegian solution developed in .NET based on Radius Server and infrastructure provided by Call Norwegian and Avinor</td>
<td>Avinor</td>
</tr>
<tr>
<td>Billing</td>
<td>NaviBilling by TeleBilling</td>
<td>TeleBilling</td>
</tr>
<tr>
<td>Network on board</td>
<td>Not yet decided</td>
<td>Norwegian and partner</td>
</tr>
<tr>
<td>Budget and finance</td>
<td></td>
<td>Norwegian</td>
</tr>
</tbody>
</table>

Table 2: The Call Norwegian project

The Director of Business Development commented:

“When we started the project the solutions were only sketched out as architectural ideas and financial opportunities. We learned a lot from the meetings with potential vendors and content providers in the spring 2008, exploring a space of possibilities.”

Said the Portal Director:

“We focused on how to make money on new services, analysing which services we should provide ourselves, which we should buy and how they should be integrated. At the same time we are very concerned about our architecture; it as an important ambition to maintain it as ‘clean’ as possible. We don’t really go for cutting-edge solutions. Rather, we combine known and stable components in new ways.”

The technical solution is illustrated in figure 3. The central element is the Call Norwegian Business Bus and interfaces, and all communication goes via the bus, as web services. This also includes communication with the Airline bus (CRM base) and the Norwegian Bank Bus (reward services). Most new services were bought, but a few was developed in-house, such as the purchasing and handling of wireless access at the airports.

The bus serves as a backbone for the various transaction systems such as billing, finance, GSM services, and Norwegian’s shared CRM system. It also connects these systems to the sales and service channels, in particular mobile, web and email communications. Technically, the bus is a piece of .net software. It enables the company to expand the number of sales nodes (above) and the service nodes (below). In technical terms we might describe the function of the bus as bridging two different standards; the standards of World Wide Web with the standards of telecom and billing systems. This service oriented architecture allows the company to establish new nodes, and get rid of unneeded ones, in a very short time span, because the services are bought or leased, not developed.
Figure 3: Call Norwegian infrastructure (Courtesy of H.-P. Aanby)

Seen as information infrastructure we may describe it as a heterogeneous network, consisting of a large number of actors, connected to the business bus.

- Users will access the services via the mobile portal. One minor (but exciting) innovation is using the mobile phone to check-in and boarding, by providing a link to a ticket barcode, which may be read electronically right from the mobile screen.
- The providers of services include TeleBilling, who provided billing, Avinor who provided airport broadband, Network Norway, who provided GSM services, and Norwegian, who provided travel and profile services.
- A large number of content providers, who will provide news and travel information. A potential new service is location and context based information to travelers.

The complexity of integrating these very specialized services may appear large, but the bus structure and web services simplify the technical solutions. More interesting, in the context of information infrastructure innovation, which socio-technical mechanisms are active in the extending this infrastructure?

5 MECHANISMS FOR INNOVATING IN THE INFRASTRUCTURE

5.1 Returning to the research question: How can an information infrastructure provide generative mechanisms for innovation?

Following DeLanda’s call of investigation of both macro-micro and micro-macro mechanism; what we are looking for is a recursive structure of mechanisms. This structure should explain how the information infrastructure is generating innovation, and also how the innovations are modifying the information infrastructure. Further, these mechanisms should be “external”, in the sense that their capabilities are not decided from the components’ internal properties. For example, an IT architecture may have certain properties, such as being layered. This property is quite important, but it is not a mechanism. A mechanism emerges from the relationship between the IT architecture and some other
component, where the IT architecture plays a role in a larger context. For example, the IT architecture may allow someone to produce a new service. This does not (in itself) change the layer structure of the IT architecture, rather its ability to play different roles in different settings is more related to its interfaces.

5.2 Macro-micro mechanism

Several of the interviewees used the expression of space of possibilities, as the starting point for innovation. What constitutes this space? The respondents gave different answers. Some emphasized the business opportunity by logic of analogy; that a successful service (such as airline booking) is similar to the business of a mobile operator. Another informant pointed to the modularization and layers of the IT architecture, while a third informant emphasized the role of external partners in idea generation.

When examined more closely, the space of possibilities more fundamentally has to do with patterns of assembling different components into new services. The point is that these new services emerge through new combinations of components in ways that generate a self reinforcing mechanism. This mechanism was also identified (although at a business level, not service level) by (Davenport and Short 1990).

We may illustrate this process in figure 4.

![Figure 4. The innovation reinforcement mechanism](image)

The established information infrastructure (of Airline and Bank) offered a space of possibilities, constituted of the following elements:

- A large user group (ca 1 million individual customers), represented in the central CRM system
- An IT architecture, where components may be reused
- A (limited) number of key persons with a thorough knowledge of the dynamics of the information infrastructure, including a strong technical knowledge of the IT architecture.

The next step was to combine these elements into ideas, and from there to assemble the components into new services. More than 30 different companies from telecom and IT were involved in these talks. The Portal Director described the process in these terms:

“Based on the overall architecture we decided which components to build, reuse, buy or outsource. We involved various actors, such as IT, Sales, Marketing, expert consultants and
external vendors. Contracts with vendors were signed and integration specifications made. The design and development process was iterative, producing design sketches, arranging workshops, evaluating new builds, optimizing and testing.”

In this description we recognize the findings of Andersson et al. (2008), that inter-organisational innovation requires a demanding set of architectural knowledge. For example, the operator of the billing services had to a) understand the inner workings of both Norwegian’s corporate bus architecture and the Telebilling system and b) be able to implement a solution which connects these two resources.

5.3 Micro-macro mechanism

At the end of this process, illustrated with the upward arrow, we find the micro-macro mechanism. As the information infrastructure has generated a new service, this will extend the information infrastructure. Technically, this will in the form of a new bus with web service interfaces, a set of new databases and a new set of terminal devices. The information infrastructure will further include the vendors and operations, and the thousands of new customers.

The extended information infrastructure will increase the space of possibilities. To get the full picture we may extend the figure with another feedback loop, the self-reinforcing mechanism of services.

Figure 5: The double set of self-reinforcing mechanisms of information infrastructures

As illustrated in figure 5 the information infrastructure has two self-reinforcing mechanisms. First, there is the innovation reinforcement mechanism, which was described above. At macro (infrastructure) level the result of the mechanism is a new service which extends the information infrastructure. The Second mechanism is the service reinforcement mechanism, which is built on Grindle’s standard’s model, described in the Introduction. The result of this mechanism is more users to the information infrastructure.

The result is also a financial profit, which may be used to invest in more innovation. This closes the double loop. The innovation loop provides new services in the information infrastructure, which in turn (in the service loop) creates more profits. And so on.
5.4 Implications

The practical implications of self-reinforcing service mechanism have been described by other researchers, in the form of guidelines for information infrastructure design (Hanseth and Lyttinen 2008). The practical implications of the self-reinforcing innovation mechanism are more elusive, but two suggestions may serve as an inspiration for further research.

First, a successful information infrastructure constitutes a considerable resource for ICT-based service innovation. The innovation mechanism and the Call Norwegian case, illustrate how these resources may be assembled into new innovations. One should however not interpret this as a friction-free process, because the whole information infrastructure may be set under pressure as it expands, both in terms of IT architectural and social complexity. Using DeLanda’s (2006) terms, we might characterize it as a struggle between territorialisation (increased homogeneity) and deterritorialisation (destabilising).

Second, as described by Yoo et al. the innovation process in networks is both cognitive, social and non-linear (Yoo et al. 2008). In line with other information infrastructure research our findings illustrates that ICT-based service innovation cannot be planned and managed in detail. But the innovation mechanism may help organisations to facilitate the innovation process in a structured way.

6 CONCLUSION

This paper investigated innovation in information infrastructures, through a case study. Building on a critical realist approach, our empirical evidence was a case study within an international airline, aiming to diversify its services. From our analysis we propose that there are two self-reinforcement mechanisms in information infrastructures.

First, we identified the innovation reinforcement mechanism, resulting in a new service. This consists of the following steps: A space of possibilities in the information infrastructure architecture and operations creates new ideas for services. Together with external partners these may be developed into innovations, which will be included in the information infrastructure as new services. The new services and components in turn increase the space of possibilities.

Second, there is the service reinforcement mechanism, resulting in more users and profits. The information infrastructure provides a number of cheap and easy-to-use services. This attracts more partners with their add-on services, which reinforces the value of the information infrastructure, and attracts more users. This growth generates a profit which may be used to invest in new innovations.

The proposed mechanisms were derived from the information infrastructure literature and the case study. Further research should look at innovation in other information infrastructures, both to test the validity of the suggested mechanisms and to discover others.

7 ACKNOWLEDGEMENTS

The author thanks the Norwegian company for allowing me to conduct the case study, and the informants for spending their time. In particular I thank Hans-Petter Aanby and Håvard Haug Hansen. I also thank Ole Hanseth and Sturla Bakke for commenting on earlier drafts, and I acknowledge the useful comments of the anonymous ECIS reviewers and AE.

References


A CONCEPTUAL MODELING APPROACH FOR SUPPLY CHAIN EVENT MANAGEMENT (SCEM)

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0600.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
</tbody>
</table>
A CONCEPTUAL MODELING APPROACH FOR SUPPLY CHAIN EVENT MANAGEMENT (SCEM)

Winkelmann, Axel, University of Koblenz-Landau, Department of Computer Science, Institute for IS Research, Universitätsstraße 1, 56070 Koblenz, Germany, winkelmann@uni-koblenz.de

Fleischer, Stefan, University of Münster, European Research Center for Information Systems (ERCIS), Leonardo-Campus 3, 48149 Münster, Germany, stefan.fleischer@ercis.uni-muenster.de

Herwig, Sebastian, University of Münster, European Research Center for Information Systems (ERCIS), Leonardo-Campus 3, 48149 Münster, Germany, sebastian.herwig@ercis.uni-muenster.de

Becker, Jörg, University of Münster, European Research Center for Information Systems (ERCIS), Leonardo-Campus 3, 48149 Münster, Germany, joerg.becker@ercis.uni-muenster.de

Abstract

The management of supply chains is a very comprehensive task. Predictability and response to incidents in process executions are challenging. With Supply Chain Event Management (SCEM), researchers propose an approach for overcoming these problems through proactive monitoring and notification of crucial process activities across the supply chain. The identification and definition of such crucial activities and information needs are necessary to handle possible abnormalities although they are challenging task which have not been solved in SCEM research so far. Hence, we propose a modeling approach which allows the conceptual specification of an adequate information exchange along the supply chain. With our approach we focus on the specification of relevant logistical objects, the definition of possible events, and the design of notifications for decision makers in the supply chain.

Keywords: conceptual modeling, supply chain event management, SCEM.
1 MOTIVATIONS

Divisions of labor and distributed value chain activities require a good management of logistic networks as a strategic competitive advantage (Steven & Krüger 2001). The role of Supply Chain Management (SCM) increases. Interorganizational integration of processes, logistics, and information supply becomes necessary in order to efficiently cooperate (Bechtel & Jayaram 1997; Holten & Schulz 2001). The provisioning of information and information systems throughout the supply chain is a key point within SCM (Barrat 2004; Bechtel 1997; Knolmayer & Mertens & Zeier 2000; Göpfert 2002).

Characteristic information systems for SCM such as Advanced Planning and Scheduling Systems are mainly used for the planning and implementation of logistic processes but offer only limited functionality for the monitoring and management of the operative execution of logistic activities. On the contrary, Supply Chain Event Management (SCEM) allows the short-term planning, management and controlling of information regarding the operative logistic processes. It is a proactive concept that allows diverging high process complexity. Therewith, it helps overcoming a bad predictability or forecast reliability within the different steps of a value chain. It fosters the interorganizational visibility of critical objects throughout the supply chain (Nissen 2002; Wieser & Lauterbach 2001). Proactive and automated monitoring and management of supply chain critical objects (e.g. palettes, customer orders) help identifying failures and disruptions in material, product, and information flows. Hence, countermeasures can be induced earlier and more systematic (Wiener 2007). An increased reactivity from various parties or stakeholders of the supply chain to possible disruptions or events can increase an effective information supply and an accomplishment of superior business objectives (Klaus 2004; Steven & Krüger 2004).

An adequate and prompt supply of information is a prerequisite for the realization and execution of SCEM. Hence, the conception and implementation of appropriate information systems is a basic condition. Nevertheless, the identification of critical objects or steps in the supply chain, the specification of information needs to cover possible incidents in process executions and the adequate provisioning of responsible decision makers with necessary information to handle such incidents as well as the implementation of an appropriate information technology are challenging tasks in SCEM (Liu & Kumar & van der Aalst 2007; Nissen 2002; Steven & Krüger 2004). Although there is much research on the overall concept of SCEM and related information technology challenges, a methodical support for identification and definition of such crucial activities and information needs of persons responsible to handle possible disruptions is neglected up to now. Conceptual models serve as an instrument for structuring and documentation of such complex problems and encourage a common understanding between different stakeholders, e.g. managers, domain experts, and technical experts (Rosemann, 1996; Kottemann & Konsynski 1984; Karimi 1988). Therefore, the exploitation of such a modeling approach for the management of supply chain events—that is for the identification and specification of incidences and for the preparation of necessary notification based on information needs as well as their implementation in information systems—seems reasonable. The article proposes an approach for the conceptual modeling of interorganizational events within the supply chain.

2 RESEARCH METHOD

In literature, the usage of conceptual models for the documentation and construction of real world events is broadly accepted (Rosemann 1996). Hence, using such models and modeling methods for an interorganizational supply chain information management seems to be reasonable. Information systems engineering is a process of change in order to achieve certain goals in information systems design (Hirschheim & Klein & Lyytinen 1995). With this article, we will present certain representation forms of the object systems that are necessary for the systematic and exact conceptual modeling of SCEM in the context of interorganizational needs.
The research process (cf. Figure 1) started with the awareness of a practical problem (Takeda & Veerkamp & Tomiyama & Yoshikawa 1990). We were faced with the question of how to specify a management system for supply chain events on a conceptual level. The aim of such a research process was to provide a conceptual modeling approach guiding the development of a SCEM information system from functional specification to technical implementation. The practical need for such a method became apparent during discussions with various retailers about their experience with Collaborative Planning Forecasting and Replenishment (CPFR) and Efficient Consumer Response (ECR) and their problems with data exchange between different stakeholders (Winkelmann & Beverungen & Janiesch & Becker 2008). It became obvious that there were many obstacles in sharing all information between the stakeholders within the supply chain. It was not possible to individually define notifications and information allocations for each member of the supply chain. Hence, a conceptual approach turned out to facilitate a consensus between the different stakeholders about the information sharing within the supply chain. To solve the identified problem we were looking for a method that:

(R1) supports the conceptual modeling of event-driven information flows according to business processes, information needs and restrictions in the supply chain (Requirement 1),

(R2) facilitates the technical implementation of an SCEM information system in accordance to address different stakeholders (e.g. domain and technical experts) (Requirement 2).

Having this in mind, a literature review was performed in order to consider available research on conceptual modeling techniques for SCEM. Based on the requirements R1 and R2 we were able to identify existing modeling methods which are possibly capable of fulfilling these requirements. However, within a detailed analysis of the approaches based on the comparison process of (Song & Osterweil 1994) we came to the conclusion that none of the existing approaches fully met our requirements. Hence, we decided to develop a new methodological approach.

Figure 1. Research Procedure

Having this in mind, a literature review was performed in order to consider available research on conceptual modeling techniques for SCEM. Based on the requirements R1 and R2 we were able to identify existing modeling methods which are possibly capable of fulfilling these requirements. However, within a detailed analysis of the approaches based on the comparison process of (Song & Osterweil 1994) we came to the conclusion that none of the existing approaches fully met our requirements. Hence, we decided to develop a new methodological approach.
The construction process of the approach is similar to the procedures proposed by Greiffenberg (2004) and Gupta & Prakash (2001). Based on previous discussions, interviews and literature reviews, we derived language concepts that were necessary for supporting our requirements R1 and R2. Afterwards, the resulting concepts were supplemented with attributes and, thus, relations between the concepts were defined in order to establish the conceptual modeling approach. All elements of the conceptual modeling language have then been consolidated in common specification.

For the evaluation of the artifact, we implemented the model in order to visualize and apply its implications. For a practical assessment and a better method refinement, we evaluated the approach with the help of a case study in retail. Based on our requirements we generated the following propositions for the evaluation:

(P1) The application of our modeling approach improves supply chain spanning information exchange in the context of SCEM. (Proposition 1)

(P2) The application of the approach facilitates the implementation of an information system for the management of events in the supply chain. (Proposition 2)

(P3) The approach is capable of modeling the appropriate information needs of the persons responsible to handle occurring events of fault. (Proposition 3)

(P4) The approach allows a linkage to existing process modeling languages in order to specify process events and information that should be promptly processed at occurrence of such events. (Proposition 4)

The results from interviews, literature reviews, method implementation, and case study research were then compared to the propositions. We were able to derive new requirements on the approach, to identify deficiencies, and refine the concept by omitting and modifying language constructs.

Our way of conducting research is based on the design science paradigm (Hevner & March 2004). We aim at developing a domain neutral IT artifact, namely a modeling approach for the management of supply chain events. The research method is built on works of Takeda & Veerkamp & Tomiyama & Yoshikawa (1990), Rossi & Ramesh & Lyytinen & Tolvanen (2004) and Greiffenberg (2004), in which a practical problem is the focus of the construction process.

3 STATE-OF-THE-ART IN SUPPLY CHAIN EVENT MANAGEMENT

The term Supply Chain Event Management (SCEM) is widely used ambiguous. On the one hand, this term is characterized as an overall concept, which abstracts from specific technical aspects. On the other hand, the term SCEM is considered as a specific technical implementation of application systems (Alvarenga & Schoenthaler 2003). In the following, we abstract from technical details and rather focus on the conceptual perspective of SCEM. Thereby, the core aspects of SCEM and especially the major tasks of the conceptual design are taken into account to derive requirements for a conceptual modeling language.

From a conceptual point of view, SCEM is primarily subdivided into the core aspects monitoring, notification, simulation, controlling, and measuring (cf. Figure 2), which are representing the core functionalities of this approach (Knickle & Kemmeter 2002). The basic aspect of SCEM is the monitoring (1) of logistic objects within supply chains in respect of occurring events. A logistical object represents a logical or physical item (e.g. a customer’s order or a container), an activity within a logistic process as well as a whole logistic process. In order to monitor such objects, they serve as reference points for possible events. An event is an indicator for differences between actual and planned conditions of a parameter, which is monitored for a logistical reference object (Wieser & Lauterbach 2001; Alvarenga & Schoenthaler 2003). Possibly, the occurrence of events has effect on the planning and execution of other (up- and downsteam) logistic processes. To handle critical effects on relating processes, the automatic notification (2) of corresponding decision makers is necessary. Hence, providing relevant information to persons responsible is an essential prerequisite in order to
take countermeasures to resolve variations while process execution. Based on provided information, simulations (3) are used to verify and benchmark handling alternatives to resolve variations and resulting effects to other processes along the supply chain. Subsequently, a selected alternative of handling is implemented in logistic processes and the supply chain is adjusted (control (4)). To ensure planned conditions while process execution, measuring (5) functionalities allow continuous observations and analysis of actual and planned conditions along the whole supply chain.

Due to the core concepts of SCEM, the major tasks of the conceptual design are the definition of supply chain processes, the identification of relevant logistical reference objects and the definition of related events, as well as the design of status notifications with relevant information (again cf. Figure 2). Further, countermeasures are defined by developing business rules (Nissen 2002; Steven & Krüger 2004).

The initial point of the conceptual design of a SCEM solution is the coverage and definition of supply chain processes. These processes are perceived as a logical order of logistical activities and described hierarchical in different levels of abstraction (Steven & Krüger 2004). According to methodical facilitation of describing supply chain processes in supply chain management, methods and languages of business process modeling are also used for the definition of supply chain processes. Modeling languages like value chain diagrams or simplified process chains are mostly proposed for the description and modeling of such processes (Arndt 2008; Kruppe 2007). These approaches have in common, that they are mostly used to describe processes on a high level of abstraction. In contrast to this, detailed descriptions of supply chain processes are necessary for the identification and specification of relevant logistical reference objects and related events (Steven 2004). In the field of business process modeling, Event-driven Process Chains (EPC) (Scheer 2000), Petri Nets (Petri 1962) or the Business Process Modeling Notation (BPMN) (White, & Miers 2008) are established universal modeling approaches, which take these requirements into account. Using these modeling approaches supply chain processes can be described in detail as well as hierarchically structured on different levels of abstraction. In addition to these universal modeling approaches, with the Supply Chain Operations Reference (SCOR) (SCC 2008) a specific method for designing and modeling of supply chains is given. Therefore, the SCOR serves as a reference model for supply chains and provides predefined building blocks of supply chain processes. Additionally, supply chain related ratios are provided by the SCOR for monitoring and measuring purposes.

Especially, the identification of relevant logistical reference objects, the definition of events, and the design of status notifications are commonly referred as challenging tasks in SCEM (Nissen 2002; Liu & Kumar & van der Aalst 2007). In the literature, these tasks are often discussed from a more general point of view, conceptual or technical aspects are mostly neglected (Steven & Krüger 2004; Nissen 2002). To define events, monitoring relevant ratios have to be specified in relation to a logistical
reference object as well as planned conditions and critical notification levels. The design of status notifications comprehends the definition of organizational units (addressees), which have to be informed in case of incidents in process execution as well as the identification of necessary information needs to cover such interruptions. Methodical assistance of these tasks is only provided for specific aspects, e.g. the definition of events. For the modeling of events Liu & Kumar & van der Aalst (2007) propose the use of time-related Petri Nets. In this approach, only the formal definition of critical notification levels is addressed. Other aspects of SCEM are neglected.

To sum up, the need of a methodical support of the conceptual design and implementation of information systems in the context of SCEM is identified. Especially, a comprehensive view of the definition of events, the identification of logistical reference objects as well as the design of status notifications and the identification of information needs is necessary. In order to overcome the identified methodical shortcomings, we propose a modeling approach which allows the conceptual specification of an adequate information exchange along the supply chain and facilitates the technical implementation of information systems to benefit the realization of SCEM.

4 CONSTRUCTION AND APPLICATION OF THE MODELING APPROACH

4.1 Language Construction

To describe the events and the triggered notifications we have developed a modeling language by applying meta-modeling as a design technique (Strahlinger 1996). The Entity-Relationship Model (ERM) (Chen 1976) is a widely established approach for the conceptual representation of modeling languages (Becker & Seidel & Pfeiffer & Janiesch 2006). As a result of the development process, we constructed a modeling language whose conceptual specification is shown in Figure 3. Additionally, the implemented language constructs, relations, and the proposed modeling process are illustrated in Figure 4. The constructs are classified into different modeling contexts according to the different components and tasks of SCEM. For a better understanding, from now on, constructs are highlighted italic with capital initial letters; the steps of the modeling process are numbered in round brackets.

Figure 3. Conceptual Specification of Modeling Approach

First of all, Events are set in relation to Logistical Reference Objects and Rules in order to trigger the particular Event. Moreover, Addressees of dedicated Information Objects are also related to Events.

Proceedings ECIS 2009
Therefore, the relevant Logistical Reference Objects (1) are defined. After that, certain Ratios (2) are defined and combined with a Rule Expression to specify the triggering Rules (3). Those Rule Expressions are certain arithmetic combinations of Basic Ratios or Calculated Ratios that are assigned to specific Rules. For each assignment specific values are defined. For example, a rule applies if the stock of a product is lower than a specified threshold. Based on this, the actual Events (4) are modeled by relating them to Logistical Reference Objects and Rules. At this point, Logical Operators may combine those Rules to a combined and hence more complex expression. The Addressees (5) are defined and arranged in a hierarchical manner. They possibly receive Information Objects (6) which are modeled afterwards. Addressees are related to Status Notifications depending on the Context of Usage. It can be interpreted as some kind of filter that specifies under which contextual conditions the Addressees receive the relevant Information Objects. In the end, those contextual conditions are related to Status Notifications (7) in order to notify Addressees about certain Events. For example, an addressee may receive an information object A in case of an event X and receive an information object B in case of another event Y. Thus, sensitive information objects can be filtered and each addressee only receives the information objects with depending attributes, such as additional amount that has to be ordered.

To integrate the different modeling contexts, the concepts of the modeling language are specified context-spanning and hence can be re-used in suitable contexts. This concept prevents from illustration conflicts of occurrences of the same object and from modeling redundant issues. This integration of contexts allows a comprehensive view on the different tasks of supply chain event definition.

**Figure 4. Representational Specification of Modeling Approach**

As mentioned before, the definition of logistical reference objects, events and status notifications are parts of the overall SCEM tasks (cf. Section 3). To provide a linkage of these tasks to the definition of supply chain processes, a methodical integration of our approach with supply chain processes is necessary. Thus, defined events can be linked to specific states in supply chain processes in order to provide a detailed view and to monitor crucial states. For example, to monitor the order status in a particular process, a defined event can be related to the different states. E.g. an event to monitor extraordinary deliveries is related to process states. In case of e.g. abnormal amounts or quantities the defined event is triggered and the contracting as well as the order management will receive status notifications (cf. Figure 5).
4.2 Exemplified Application based on a Retail Example

For the purpose of application, the constructed modeling language is implemented using a meta-modeling tool. Figure 6 displays a screenshot of the (meta-) modeling software containing the corresponding model. A general introduction to the tool support is given in Becker & Janiesch & Seidel & Brelage (2005) and Delfmann & Janiesch & Knackstedt & Rieke & Seidel (2006). The exemplary case evaluation deals with two (unplanned) events. Both events are regular incidents at wholesalers. The first event is the result of a stock monitoring. In this example, the stock falls below a certain limit. With the help of our approach, we are able to specify an automatic preparation of a further delivery to the customer and an additional order at the pre-supplier of the wholesaler. The second event is triggered by the arrival of a delivery, whose billing amount differs from the declared amount. The same event is activated if the scheduled delivery date is not met, the quantity delivered is not as ordered, or the delivery has happened without any preceding order.

The logistical reference objects (1) relevant for this example are defined in a specific context (e.g. Commodities, Customers, Suppliers). As a next step, relevant information objects (6) are to be defined. Relevant information objects in this context are Preliminary Order Forms (Customer), Order Forms (Supplier), and Packaging Notes (Supplier). To monitor events, specific ratios are to be defined (2). For example, a particular ratio Stock representing a customer's stock is required. To determine whether an incoming delivery is a planned or an unplanned delivery other ratios are needed. E.g.
quantity, billing amount, and scheduled delivery dates have to be checked. For this purpose the definition of calculated ratios (e.g. Billing Amount as result of the quantity multiplied by the price of a single item plus a possible flat rate) are necessary. Moreover, the check criteria are defined by rules (3). For example, the billing amount has to be matched against the expected billing amount as declared. Based on the matching results, the rules trigger respective events. By combining logistical reference objects with those rules, the related events (4) have to be defined (e.g. Extraordinary Delivery). To this event definition four rules are assigned by combining them with a logical AND-operator (and negating by a NOT-operator). The rules are evaluated in a cascading manner. To handle possible event occurrences, organizational units, namely addresses, have to be defined (5). The addresses will receive individual notifications and information about abnormality in process execution. Defining such notifications (7), relevant information objects are related to particular addresses in order to handle a specific situation (event).

Figure 6. Exemplified Application of Modeling Approach

5 CONCLUSION, LIMITATIONS AND OUTLOOK

The concept of supply chain event management is a further development of existing supply chain management concepts. It does not deal with goods themselves but with the handling of incidents (events) concerning the flow of goods. SCEM allows the notification of people and organizations across organizational boundaries to provide a prompt response on defined events. For example, SCEM can help reducing the bullwhip effect because information cannot only be forwarded to the next stage in the supply chain but information can also be forwarded to pre-suppliers or pre-pre-suppliers. So far, the discussion on SCEM has been focused on a managerial or a process level. However, a conceptual approach for the identification and definition of information needs across a supply chain to handle
possible occurring incidents is not provided in the literature up to now. With our conceptual approach we propose a concept for the identification of suitable logistical reference objects and for the definition of relevant events with regard to this reference objects. It is possible to design status notification based on reference objects and events.

As our exemplified application indicates, our initial approach supports the conceptual modeling of event-driven information flows according to the needs and restrictions in supply chain communication. Furthermore, it facilitates the technical implementation in accordance to address different stakeholders (e.g. domain experts and technical experts) in the development process and to effect a common understanding of the treated object. The application of our conceptual approach based on a tool implementation is a first evaluation step on our research agenda (cf. Table 1).

<table>
<thead>
<tr>
<th>Proposition</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>The proposed modeling approach can improve a supply chain spanning information exchange in the context of SCEM by facilitating the definition of relevant logistical objects, events, rules, notifications and addressees.</td>
</tr>
<tr>
<td>P2</td>
<td>The application of the approach facilitates the implementation of information systems for the management of supply chain related events. Based on this conceptual approach the technological independent specification of relevant objects and entities is possible. This allow for the involvement of different stakeholders like domain and technical experts to ensure a consensus about the requirements on such an information system.</td>
</tr>
<tr>
<td>P3</td>
<td>By defining rules and triggers that can be allocated to individual needs of specific actors in the supply chain, the proposed approach is capable for the specification of appropriate information needs of persons responsible to handle occurring events of fault.</td>
</tr>
<tr>
<td>P4</td>
<td>Due to the process modeling language independent definition of events and triggers, a combination with existing process modeling languages in order to specify process events and process related information is possible.</td>
</tr>
</tbody>
</table>

Table 1. Approval of the proposed approach

However, with respect to the intended objective different assumptions are made during the development process of the conceptual modeling method. Consequently, different limitations are attended with the proposed modeling method and have to be taken into account for further applications. First of all, the approach is limited to a conceptual perspective on SCEM, excluding specific implementation aspects. In this way, the managerial, functional, and technical perspective of SCEM is addressed in common. Nevertheless, interrelations to the organizational and structural implementation and management of supply chains are not covered by the proposed approach. I.e. in our approach linkage between supply chain processes and events are only defined in a logical way. A methodical integration with process modeling languages is still outstanding. As a result of this, effects of changes in supply chain arrangements and process could not be taken into consideration of the specification of SCEM. Reverse, the derivation of further details for development and improvement of existing supply chain structures is not methodically supported up to now. Furthermore, the proposed approach is focused on the field of supply chain management. With respect to unstructured domains like creative industries or research and development, the processes in supply chains are commonly structured. Hence, the application of our approach is focused on structured processes with clearly identifiable decision making entities, clearly delineated areas of authority, classifiable and predictable (planned or unplanned) events and event variations and predictable responses to such events. Unpredictable events should not be taken into consideration. An ex-ante definition of all possible but not predictable events seems to be not appropriate and would lead beyond the intention of the SCEM approach. Even though, results of such events can still be regarded.

Further research will focus on the evaluation of the provided modeling approach and on aspects of the before mentioned limitations. In the short-term, the approach will be applied in different domains and different application scenarios. In particular, the capability of our approach to cover the requisite aspects of SCEM to improve the event-driven information flows and handling of incidents across supply chains is evaluated. In a next step, possibilities and requirements for the connection with or
integration into supply chain tools are reviewed. Middle-term research will address the methodical integration with common process modeling languages. Furthermore, consequence of changes in supply chain structures and processes to the definition of SCEM will be addressed by identifying interdependencies and developing response mechanisms. As a long-term research field, the approach would be enhanced by a methodical support for the improvement of existing supply chain arrangements, e.g. the identification of potential risks in supply chain execution.

References


A REAL OPTIONS APPROACH FOR VALUATING INTERTEMPORAL INTERDEPENDENCIES WITHIN A VALUE-BASED IT PORTFOLIO MANAGEMENT - A RISK-RETURN PERSPECTIVE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0518.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business value of IT, Design Science, Portfolio management, Real Options</td>
</tr>
</tbody>
</table>
A REAL OPTIONS APPROACH FOR VALUATING INTERTEMPORAL INTERDEPENDENCIES WITHIN A VALUE-BASED IT PORTFOLIO MANAGEMENT – A RISK-RETURN PERSPECTIVE

Diepold, Dennis, University of Augsburg, Universitätsstraße 16, 86159 Augsburg, Germany, dennis.diepold@wiwi.uni-augsburg.de
Ullrich, Christian, University of Augsburg, Universitätsstraße 16, 86159 Augsburg, Germany, christian.ullrich@wiwi.uni-augsburg.de
Wehrmann, Alexander, Senacor Technologies AG, Vordere Cramergasse 11, 90478 Nürnberg, Germany, alexander.wehrmann@senacor.com
Zimmermann, Steffen, University of Innsbruck, Universitätsstraße 15, 6020 Innsbruck, Austria, steffen.zimmermann@uibk.ac.at

Abstract

Value-based IT portfolio management requires the consideration of intertemporal interdependencies that may exist among IT projects. Therefore, several papers suggest adopting the real options approach in order to include intertemporal interdependencies within the valuation of IT projects. However, this paper shows that the standard Black-Scholes model, which is often used for valuating real options, is not appropriate to correctly account for project-specific private risks due to its restrictive assumptions. Since this can have major impacts on the value of IT projects, we develop an approach – based on the Black-Scholes model – to consider private risks properly within project valuation. A comparison of the results of the standard Black-Scholes model used today and our approach finally reveals that the neglect of private risks results in a systematic underestimation of both risk and return of IT projects, which may lead to wrong investment decisions.

Keywords: Value-based IT portfolio management, intertemporal interdependencies, real options, Black-Scholes model, private risks
1 INTRODUCTION

Since firms with superior IT governance have at least 20 percent higher profits than firms with poor governance (Weill & Ross 2004), it is not surprising that in practice much effort is put towards implementing IT governance structures (IT Governance Institute 2008). Amongst others, companies seek to implement methods to plan and manage IT investments aligned with their business objectives. Since a lot of firms primarily seek to maximize shareholder value, they need to determine the value proposition of each IT investment considering its risk and return (IT Governance Institute 2008). But it is not sufficient to valuate each IT investment separately, because firms usually conduct several IT investments simultaneously or consecutively, which may cause the existence of interdependencies among these investments. Due to the fact that those interdependencies affect the value of a single investment, firms should rather implement methods to valuate the overall IT portfolio and make sure that interdependencies are considered correctly.

But as of today, only about half the firms are capable of measuring risk and return of their IT investments (IT Governance Institute 2008) and much less they are capable of accounting for interdependencies among multiple investments within a value-based IT portfolio management (ITPM). But interdependencies and conflicts among IT projects in particular are one of the primary reasons for budget over-spending, which affects about one third of all IT investments according to a study of CA Inc. (2007). Interdependencies also play a decisive role when investments in IT infrastructure (e.g. operating systems or core banking systems) are considered, which account for 31% of all IT investments (CIO Insight 2004). IT infrastructure investments are typically characterized by high cash out-flows and – if at all – only low direct cash in-flows. If firms valuate these investments without considering interdependencies, it is likely that they will be rejected from an economic perspective. Given that business objectives may require IT infrastructure investments (base projects) that provide an option to launch future value added projects (follow-up projects), the profits of the follow-up projects can be attributed to the base projects to some extent. Such intertemporal interdependencies are often modelled as real options. Research literature suggests adapting approaches from financial theory to the valuation of real options, but does not thoroughly verify the applicability of these approaches.

Therefore, this paper contributes to valuating IT base projects that contain intertemporal interdependencies with an advanced real options approach. After a discussion about the applicability of the Black-Scholes model (BSM) to the valuation of real options, we conclude that strict assumptions of the BSM prevent its application to a correct valuation of base projects, since project-specific risks (e.g. quality risks) that influence the value of possible follow-up projects cannot be considered. Hence, we extend the BSM in a way that it will be capable of valuating base projects correctly. Furthermore, we will show that the BSM – as it is used today for valuating base projects – underestimates the return of any base project as well as the associated risk due to the disregard of project-specific risks.

This paper is organized as follows: In chapter 2 we provide an overview of existing ITPM approaches. Thereby, we analyze how intertemporal interdependencies are considered today and question the applicability of the BSM to the valuation of IT projects that contain intertemporal interdependencies. In chapter 3, a model extending the BSM by a correct consideration of project-specific risks is proposed. The article concludes in chapter 4 with a recapitulation of the achieved results. Thereby, the limitations of the model as well as perspectives for further research are discussed. A real-world example will serve as running-example throughout the paper and underpin the relevance of our findings.
LITERATURE SURVEY AND RESEARCH QUESTION

IT governance, which is defined as „structure of relationships and processes to direct and control the enterprise in order to achieve the enterprise’s goals by adding value while balancing risk versus return over IT and its processes“ (IT Governance Institute 2008), postulates value-based approaches to manage IT investments. This is in accordance with the definition of ITPM by Kaplan (2005), who refers to ITPM as a „method for governing IT investments across the organization, and managing them for value“. According to these definitions firms have to valuate their IT investments as well as their overall IT portfolio under a risk-return perspective. Since existing interdependencies among IT investments can affect the value proposition of the investments (Santhanam & Kyparisis 1996), firms also have to consider them. These interdependencies can be categorized as follows:

- **Intratemporal interdependencies:**
  These interdependencies occur due to resource conflicts or structural bottlenecks (e. g. use of same processes or IT functionalities) in case that multiple IT projects are conducted at the same time.

- **Intertemporal interdependencies:**
  These interdependencies occur if IT projects serve as basis for potential follow-up projects.

As a result, value-based ITPM approaches require firms to focus on a risk-return perspective of their IT investments on the one hand, but also to consider inter- and intratemporal interdependencies among IT investments on the other hand.

2.1 Value-based ITPM: Status quo

Since value-based ITPM requires the quantification of both risk and return of IT investments, those approaches are also referred to as quantitative ITPM approaches. Verhoef (2005) for example uses the Net Present Value to valuate IT projects. He includes risk by introducing the “weighted average cost of IT” as discount factor, but he does not consider any interdependencies among IT investments, which makes this approach insufficient for a value-based ITPM due to the requirements mentioned above.

Santhanam & Kyparisis (1996), Butler et al. (1999), and Asundi & Kazman (2001) include intratemporal interdependencies among IT projects in their approaches. They use modern portfolio theory by Markowitz to aggregate IT projects to IT portfolios and intratemporal interdependencies are represented by correlations among IT projects. But these approaches still neglect intertemporal interdependencies, which leads to poor or incorrect valuations especially for infrastructure projects that are characterized by low or even none direct cash in-flows.

Bardhan & Bagchi & Sougstad (2004) in contrast focus on intertemporal interdependencies among IT projects. They assume that a firm has the right – but not the obligation – to conduct possible follow-up projects after the completion of a base project. This right is modelled as a real option.

2.2 Real Option Approaches for IT portfolio management

There exist many other approaches for IT portfolio management using real options theory. E. g. Benaroch (2002) suggests a real options approach called option based risk management (OBRIM) to mitigate risks of IT projects, which is empirically validated by Benaroch et al. (2006) and Hilhorst et al. (2008). But this approach is “not concerned with determining the monetary value that embedded options add to an investment” (Benaroch et al., 2006). Thus it is insufficient according to the requirements and concerns of this paper.

On the contrary, Taudes & Feurstein & Mild (2000), Benaroch & Kauffman (1999) and Fichman & Keil & Tiwana (2005) suggest the application of real options analysis (ROA) to the valuation of IT
projects analogical to the approach of Bardhan & Bagchi & Sougstad (2004) and fulfil consequently the requirements of a value-based ITPM. All these approaches use the standard BSM or binomial trees to valuate existing real options.

But the application of valuation models like the BSM, which is adapted from financial options to real options theory, is heavily criticized due to its restrictive assumptions (Emery et al. 1978, Schwartz & Zozaya-Gorostiza 2003). Thus, the differences between financial and real options must be regarded properly, otherwise the application of the BSM to the valuation of intertemporal interdependencies can lead to skewed results. We therefore discuss the applicability of the BSM to the valuation of intertemporal interdependencies in the next chapter.

2.3 Applicability of the Black-Scholes model to the valuation of intertemporal interdependencies

The BSM is based on a riskless valuation of the option, whereby systematic risks are eliminated through a replicating portfolio consisting of the underlying and the option (Hull 2003). In order to be able to build this replicating portfolio and to continually hedge it during the runtime of the option, liquidity of the involved assets is a key requirement. But since real options usually cannot be traded and thus are illiquid (especially in the case of IT projects), the replicating portfolio cannot be built, which raises critics about the applicability of the BSM to the valuation of real options.\[1\]

Sick (2001) picks up this criticism and argues that the replicating portfolio does not necessarily have to consist of the option and its underlying per se. In fact, any liquid assets can be used for constructing the replicating portfolio, as long as they possess the same systematic risk. Therefore, trading real options is not necessary for a correct application of the BSM. Hence, the BSM can be used for the valuation of real options in case that the systematic risks can be replicated by tradable assets, which accordingly requires a complete market.

Therefore, the application of the BSM implies that the underlying investments contain only systematic risks (market risks). This issue is also raised by Copeland & Antikarov (2003), who argue that with both financial and real options, risk - the uncertainty of the underlying - is assumed to be exogenous. This represents one major weakness of today’s application of the BSM to the valuation of real options, since there are also unsystematic risks inherent in every IT investment, which are referred to as “private risks” by Smith & Nau (1995). Those private risks or project-specific risks, like for instance deficient software quality, incorrect interpreted specifications, or problems with new technologies or frameworks, account for the major source of all risks concerning IT investments.

These risks cannot be considered within the replicating portfolio, since there are no liquid assets that perfectly replicate the private risks of the base project due to their uniqueness. As a result, the BSM cannot account for private risks and therefore neglects a major source of risks in the valuation of IT investments (Smith & Nau 1995).

In order to address this weakness, Smith & Nau (1995) suggest changing the assumption of a complete market into a partially complete market, which still provides liquid assets to account for market risks. Thus, the BSM can still be applied to the valuation of real options and does consider market risks, but private risks must be incorporated otherwise.

Irrespective of these facts, several articles apply the BSM to the valuation of real options without paying attention to its applicability and thus disregard the differences between financial and real options. Mason & Merton (1985) argue that although the underlying is not traded, firms rather seek to determine what the project cash flows would be worth if they were traded. A similar qualitative

\[1\] For a thorough explanation of the BSM see Hull (2003).
discussion, which eventually equates real options theory with financial options theory, can also be found in Benaroch & Kauffman (1999) and Taudes & Feurstein & Mild (2000).

In these articles the expected value of the cash in-flows (of the follow-up project in our case) usually serves as underlying of the option, since there is no observable market value (Schwartz & Zozaya-Gorostiza 2003). This expected underlying value contains the potential impacts of market and private risks. It can be determined by specifying scenarios and valuating them through a decision tree analysis. Possible deviations from the expected underlying value, which are caused by market risks as well as private risks, are oftentimes considered solely within the volatility of the BSM (cp. Bardhan & Bagchi & Sougstad 2004). But – as we have discussed above – this is not valid since only market risks can be hedged in a replicating portfolio of liquid assets. Taudes & Feurstein & Mild (2000) and Benaroch & Kauffman (1999) do not pick up the different risk types as a central theme in their articles. If we assume benevolently that the authors consider only market risks within the volatility of the BSM, their application would be consistent to the assumptions of the standard BSM. However, in this case they would completely disregard private risks in their valuation, which leads to skewed results because a significant part of IT project risks are neglected.

2.4 Research Question

Because of the major weakness of the standard BSM regarding the consideration of private risks, we will answer the question, how intertemporal interdependencies can be correctly valuated using the BSM within the scope of a value-based ITPM. Therefore, it is necessary to extend the BSM by a correct consideration of the impacts of private risks of the underlying (follow-up project) on the risk-return position of a base project.

But before answering this research question using our approach, we introduce a real world example to illustrate the relevance of this question. This example is taken from an IT Portfolio of a German retail bank, which invests a high binary million amount per year into IT projects. For reasons of confidence we changed the data proportional to the original values. The example will be continued throughout the paper.

A multi-channel retail bank wants to enhance its market position (relative market share) in distributing consumer credits. To reach this strategic objective, they want to increase the level of automation of their credit processes and enable a risk-adjusted pricing of consumer credits. Therefore, existing credit processes have to be redesigned, which requires the adaption of the IT landscape. First of all, their core banking backend-systems has to be changed. The costs for this infrastructure project are estimated at 2 million Euros. Furthermore, the investment into the backend-system does not generate any direct cash in-flows, which leads to the fact that it should be rejected from an economic perspective if the project is considered independently. But the bank decided to base their investment decision not solely on the NPV of this infrastructure project. The company rather considers this project to serve as a base project that provides the launch of future value adding project opportunities, which can be realized once the backend-systems are implemented successfully. Thereby, the company identifies a follow-up project that integrates the new credit-pricing into the existing retail frontends (i.e. in-store, online, and call-center) as a lucrative opportunity and decides to include this possible follow-up project within their investment decision. In contrast to the base project, the bank anticipates high cash in-flows from the follow-up project due to the involved launch of new credit products as well as savings on human resources due to the automation of credit processes. In order to integrate this possible follow-up project into the investment decision, the company wants to use the real options approach to evaluate the base project with the follow-up project being an option to expand.
3 CONSIDERATION OF PRIVATE RISKS WITHIN THE VALUATION OF INTERTEMPORAL INTERDEPENDENCIES

To illustrate our model and the impact of private risks on the risk-return position of a base project, we firstly have to introduce some notations and assumptions.

3.1 Notations and Assumptions

A firm has to decide, whether it invests in an IT project (base project) at time $t = 0$ with a runtime of $T$ periods. This base project creates the technical requirements for a possible follow-up project. In case of conducting the base project, the firm can decide in $t = T$ whether it invests in the follow-up project or not. This can be interpreted as an intertemporal interdependency and thus be modelled as a real option (option to expand) on the follow-up project. Because the focus of this paper is the correct valuation of the real option using the BSM, we state the following simplistic assumption:

(A1) The direct cash flows and thus the isolated Net Present Value ($NPV$) of the base project (without considering the impacts of the follow-up project) are known and fixed.

In order to valuate the real option under a risk-return perspective we have to consider the risks concerning the follow-up project. During the runtime of the base project two major types of risk exist which cause uncertainty regarding the cash in-flows of the follow-up project. The first type of risk can be described as market risks, which – as we mentioned earlier – can be considered by the volatility of the standard BSM. Examples for market risks are uncertainties regarding economic conditions like the prime rate or the demand since they are subject to fluctuations. The second type of risk – private risks – cannot be considered in the standard BSM as discussed above. It results from uncertainties regarding the implementation quality of the base project. Some examples of those uncertainties are:

- Uncertainty regarding the requirements of the base project: At the beginning of the base project it is not conceivable whether the functional or technical specifications describe the requirements unambiguously. Missing functionalities due to missing or incomplete functional specifications will limit the amount of possible subsequent applications.
- Uncertainty regarding the replacement of legacy systems: If (poor documented) legacy systems have to be replaced, unpredictable side effects can occur and reduce the scope of available functionalities.
- Uncertainty regarding the product quality: Irrespective of the uncertainties mentioned before, the implementation itself can be inaccurate. If too many critical mistakes are made during the implementation, it is likely that the scope of available functionalities is reduced.

These uncertainties can be responsible for providing an insufficient amount of functionalities at time $t = T$. The fact that the cash in-flows of the follow-up project (underlying of the real option) depend on the implementation quality of the base project leads to our second assumption:

(A2) At time $t = 0$, there is a known functional connection between the achieved implementation quality of the base project at time $t = T$ and the cash in-flows of the follow-up project. The present value of the cash in-flows of the follow-up project at time $t = 0$ is represented by the non-negative random variable $\tilde{S}_0$ with its known corresponding density function $f(s)$.\footnote{$\tilde{S}_0$ can also be a discrete random variable with probability mass function $f(s)$.}
Function $f(s)$ represents the potential impacts of the private risks on the present value of the cash inflows of the follow-up project.

Since the follow-up project will be conducted by an IT services provider and is already contractually fixed in $t = 0$, the cash out-flows of the follow-up project can be considered as independent of market and private risks, which leads to the following assumption:

\[(A3)\] The present value of the cash out-flows $X_0$ of the follow-up project is known and fixed at time $t = 0$.

This paper focuses on the correct consideration of private risks during the runtime of the base project, which affect the cash in-flows of a follow-up project. Therefore we abstract away from the existence of risks during the runtime of the follow-up project with the following assumption.

\[(A4)\] The present value of the cash in-flows of the follow-up project is known and fixed at time $t = T$ (depending on the implementation quality of the base project).

On the basis of these assumptions we will discuss the impacts of private risks on the risk–return position of the real option and consequently on the base project in the next chapter. Therefore, we compare the results of the standard ROA with our approach.3

3.2 Impacts of private risks on the risk-return position of the base project

According to standard ROA, the extended net present value ($ENPV$) denotes the return of the base project. It consists of the isolated net present value of the base project ($NPV$) and the value of the option ($C_0$) to extend the base project with a follow-up project (Trigeorgis 1996). In order to calculate $C_0$ the value of the underlying is required. But since the underlying is not traded on a market, the expected value of the underlying ($E(\tilde{S}_0)$) is often being used instead (Schwartz & Zozaya-Gorostiza 2003). Bardhan & Bagchi & Sougstad (2004) therefore approximate $E(\tilde{S}_0)$ by conducting a scenario analysis. According to the BSM function $c(s)$, which is described in appendix A-2, the value of the option ($C_0$) then equals $c[E(\tilde{S}_0)]$ according to our notation. As a result, the return of the base project calculated with the standard BSM can be obtained through the following equation:

\[
ENPV = NPV + C_0 = NPV + c[E(\tilde{S}_0)]
\]

According to this approach the value of the option depends on the risk, but it still can be calculated deterministically, because all risks are hedged by the replicating portfolio. But note that equation (1) is only valid in case of a complete market, where only market risks exist, which can be replicated by traded securities, and private risks do not. But in a partially complete market private risks have to be considered separately.

Bardhan & Bagchi & Sougstad (2004), for instance, consider private risks within their scenario analysis for the calculation of $E(\tilde{S}_0)$. They include the deviation of the underlying value within the volatility of the BSM. But unfortunately, this approach cannot account for private risks since only market risks can be considered within the BSM (cp. chapter 2). Other approaches neglect the impacts of private risks completely (e. g. Benaroch & Kauffman 1999), which leads to wrong investment decisions if a value-based valuation takes place.

---

3 The notation used in this paper is summarized in appendix A-1.
4 Even though the standard ROA does not explicitly account for $\tilde{S}_0$, we denote the expected value of the underlying as $E(\tilde{S}_0)$ according to our notation, as it provides an intuitive description as well as a better comparison of the two approaches.
In order to correctly account for private risks, we first consider the present value of the cash in-flows of the follow-up project being a non-negative random variable \( S_0 \) due to its uncertainty according to \((A2)\). Since the corresponding and known density function \( f(s) \) denotes the different realizations of \( S_0 \) \( (s^*) \), we can picture all possible impacts of the private risks on the cash in-flows of the follow-up project.

In a next step we need to find out how this uncertainty of the underlying affects the option value. Therefore, we need to focus on the functional connection (in our case the BSM function \( c(s) \)) between the underlying and the option value. As the BSM function cannot handle private risks, an option value has to be generated for every possible realization \( s^* \). Since every option value \( c(s^*) \) has the same cumulated probability as its corresponding underlying value \( s^* \), we obtain \( \int_0^{c(s^*)} g(c)dc = \int_0^{s^*} f(s)ds \) for all \( s^* > 0 \), and thus we can sufficiently approximate the density function \( g(c) \) of the option value \( C_0 \), which is also a random variable. According to this transformation we now know the different impacts of the private risks on the option value. Finally, we have to add the option value \( C_0 \) to the net present value of the base project \( (NPV) \) to get the extended net present value of the base project \( (E\tilde{NPV}) \), which is consequently represented by a random variable:

\[
E\tilde{NPV} = NPV + \tilde{C}_0 = NPV + c(\tilde{S}_0)
\]

At the time when the decision about the base project is made \( (t = 0) \) we only know the approximated density function \( g(c) \) of the option value and thus the probability for every possible option value, but we do not know the actual realization in \( t = T \). In order to account for this uncertainty caused by private risks, we will use the expected option value \( E(C_0) \) to calculate the expected extended net present value of the base project \( (E\tilde{NPV}) \), since it includes all possible realizations of the option value that are caused by private risks. By doing so, \( E(\tilde{NPV}) \) represents the return of the investment decision and can be obtained by the following equation:

\[
E(\tilde{NPV}) = NPV + E(C_0) = NPV + E[c(\tilde{S}_0)]
\]

In order to compare the return of the base project obtained by our approach (equation (3)) with the standard BSM (equation (1)), we need to go into a more detailed analysis of the BSM function. In case of call options the first and second derivative (i.e. the greeks “delta” and “gamma”) of the BSM function are positive (Hull 2003), thus we know that the BSM function is strictly monotonically increasing and strictly convex. Therefore, Jensen’s inequality, as is denoted in equation (4), becomes valid:

\[
E(c(\tilde{S}_0)) \geq c[E(\tilde{S}_0)]
\]

Since the BSM function is strictly convex, Jensen’s inequality can be rewritten in our case as \( E(c(\tilde{S}_0)) > c[E(\tilde{S}_0)] \), which implies:

\[
E(\tilde{NPV}) = NPV + E[c(\tilde{S}_0)] > NPV + c[E(\tilde{S}_0)] = ENPV.
\]

This leads to our first result:

(R1): The expected value of the real option and therefore the return of the base project are underestimated if the standard BSM (equation (1)) is used to valuate intertemporal interdependencies.
The return of the option derived by both, the standard BSM (equation (1)) and our approach (equation (3)), are pictured in Figure 1. It further shows the density function of the underlying $f(s)$ and the density function of the real option $g(c)$, as well as their functional connection through the BSM function $c(s)$.

![Graph showing the connection between $f(s)$, $g(c)$, and $c(s)$](image)

Figure 1 – Impacts of the private risks on the option value

Figure 1 also visualizes the consequence of Jensen’s inequality: The density function $g(c)$ results from a compression of $f(s)$, which is stronger for small values than it is for larger ones due the convexity of the BSM function $c(s)$ (The skewness of $g(c)$ is greater than the skewness of $f(s)$). This leads to the fact that the cumulated probability for $E(\tilde{C}_0)$ is greater than the cumulated probability for $E(\tilde{S}_0)$. But since the cumulated probability for $s^* = E(\tilde{S}_0)$ equals the cumulated probability for $c(s^*) = c[E(\tilde{S}_0)]$, $E(\tilde{C}_0)$ must be greater than $c[E(\tilde{S}_0)]$.

The bank decides to evaluate the possible follow-up project (option to expand) with the BSM, which created some challenges due to the collection of the data. Although the costs of the follow-up project, the risk-free rate, and the runtime of the base project could be estimated easily, the estimation of the discounted cash in-flows of the follow-up project (underlying of the option) seemed to be a major challenge, because they are based on the sale of new credit products. In order to address the uncertain success of the project, a scenario analysis (consisting of a worst-case, most-likely, and best-case scenario) was conducted (cp. Table 2) that should cover possible outcomes. By doing so, the bank derived an expected value of the discounted cash in-flows in the amount of 5.7 million Euros. Furthermore, the estimation of the volatility of the discounted cash in-flows were another major challenge: Since the volatility can only account for market risks, the bank used the volatility of a credit derivative index, because it was believed that this volatility represents all relevant risks of the credit markets.

An overview of the estimated values of the parameters needed for the BSM is provided below:
Expected Present Value of Cash In-flows ($E(S_0)$) & Costs ($X_0$) & Risk-free rate ($r$) & Runtime ($T$) & Volatility ($s$) \\
5.7 mio. Euros & 4 mio. Euros & 5% & 1 year & 40%

Table 1. BSM Parameters

According to the standard BSM the option value obtained by the bank was 2.05 million Euros ($C[E(S_0)]$). Since the option value outweighs the negative NPV, the consideration of the intertemporal interdependency resulted in a positive value for the investment decision.

(a) $ENPV = -2$ mio. Euros + 2.05 mio. Euros = 0.05 mio. Euros

In the following, this result will be compared to the result the bank would have achieved if they had considered private risks correctly. According to our approach, the different scenarios represent the private risks, and thus an option value has to be determined for each scenario. Based on the resulting realizations of the option values the expected value of the real option has to be derived by summerizing the realizations weighted with their adherent probabilities.

<table>
<thead>
<tr>
<th>Scenario:</th>
<th>Present Value of Cash in-flows</th>
<th>Expected Present Value of the cash in-flows</th>
<th>Option value</th>
<th>Expected value of the option</th>
</tr>
</thead>
<tbody>
<tr>
<td>most-likely</td>
<td>6 mio. Euros</td>
<td>5.70 mio. Euros</td>
<td>2.31 mio. Euros</td>
<td></td>
</tr>
<tr>
<td>worst-case</td>
<td>3 mio. Euros</td>
<td>5.70 mio. Euros</td>
<td>0.23 mio. Euros</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Scenario Analysis and Option Valuation

As we can see from Table 2, the expected option value calculated with our approach equals 2.26 million Euros. This leads to the following equation:

(a') $E(ENPV) = -2$ mio. Euros + 2.26 mio. Euros = 0.26 mio. Euros

Illustrated by this real world example we can state, that considering private risks correctly can lead to a fundamental increase of the return of the base project (in this case by approx. 420 %).

The standard ROA approach also suggests that the value of the option ($C_0 = c[E(S_0)]$) is riskless, which means the resulting value of the base project ($ENPV$) is a fix return on the investment. Thus, no consideration of further risks concerning the investment decision is required. Our approach enables the consideration of private risks, which are the major source for uncertainty regarding IT projects. Hence, we modeled the value of the option as a random variable ($\tilde{C}_0 = c(\tilde{S}_0)$), which means the value of the base project is also a random variable ($\tilde{ENPV}$) with its expected value $E(\tilde{ENPV})$. This expected value represents the return on the base project. But the possible deviations of $\tilde{ENPV}$ and so the private risk must be included in the investment decision about the base project. This leads to our second result:

5 The number for the volatility was adapted from Wigan (2006).
3.3 Consequences for the investment decision

Because of the uncertainty of the value of the base project resulting from private risks, the return or expected extended net present value is not sufficient for a risk-return integrated valuation within the scope of a value-based ITPM. Hence, it is necessary to balance risk versus return of the investment decision in order to get a thorough result. Therefore, the deviation of the project value resulting from our approach has to be opposed to the higher return in comparison to the standard BSM to get a risk adjusted value of the base project. The impact of this uncertainty on the risk adjusted value (value proposition), which results from private risks, heavily depends on the risk attitude of the decision maker.

If the decision maker is risk neutral or risk seeking, our approach leads to a higher value proposition for the base project compared to the standard ROA. This results from the increase of the return on the investment, which is shown in equation (4). However, if companies put efforts towards identifying and considering risks, they are aware of risk. Through this risk awareness and the consequent risk management such companies can be assumed to be risk averse. In this case, the consideration of the private risks within a risk-return integrated base project valuation will lower the return on the investment. In this case, the value of the base project can be lower than using the standard ROA, depending on the degree of risk aversion – which symbolizes the perception of the weight of private risks – and the specific utility function that the decision maker chooses in order to integrate risk and return.

Using the standard BSM, the bank based its investment decision on one deterministic value for the ENPV (cf. equation (a)) and thus neglected the existence of private risks. Based on this positive ENPV the bank decided to conduct the project.

But, according to our approach, the Bank should have considered private risks separately and carefully balanced their decision dependent on their risk attitude. As the bank also suffers from the subprime crisis, the company is currently risk averse. Therefore, a monetary valuation of the private risks regarding the degree of risk aversion would have been required. Based on such a risk valuation, the bank could have also rejected the base project, if the private risks had been assigned according to the degree of risk aversion with a monetary value greater than 0.26 million Euros. This would have led to the fact that the risk adjusted value of the base project is negative – although a higher return was realized compared to the result of the standard BSM.

Finally, we can state that the valuation of the intertemporal interdependencies with the BSM, as it is done today, can lead to false estimations and therefore to wrong investment decisions due to the neglect of the impacts of private risks on the risk-return position of the investment.

4 CONCLUSION

In this paper we propose an approach to consider intertemporal interdependencies within a value-based IT portfolio management. Therefore we analyzed current approaches of valuating intertemporal interdependencies among IT projects. Since intertemporal interdependencies are usually modelled as real options and valuated with the BSM, which is adapted from financial options theory, we examined the applicability of the BSM to real options. Through this examination we revealed that the BSM is only able to consider market risks, but not to consider existing private risks of IT projects. Since private risks mostly preponderate in case of IT projects, we developed an approach based on the BSM that allows the consideration of private risks and thus provides a correct valuation of intertemporal
interdependencies. The main findings in this paper are that today’s real option approaches based on the BSM underestimate the risk as well as the return of IT projects systematically. Hence, the application of the standard BSM can lead to false investment decisions.

4.1 Limitations

But there are still limitations that come along with the introduced model.

E.g. the assumption of fixed cash in-flows of the follow-up project at time $t = T$ (A4) leads to a neglect of any risks that can occur during the follow-up project. As software is usually developed in sequent releases which contain intertemporal dependent projects and in order to provide a thorough assessment of all risks involved in the IT portfolio, the consideration of follow-up project risks would further increase the results of this approach.

Another limitation comes along with assumption (A2), where we assume, that the cash inflows of the follow-up project depend only on the implementation quality of the base project at time $t = T$. But there may exist different other aspects which influence the cash inflows of a follow-up project. These aspects cannot be considered within our approach.

Furthermore, we are not able to calculate a concrete risk adjusted value of the base project, since we do not model the risk attitude and the risk itself by a specific risk measure.

4.2 Further Research

In order to overcome these limitations further research is required. To cope with the last mentioned limitation, we want to extend this paper by developing and introducing a decision model. Therefore, private risks need to be quantified by an adequate risk measure in a first step. Since firms usually focus on downside risks, the lower partial moment (LPM) can serve as an adequate risk measure. Otherwise, if firms also want to take the upside chances of a base project into account, a symmetric risk measure like the standard deviation can be chosen. Once a risk measure is selected, a preference function will be introduced that accounts for the investor’s risk preference. Founded preference functions for different risk measures can be found in decision theory literature.

Since firms not only conduct projects consecutively but also parallel at the same time, further research is needed to additionally consider intratemporal interdependencies among projects. The approach suggested by Santhanam & Kyparisis (1996) can serve as a fundament for this extension.

Further research is also needed regarding the consideration of multiple real options within an IT project and their impact on its risk and return. Smit and Trigeorgis (2006) already propose an approach to manage a portfolio of real options, on which our further analysis can be based. On the one hand, compound options, i.e. options on options, which occur if a follow-up project again serves as basis for another follow-up project, can be considered. On the other hand, the impacts of either deferral options or abandonment options on the risk-return position of IT projects deserve a more detailed analysis.

Appendix:

Due to space restrictions we provide the appendix on http://www.uibk.ac.at/iwi2/

References:

Due to space restrictions we provide the list of references on http://www.uibk.ac.at/iwi2/
Stock Market Reaction to Information Technology Investments: Towards an Explanatory Model

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Journal:</strong></td>
<td>17th European Conference on Information Systems</td>
</tr>
<tr>
<td><strong>Manuscript ID:</strong></td>
<td>ECIS2009-0243.R1</td>
</tr>
<tr>
<td><strong>Submission Type:</strong></td>
<td>Research Paper</td>
</tr>
<tr>
<td><strong>Keyword:</strong></td>
<td>Business value of IT, Economics of IS, IS Investments, Theory Building</td>
</tr>
</tbody>
</table>
STOCK MARKET REACTION TO INFORMATION TECHNOLOGY INVESTMENTS: TOWARDS AN EXPLANATORY MODEL

Roztocki, Narcyz, School of Business, State University of New York at New Paltz, 75 S. Manheim Blvd., New Paltz, NY 12561-2443, USA, roztockn@newpaltz.edu

Weistroffer, Heinz Roland, School of Business, Virginia Commonwealth University, 1015 Floyd Avenue, Richmond, VA 23284-4000, USA, hrweistr@vcu.edu

Abstract

Investments in information technology (IT) do not always result in the expected tangible payoffs, and the factors which influence the effect of IT investments on organizational performance are not well understood. Stock market reaction is one approach to appraising IT investments. In this paper we propose a conceptual model describing the factors that impact IT investments based on market reaction findings of major event studies on IT implementation announcements. This preliminary model may serve as a starting point to better understand the complex issue of stock movements related to IT investments.

1 INTRODUCTION

In light of the apparent importance of IT to organizational performance, combined with the lack of clarity as to what makes IT investments successful (Oz, 2005), it is not surprising that much research has been dedicated to identifying and understanding the factors that lead to improved payoffs from IT. IT productivity has implications not only for the firm and its stakeholders, but also for the economy at large. Much current research is dedicated to the impact of IT on global competitiveness and development, particularly in emerging and developing countries (Roztocki & Weistroffer, 2008b).

In addition to the more traditional approaches to investigate IT productivity, such as case studies, surveys, and research databases, event studies are also increasingly being used (Roztocki & Weistroffer, 2008a; Roztocki & Weistroffer, 2009). According to the efficient market theory (Fama, 1970; Fama, 1991), which provides the foundation for the event studies methodology, all available information to investors is reflected in the stock prices. When unanticipated news reach the financial markets, investors assess their relevance and potential effects on particular firms, industries, and economic regions. Stock prices of a company will move up when the news are perceived to be favorable, and bad news, i.e. news indicating the possibility of diminishing future cash flows for a company, will result in decreasing stock prices. In the event studies approach, the reaction of stock markets to reported events is used to explore the perceived relevance and implications of these events (McWilliams & Siegel, 1997).

Event studies are now widely-used in business research areas such as accounting, finance, and strategic management. More recently event studies have become quite popular in IT research to identify specific factors which impact the outcomes of IT investments. Most of these studies, however, have taken an oversimplified approach, looking at particular factors in isolation – the interaction among the influential factors has hardly been considered at all (Oh et al., 2006). Frequently, this simplistic approach may lead to a perceived absence of market reaction to IT investments in a large sample of announcements, as the interaction of various factors can have a nullifying effect. Despite the fact that interaction among factors seems apparent and has been observed by several authors (Hayes et al., 2001; Oh et al., 2006), according to our knowledge, no attempts have been made to construct a model to explain the interaction among multiple factors.
In this paper we try to close or at least narrow this gap and propose a model which encompasses a variety of factors and possible interactions between these factors. The proposed model is based on a systematic literature review and backed by experiences from conducting our own event studies.

2 METHODOLOGY AND DATA COLLECTION

Our conceptual model is constructed from a systematic review of previous event studies investigating stock price reactions to IT investments. To identify appropriate published studies, we checked several literature databases, such as ABI/Inform (Proquest), Business Source Premier (EBSCO), JASTOR, and Science Direct. To find papers relevant to our proposed model, search queries including keywords such as “event study”, “stock market reaction”, “market value”, “announcement”, “information technology” and “information system” were used. This literature review was first conducted in spring 2007, with several up-dates, the last one in fall 2008.

The selection of research papers describing relevant event studies was topic driven and not particularly focused on specific publication outlets. However, to be included for the construction of our model, an article needed to satisfy the following three criteria: First, the article needed to be published in an academic, peer-reviewed journal, or in refereed proceedings of a major IT conference. Second, the article needed to use an event studies approach as the primary research method. Third, the topic of investigation presented in the paper had to be some kind of investment in IT, where investment in IT is defined as any large, non-routine expense for implementing new technology or aimed at making better use out of existing technology. This definition basically follows that proposed by Bacon (1992), though we broaden it slightly beyond hardware and software to also include human resources. Thus, implementing a new enterprise information system qualifies, but so does creating a new executive position dedicated to the administration and management of existing technology.

After assembling the appropriate papers, we conducted a systematic review of the reported studies, following the steps suggested by Rosenthal and DiMatteo (2001). Consequently, each paper was examined for independent, dependent, and control variables, and the reported results across the different studies were compared and analyzed for possible explanations when the outcomes differed or seemed to contradict each other.

Overall, our final sample includes at total of twenty-three studies.

3 REVIEW OF EVENT STUDIES ON IT INVESTMENTS

3.1 General Overview of the Event Studies

Twenty-two of the twenty-three papers included in our sample were published in the time period 2001-2008; one paper was published in 1993. The twenty-three papers in our sample represent a combined authorship of forty-eight authors, most (thirty-four) affiliated with US universities; the remaining fourteen authors were from Australia (three), Canada (two), China (two), Netherlands (three), South Korea (one) and Taiwan (three). The main focus of the investigations was stocks of US based companies.

The majority (nine) of the twenty-three papers investigated the stock market reaction to IT investments in general, while most of the others focused on investment in a particular technology, such as e-commerce initiatives (six), enterprise resource planning (ERP) systems (two), enterprise application integration (EAI) applications (one), supply chain managements systems (one) and customer-related systems (one). One paper examined both integration solutions (ERP and EAI) and two investigated IT related investments in human resources, specifically the creation of a CIO position. Table 1 summarizes the event studies in chronological order.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Type of IT Investments</th>
<th>Specific Factors Investigated</th>
</tr>
</thead>
</table>
| Dos Santos et al. (1993) | general IT investments 97 announcements for 1981-1988 | - manufacturing vs. financial industry  
- innovative vs. non-innovative IT |
- IT driven vs. not IT driven industry  
- external vs. internal CIO hires |
| Hayes et al. (2001) | ERP implementation 91 announcements for 1990-1998 | - small vs. large company  
- financially healthy vs. unhealthy  
- large vs. small vendor |
| Im et al. (2001) | general IT investments 238 announcements for 1981-1996 | - manufacturing vs. financial industry  
- small vs. large company  
- B2B vs. B2C  
- tangible vs. digital goods |
| Chatterjee et al. (2002) | general IT investments 112 announcements for 1992-1995 | - IT infrastructure vs. IT applications |
| Geyskens et al. (2002) | e-commerce investments 93 announcements | - channel power and experience  
- time of entry |
| Dehning et al. (2003) | general IT investments 353 announcements for 1981-1996 | - transformational vs. non-transformational  
- industries with substantial structural changes vs. industries with modest changes  
- leaders vs. laggards |
| Hunter (2003) | general IT investments 150 announcements for 1990-1997 | - exploitative vs. exploratory |
| Filbeck et al. (2005) | supply chain management IT 247 announcements for 1995-2000 | - general only |
| Dardan et al. (2006) | customer-related IT investments 57 announcements for 1996-2001 | - general only |
| Oh et al. (2006) | general IT investments 340 announcements for 1981-1999 | - high growth vs. slow growth firm  
- strategic vs. non-strategic investment  
- asset specific vs. non-asset specific |
| Ranganathan and Brown (2006) | ERP implementations 116 announcements for 1997-2001 | - full suites vs. small number of modules  
- leading vs. non leading ERP vendors |
| Roztocki and Weistroffer (2006) | IT investments by companies using ABC 81 announcements | - automate vs. transform |
- tangible vs. digital goods and services  
- B2B vs. B2C |
| Khallaf and Skantz (2007) | CIO position 443 announcements for 1987-2002 | - CIO hires: new vs. existing position |
| Lin et al. (2007) | e-commerce investments 179 announcements for 1999-2002 | - large vs. small firms  
- early entrants vs. late entrants |
| Meng and Lee (2007) | general IT investments 128 announcements for 1999-2002 | - company location in China vs. USA |
- companies with low beta vs. high beta  
- bull vs. bear market condition |
| Roztocki and Weistroffer (2007b) | general IT investments 179 announcements for 1989-2005 | - companies using ABC vs. not using ABC  
- companies with low beta vs. high beta  
- bull vs. bear market conditions |
- companies with low beta vs. high beta  
- bull vs. bear market conditions |

Table 1. Summary of Event Studies Reviewed
3.2 Variables Used in the Studies

Following the meta-analysis approach of Rosenthal and DiMatteo (2001) we examined the independent and dependent variables in the studies of our sample. Almost all studies used changes in stock price as dependent variable. One study used changes of systematic and unsystematic company risk as consequence of investments in IT (Dewan & Ren, 2007). Additionally, a few studies used changes in trading volume as dependent variable. We also identified the control variables reported in the studies. Overall, the independent and control variables used in the studies fall into five major categories: company characteristics, type of IT investments, vendor characteristics, economic conditions, and characteristics of the announcements. Tables 2 and 3 summarize the independent and control variables.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Studies</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Company Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry</td>
<td>(Dos Santos et al., 1993)</td>
<td>1: manufacturing; 0: otherwise</td>
</tr>
<tr>
<td></td>
<td>(Im et al., 2001)</td>
<td>finance versus non-finance</td>
</tr>
<tr>
<td></td>
<td>(Chatterjee et al., 2001)</td>
<td>1: high level of IT driven transformation; 0: otherwise</td>
</tr>
<tr>
<td></td>
<td>(Dehning et al., 2003)</td>
<td>1: financial (SIC code 6000-6299); 0: otherwise</td>
</tr>
<tr>
<td></td>
<td>(Meng &amp; Lee, 2007)</td>
<td>manufacturing versus finance</td>
</tr>
<tr>
<td></td>
<td>(Meng &amp; Lee, 2007)</td>
<td>IT-using versus IT-producing companies</td>
</tr>
<tr>
<td>Company Size</td>
<td>(Hayes et al., 2001)</td>
<td>small and large companies, based on total assets</td>
</tr>
<tr>
<td></td>
<td>(Im et al., 2001)</td>
<td>small, middle, and large companies, based on total assets</td>
</tr>
<tr>
<td></td>
<td>(Dehning et al., 2003)</td>
<td>estimated by total assets</td>
</tr>
<tr>
<td></td>
<td>(Meng &amp; Lee, 2007)</td>
<td>estimated by assets</td>
</tr>
<tr>
<td>Company Location</td>
<td>(Meng &amp; Lee, 2007)</td>
<td>companies located in China versus companies located in the USA</td>
</tr>
<tr>
<td>Financial Health</td>
<td>(Hayes et al., 2001)</td>
<td>estimated by Altman’s Z score</td>
</tr>
<tr>
<td>Company Growth</td>
<td>(Oh et al., 2006)</td>
<td>estimated by market-to-book ratio</td>
</tr>
<tr>
<td>Systematic Company Risk</td>
<td>(Roztocki &amp; Weistroffer, 2007a; Roztocki &amp; Weistroffer, 2007b; Roztocki &amp; Weistroffer, 2008c)</td>
<td>estimated by beta factor</td>
</tr>
<tr>
<td>Variance of Daily Stock Returns</td>
<td>(Oh et al., 2006)</td>
<td></td>
</tr>
<tr>
<td>Costing Systems</td>
<td>(Roztocki &amp; Weistroffer, 2007b)</td>
<td>companies using activity-based costing versus not using activity-based costing</td>
</tr>
<tr>
<td><strong>Type of IT Investments</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Innovation Content</td>
<td>(Dos Santos et al., 1993)</td>
<td>two dummy variables: 1: innovative investment; 0: otherwise</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1: non-innovative investment; 0: otherwise</td>
</tr>
<tr>
<td></td>
<td>(Ferguson et al., 2005)</td>
<td>innovative and non-innovative e-commerce initiatives</td>
</tr>
<tr>
<td>Benefit to IT Infrastructure</td>
<td>(Chatterjee et al., 2002)</td>
<td>1: infrastructure investment; 0: otherwise</td>
</tr>
<tr>
<td>CIO Position</td>
<td>(Chatterjee et al., 2001)</td>
<td>1: external candidate; 0: otherwise</td>
</tr>
<tr>
<td>Strategic Role</td>
<td>(Oh et al., 2006)</td>
<td>1: transformative investments; 0: otherwise</td>
</tr>
<tr>
<td>Tangible/Digital Goods</td>
<td>(Dewan &amp; Ren, 2007)</td>
<td>1: for tangible goods electronic commerce initiatives; 0: for digital goods or services initiatives</td>
</tr>
<tr>
<td></td>
<td>(Dewan &amp; Ren, 2007)</td>
<td>1: for a B2B type of electronic commerce initiative; 0: for B2C</td>
</tr>
<tr>
<td>Asset Specificity</td>
<td>(Oh et al., 2006)</td>
<td>1: specific purpose investment; 0: high flexibility investment</td>
</tr>
<tr>
<td></td>
<td>(Dehning et al., 2003)</td>
<td>number of days passed from the first announcement</td>
</tr>
<tr>
<td></td>
<td>(Ferguson et al., 2005)</td>
<td>pre-10 March 2000 versus post-10 March 2000 electronic commerce initiatives</td>
</tr>
<tr>
<td></td>
<td>(Dewan &amp; Ren, 2007)</td>
<td>1: 1998 or 2000; 0: 1996 or 2002</td>
</tr>
<tr>
<td>Variables</td>
<td>Studies</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>Scope of Investment</td>
<td>(Dewan &amp; Ren, 2007)</td>
<td>1: new electronic commerce initiative; 0: expansion of existing e-commerce capability</td>
</tr>
<tr>
<td>Vendors Characteristics</td>
<td>ERP Vendor Size</td>
<td>(Hayes et al., 2001)</td>
</tr>
<tr>
<td></td>
<td>ERP Vendor Leadership</td>
<td>(Ranganathan &amp; Brown, 2006)</td>
</tr>
<tr>
<td>Economics Conditions</td>
<td>Stock Market Conditions</td>
<td>(Roztocki &amp; Weistroffer, 2007a; Roztocki &amp; Weistroffer, 2008c)</td>
</tr>
</tbody>
</table>

**Table 2. List of Independent Variables**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Studies</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company Characteristics</td>
<td>Industry</td>
<td>(Oh et al., 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Chatterjee et al., 2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Ranganathan &amp; Brown, 2006)</td>
</tr>
<tr>
<td></td>
<td>Company Size</td>
<td>(Chatterjee et al., 2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Hunter, 2003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Oh et al., 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Ranganathan &amp; Brown, 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Dewan &amp; Ren, 2007)</td>
</tr>
<tr>
<td></td>
<td>Company Growth</td>
<td>(Chatterjee et al., 2002)</td>
</tr>
<tr>
<td></td>
<td>Organizational Slack</td>
<td>(Hunter, 2003)</td>
</tr>
<tr>
<td></td>
<td>Firm Performance</td>
<td>(Dewan &amp; Ren, 2007)</td>
</tr>
<tr>
<td></td>
<td>Systematic Company Risk</td>
<td>(Dewan &amp; Ren, 2007)</td>
</tr>
<tr>
<td></td>
<td>Unsystematic Company Risk</td>
<td>(Dewan &amp; Ren, 2007)</td>
</tr>
<tr>
<td>Type of IT Investment</td>
<td>Time of Announcement</td>
<td>(Hunter, 2003)</td>
</tr>
<tr>
<td>Announcement Characteristics</td>
<td>Source of Announcement</td>
<td>(Oh et al., 2006)</td>
</tr>
</tbody>
</table>

**Table 3. List of Control Variables**

### 3.3 Comparison of Reported Stock Market Reactions

Typically, the stock market reaction to IT investments is assessed by calculating abnormal returns (AR) around the time of the announcement (Roztocki & Weistroffer, 2008c). The day of the announcement is denoted as day 0. The period around the announcement day is called event window and may include several days before and/or after. The most commonly used event window includes one day before (day -1), the day of announcement (day 0), and the day after the announcement (day 1). The AR for the days included in the event window are added up to get the cumulative abnormal returns (CAR). Frequently, the AR are standardized and the stock market reaction for particular event windows is measured by cumulative standardized abnormal returns (CSAR). (The formula for
calculating the CSAR can be found in, for example, Roztocki & Weistroffer (2008c). Table 4 shows the stock market reactions reported by the studies included in our sample.

<table>
<thead>
<tr>
<th>Reaction (%)</th>
<th>Measured by</th>
<th>Sample Size</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General IT Investments</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.09</td>
<td>CSAR (-1,0)</td>
<td>97</td>
<td>(Dos Santos et al., 1993)</td>
</tr>
<tr>
<td>0.02</td>
<td>CSAR (-1,0)</td>
<td>238</td>
<td>(Im et al., 2001)</td>
</tr>
<tr>
<td>1.224</td>
<td>CAR(-1,1)</td>
<td>112</td>
<td>(Chatterjee et al., 2002)</td>
</tr>
<tr>
<td>-0.85</td>
<td>CAR(-1,1)</td>
<td>150</td>
<td>(Hunter, 2003)</td>
</tr>
<tr>
<td>0.35</td>
<td>CAR(-1,1)</td>
<td>340</td>
<td>(Oh et al., 2006)</td>
</tr>
<tr>
<td>0.0037</td>
<td>CAR(0,2)</td>
<td>63</td>
<td>(Meng &amp; Lee, 2007)</td>
</tr>
<tr>
<td>1.0778</td>
<td>CAR(0,2)</td>
<td>65</td>
<td>(Meng &amp; Lee, 2007)</td>
</tr>
<tr>
<td>-0.09</td>
<td>CSAR (-1,1)</td>
<td>179</td>
<td>(Roztocki &amp; Weistroffer, 2007b)</td>
</tr>
<tr>
<td><strong>Specific Types of Investments</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Innovative</td>
<td>1.03</td>
<td>CSAR (-1,0)</td>
<td>25</td>
</tr>
<tr>
<td>Not Innovative</td>
<td>-0.09</td>
<td>CSAR (-1,0)</td>
<td>43</td>
</tr>
<tr>
<td>CIO Position</td>
<td>1.16</td>
<td>CAR(-1,1)</td>
<td>96</td>
</tr>
<tr>
<td>Customer Related</td>
<td>0.366</td>
<td>CAR(-1,1)</td>
<td>57</td>
</tr>
<tr>
<td>EAI</td>
<td>-0.084</td>
<td>CSAR(-1,1)</td>
<td>81</td>
</tr>
<tr>
<td>Electronic Commerce</td>
<td>4.2</td>
<td>CAR (-1,1)</td>
<td>251</td>
</tr>
<tr>
<td></td>
<td>-1.9</td>
<td>CAR (-1,1)</td>
<td>542</td>
</tr>
<tr>
<td></td>
<td>0.48</td>
<td>CAR (-1,1)</td>
<td>232</td>
</tr>
<tr>
<td>ERP</td>
<td>0.6</td>
<td>CAR(0,1)</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>1.49</td>
<td>CAR (-1,1)</td>
<td>116</td>
</tr>
<tr>
<td></td>
<td>-0.113</td>
<td>CSAR(-1,1)</td>
<td>48</td>
</tr>
<tr>
<td><strong>Specific Company Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manufacturing</td>
<td>0.40</td>
<td>CSAR (-1,0)</td>
<td>33</td>
</tr>
<tr>
<td>Finance</td>
<td>-0.08</td>
<td>CSAR (-1,0)</td>
<td>64</td>
</tr>
<tr>
<td>No-Finance</td>
<td>-0.03</td>
<td>CSAR (-1,0)</td>
<td>115</td>
</tr>
<tr>
<td>Companies Using ABC</td>
<td>0.066</td>
<td>CSAR (-1,0)</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>-0.097</td>
<td>CSAR(-1,1)</td>
<td>81</td>
</tr>
</tbody>
</table>

Table 4. Stock Market Reaction

4 SUMMARY OF FINDINGS

A comparison across different studies suggests that there are a large number of factors, which may influence stock market reaction with varying contributions. Company size and timing of the investments appear to be highly influential, whereas industry type seems to have a more moderate effect. As stated earlier, the variables examined in the reviewed event studies can be grouped into five major categories: company characteristics, type of IT investments, vendor characteristics, economic conditions, and announcement characteristics. The market reactions observed in the twenty-three studies reviewed are summarized in Table 5.
<table>
<thead>
<tr>
<th>Category</th>
<th>Factor</th>
<th>Stock Reaction</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Company Characteristics</strong></td>
<td>Industry type</td>
<td>Indifferent</td>
<td>Industry effects tend to be insignificant</td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td>Negative</td>
<td>The magnitude of stock market reactions seems to be negatively related to the company’s size; larger for small companies; smaller for large companies. Stock reactions seem to be more positive for smaller companies</td>
</tr>
<tr>
<td></td>
<td>Financial health</td>
<td>Indifferent</td>
<td>For ERP investments, when the financial health of a company worsens, the stock price reaction seems to become more positive for large firms and become more negative for smaller firms</td>
</tr>
<tr>
<td></td>
<td>Costing system used</td>
<td>Indifferent</td>
<td>ABC implementation seem to benefit investments to automate business processes</td>
</tr>
<tr>
<td></td>
<td>Industry leadership</td>
<td>Indifferent</td>
<td>Stock market seems to react more positively when the announcing company that lead its industry sector, but this reaction is statistically insignificant</td>
</tr>
<tr>
<td><strong>IT Investment Type</strong></td>
<td>Innovation content</td>
<td>Positive</td>
<td>Highly innovative announcement are better received</td>
</tr>
<tr>
<td></td>
<td>Category</td>
<td>Indifferent</td>
<td>ERP implementation seem to be received positively Infrastructure investments also positively There seems to be large variation in stock respond to e-commerce initiatives.</td>
</tr>
<tr>
<td></td>
<td>Strategic importance</td>
<td>Positive</td>
<td>Transformative IT investments seem to be received more positively</td>
</tr>
<tr>
<td></td>
<td>Focus</td>
<td>Indifferent</td>
<td>Asset specific IT investments seem to result in insignificantly negative stock reaction</td>
</tr>
<tr>
<td><strong>Vendor Characteristics</strong></td>
<td>Size</td>
<td>Positive</td>
<td>Large, leading vendors with established reputation seem to benefit investment in IT</td>
</tr>
<tr>
<td><strong>Economic Conditions</strong></td>
<td>Bear/Bull stock market</td>
<td>Bull: Positive</td>
<td>Conditions of the stock market seem to be important factor</td>
</tr>
<tr>
<td></td>
<td>Bear</td>
<td>Bear: Negative</td>
<td></td>
</tr>
<tr>
<td><strong>Announcement Characteristics</strong></td>
<td>Source of announcements</td>
<td></td>
<td>Stocks seem to respond more favorably when the announcement is released by the investing company as opposed to the vendor.</td>
</tr>
</tbody>
</table>

Table 5. Observed Market Reaction

4.1 Company Characteristics

Several event studies looked at a number of different company characteristics as potentially influential factors to explain stock market reaction. The industry of the investing company was most often used as a variable, with the idea that companies in certain industries, such as finance, will benefit more from IT investments. This assumption was derived from the fact that banking is an information intensive industry. Most event studies, however, fail to provide evidence that industry is an influential factor.

Firm size effect was also examined by several studies. Overall, it appears that the magnitude of stock reactions diminish with company size. This makes sense, as the same size investment will have larger impact on a small firm than a large firm, i.e. the size of the investment relative to the size of the company or its capital assets is important.

4.2 Type of IT Investments

Type of IT investments investigated includes those with innovative content, transformative IT investments (as opposed to investments for operational efficiency only), investments in specific types of IT such as ERP or e-commerce, and asset specific focused investments. Only innovative IT investments and investments of strategic significance (transformative IT investments) resulted in positive stock market reactions. Innovative and transformative IT investments may result in competitive advantages for the investing company, and seem to be rewarded by the stock market.
4.3 Vendor Characteristics

Large, established vendors seem to instill trust, and investments in IT from large vendors are more likely to result in positive market reactions. In other words, it seems that stock market investors believe that large vendors are more likely to possess the technical expertise and resources to make IT investments successful.

4.4 Economic Conditions

Investments in IT in times of bull market conditions are more likely to result in positive stock market reactions than investments during bear market conditions. It appears that stock market investors are more doubtful about investments in IT and their effects on financial performance during bear markets.

4.5 Announcement Characteristics

Relatively few event studies looked at the characteristics of the announcement itself. However, it may reasonably be expected that the way the investments are communicated to the investors is of some relevance. The study by Oh et al. (2006) compared the stock market reaction to announcements made by the investing companies and announcements made by the vendors, and found a significant difference in the reaction. Financial markets appear to respond more positively to announcements made by the investing company. Interestingly, a large number of companies (approximately 60 percent) as reported by Oh et al. (2006) do seem to prefer that the announcements are made by vendors or service providers.

In addition to the source of announcements, there may also be influential factors related to the wording used in the announcements themselves. For example, investors may interpret specific language used in the announcements as an indication of presence or absence of clear objectives, technical competence, or support by management. No published studies have investigated this aspect, to the knowledge of the authors.

5 PROPOSED MODEL

The results from the literature review also suggest that there may be complex interactions between the factors that impact abnormal stock price returns. For example, the study by Hayes et al. (2001) implies that there is an interaction between the financial health and the size of a company. The most positive reactions to ERP investments were observed for small, financially healthy companies. As a company’s size increases while the financial health remains strong, the magnitude of the stock market reaction diminishes. When the financial health of a company worsens, the stock price reaction seems to become more positive for large firms and become more negative for smaller firms.

While the study by Hayes et al. (2001) looked at the interaction of two factors related to company characteristics, a comparison across different studies provides evidence of more complex interaction between factors. For example, while most studies were unable to find positive stock price reactions to non-innovative, automate IT investments (Dos Santos et al., 1993), Roztocki and Weistroffer (2006) reported positive reaction to automate investments when a company is using activity-based costing (ABC).

5.1 General Model

Overall, the reviewed studies confirm that there are a large number of influential factors that may affect investments in IT, and that these factors seem to be subject to complex interactions. As stated earlier, the identified factors can be categorized into five major groups: company characteristics, IT investment type, vendor characteristics, economic conditions, and announcement characteristics, thus suggesting the general model as depicted in Figure 1.
5.2 Application of the model

For the purpose of illustrating the use of the model and complexity of multiple factors interacting, we look at innovation content for a possible investment type. For vendor characteristics we look at size and for economic conditions we use the conditions of the stock market (i.e. bear or bull market), as shown in Figure 2.

Regarding the innovation content of IT investments, the stock prices seem to react more favorably when the investments are innovative and move the company ahead of its competitors. Therefore we show a positive effect for IT investments for innovation. Regarding vendor characteristics, a positive stock market reaction is more likely for large established vendors than for smaller and less established ones. Therefore again, we show a positive effect in our model. Regarding economic conditions, a bull market seems to benefit a positive stock market reaction. While declining stock markets, i.e. bear
markets, are more likely to result in less favorable stock price reactions. Again these are shown as positive and negative effects respectively in our model.

Our model, at least to some extent, explains the inconclusive findings of previous productivity studies (Oz, 2005). It is for example possible, for exactly the same type of IT investment to obtain significantly positive or negative stock price reactions when not controlling for other influential factors. A significantly positive stock price reaction is likely for a sample including mostly smaller companies buying from large established leading vendors during bull markets. For the same type of IT investment, the stock market reaction could be significantly negative when the companies in the sample are buying from small, non-leading vendors during bear market conditions. In both situations, investments from smaller companies seem to result in heftier stock market reactions.

5.3 Possible Business Implications

The model and the observed effects as shown in Figure 2 may have some important implications for businesses that are considering new investments in IT. Companies, particularly large companies, should not expect a positive reaction by the stock market to new IT investments, unless it is clearly communicated that the IT investment will likely result in innovation and provide strategic value to the investing firm. Furthermore, IT investments realized through small, less established vendors, are perceived as particularly risky and likely to result in negative market reactions. Using large, established vendors seems to be a safer way to go, if negative market reactions are to be avoided. If possible, IT investments are better done during bull market conditions; thus if market conditions are unfavourable, it may be wise to hold off with any non-critical new investments. Finally, the investing company should take the initiative in communicating the news to its shareholder and make the initial announcement of the new IT investment itself, rather than let the vendor make the announcement.

6 CONCLUSIONS

6.1 Contribution

We believe that our systematic review of event studies related to IT investments and our model presented in this paper make a substantial contribution to the body of knowledge in that this is perhaps the first meta analysis of event studies in the field of IT investments and the first attempt to construct a model to explain the impact and interactions of various factors on market reaction to IT investments. This model and the results of our review of twenty-three papers should greatly benefit other scholars, as it may serve as a foundation for further research on IT productivity. Future research building on our model need not be limited to event studies, as the compiled list of potentially influential factors may impact other measures of IT productivity, besides stock market reaction, and thus serve as a foundation for other types of research as well. Furthermore, as mentioned earlier, our model may help explain some of the inconclusiveness and inconsistency in the results of earlier IT productivity studies.

We believe that our model is also useful to business executives, as stock performance is often seen by stakeholders as a crucial indicator of firm performance. It is difficult for executives to ignore unfavorable stock movements of their companies, as such movements may lead to loss of confidence by employees, customers, suppliers, creditors, etc. In this regard, our model may help managers better understand favorable or unfavorable conditions for making IT investments.

Moreover, since the ideas presented in this paper are based on a comprehensive literature review, the proposed model may also lead to increased and improved usage of event studies in IT research.

6.2 Limitations

Although event studies in the field of IT have become more common, as compared to other disciplines, the absolute number of such studies is still small, which constitutes our first limitation.
Basically, our conclusions are drawn from event studies published in twenty-three academic papers. It is likely that as more studies are conducted and published, more influential factors will emerge.

A second limitation is related to the methods of estimating the stock market reaction. There is variation in estimation periods and event windows used by the different studies. Therefore the comparison of the findings from different studies may be limited. In addition, there is always a chance that the particular data set is contaminated. For example, Dehning et al. (2003) reported finding two outliers in their earlier event studies. Some outliers may go undetected and lead to faulty conclusions.

A further limitation of this study is that our model was constructed mostly based on findings derived from US companies and stock data. It is possible that for international companies some factors could vary in importance and the model would need further refinement to accommodate country characteristic.

6.3 Future Research

The results presented in this paper are not final but provide a more complete picture and new ideas for possible research avenues. Overall, it seems that previous event studies in the field of IT call for substantial revalidation. Future research may validate and enhance or improve our model by looking at additional factors that influence market returns. It is quite possible that other economic factors, such as interest rates, inflation level, and exchange rates substantially influence the stock market reaction. Also, with respect to announcement characteristics, there may be influential factors related to the communication of the investments. For example, the wording used in the announcements themselves could impact the investors’ reactions to IT investment announcements. Thus, investors may interpret overly use of some words as a sign of lacking decisiveness, lacking technical competence, or lacking support by management.

7 REFERENCES


THE HEDONIC AND UTILITARIAN VALUE OF DIGITAL GAMES AT PRODUCT CATEGORY LEVEL

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0354.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Computer Games, Evaluation, Brands, Information goods</td>
</tr>
</tbody>
</table>
THE HEDONIC AND UTILITARIAN VALUE OF DIGITAL GAMES AT PRODUCT CATEGORY LEVEL

Storgårds, Jan Henrik, Helsinki School of Economics, POBOX 1210, FI-00101 Helsinki, Finland, jan.storgards@hse.fi

Tuunainen, Virpi Kristiina, Helsinki School of Economics, POBOX 1210, FI-00101 Helsinki, Finland, virpi.tuunainen@hse.fi

Öörni, Anssi, Helsinki School of Economics, POBOX 1210, FI-00101 Helsinki, Finland, anssi.oorni@hse.fi

Abstract

Consumers’ product evaluation, choice, and use are driven by both utilitarian and hedonic considerations. Digital games, that are typically considered to be a homogenous product category, are often claimed to be high on hedonic value. However, there is a multitude of digital games genres available. Differences in their appeal, gameplay, and purpose in practice indicate that digital games vary significantly in their proposed outcome.

In this research paper, we present results from an observational survey study in which digital games were investigated at subcategory level. We found differences in weights consumers place on hedonic and utilitarian value at the product subcategory level, and also observed differences between evaluations of experienced and inexperienced consumers. The results indicate that subcategories act as a more assuring source of information than the general product category of digital games. Furthermore, higher gaming experience may lead to more optimistic evaluation on subcategories.

Rather than grouping digital games into one category and treating them all as hedonic products, they should be discussed at their specific subcategory level, by researchers and practitioners, alike.

Keywords: digital games, hedonic value, utilitarian value, product categories
1 INTRODUCTION

Since the 1980’s, various studies in consumer goods and services (Dhar et al. 2000; Okada 2005); information systems (Davis 1989; Van der Heijden 2004), and digital games (Choi et al. 2004; Hsu et al. 2005) have shown that consumer choice and use of different products and services are driven by both utilitarian (UT) and hedonic (HED) considerations.

Hirschman and Holbrook’s (1982, p. 99) early outlook to hedonistic consumer behavior is that “hedonic consumption designates those facets of consumer behavior that relate to the multisensory, fantasy and emotive aspects of one’s experience with products”. In other words, hedonic or pleasure oriented consumption is expected to be motivated by the desire for sensual pleasure, fantasy and fun (Strahilevitz et al. 1998, p. 436). In contrast, utilitarian or goal oriented consumption is “more cognitively driven, instrumental, and goal oriented and accomplishes a functional or practical task” (Dhar et al. 2000 p. 61; Strahilevitz et al. 1998).

Earlier research has found major differences between the perceived value of utilitarian and hedonic software applications, also in the context of digital games (Davis 1989; Raessens et al. 2005). Digital games are most often assumed to be high on hedonic value (Batra et al. 1990; Chen 2007; Hirschman et al. 1982; Hsu et al. 2005; Voss et al. 2003), and the motives for using hedonic systems, such as digital games, are different than those for utilitarian systems, such as office information systems (Van der Heijden 2004).

Digital game production has grown to be a significant sector of software business (Crandall et al. 2006; Siwek 2007). Even though digital games are often treated as a product category among other, relatively heterogeneous consumer products (e.g. paper clips, beer, blue jeans) Batra et al., (1990), Voss et al., (2003), there are multitudes of different types of games, and the reasons to play them vary greatly: games can be played, for example, for educative purposes, or mainly just for fun. The great differences in the appeal and use of games in practice suggest that instead of studying games as one general category, we should be looking at the subcategories.

In this study, we will analyze the perceived hedonic and utilitarian value of digital games in different game subcategories. We will also explore the differences in these values between experienced and inexperienced players, as expressed in recommendations to others.

The structure of this paper is as follows: In section 2, we present our theoretical background. In section 3 we introduce the hypotheses and research model. We then describe our research method and the empirical study setting in section 4, and present the results of our study in section 5. Section 6 summarizes and concludes the paper.

2 THEORETICAL BACKGROUND

Our study builds on the Motivational theory (Deci 1975) to understand the motivation of players of digital games, and on the theories of consumer behavior (Ajzen et al. 1980, see e.g.; Bettman et al. 1980) to understand the effect of prior knowledge and experience on the perceived value of the games.

2.1 Motivation of Game Playing

One of the objectives of game developers is to optimize game experience by designing elements of gameplay that motivate the player to continue playing without too much anxiety or boredom (Chen 2007). Specifically, in digital games production, this concept of flow experience (Csikszentmihalyi 1975) is widely used to provide outcomes such as enjoyment, pleasure, and

Proceedings ECIS 2009
fun, and to maintain the flow at the desired level. The flow experience is a part of an individual’s motivation to play games and has been defined as “an extremely enjoyable experience, where an individual engages in an on-line game activity with total involvement, enjoyment, control, concentration and intrinsic interest.” (Hsu et al. 2004).

Therefore, the motivational theory by Deci, (1975) lays the basis for the understanding of how digital games are chosen and why they are played. From motivational perspective of consumption, hedonic goods entail intrinsic value, whereas utilitarian entail more extrinsic values. Intrinsic motivation has been defined as “the inherent tendency to seek out novelty and challenges, to extend and exercise one’s capacities, to explore, and to learn it is performing an activity for the satisfaction of the activity itself” (Ryan et al. 2000, p. 70). Extrinsic motivation, in turn, is expected to lead to performance of an activity, in order to attain some separable outcome (Ryan et al. 2000).

Different products and services often vary greatly in their proposed outcomes (Hirschman et al. 1982). For instance, many services intend to provide an outcome closer to hedonistic value (e.g. movies, concerts) rather than utilitarian value provided by many packaged goods (e.g. shoe laces, hammers). Different digital games are similar in their delivery format, but many times distinct in their proposed outcome, making the analysis of the game subcategories necessary.

A category exists “when two or more distinguishable objects or events are treated equivalently” (Mervis et al. 1981, p. 89). This equivalent treatment means different ways of labeling distinct objects or events with the same name, or performing the same action on different objects. Consumers have been found to rely on the categorizing process: Evaluation of a product depends on the particular category to which it is perceived to belong (Blackwell et al. 2006, p. 110). Given this, specific brands can be built around these consumer segments (Rust et al. 2004).

We define digital games as examples of social systems which have information technology embedded in them (Land 1992). In practice, digital games are software applications, the purpose of which is to entertain (Hsu et al. 2004) – or with some games, educate – the users.

There is a multitude of ways to categorize digital games: gameplay, technology platform, delivery channel, age limit, language, graphics, user type, purpose, producer, temporality, price, and character, to name a few (Mäyrä 2008; Rutter et al. 2006). In this study, we classify different subcategories of digital games by their proposed gameplay experience, which has been defined as “a complex dynamics of interaction between the player and a game in which the structure of game including characters, virtual space, rules and story elements are at central focus” (Ermi et al. 2005). Digital game sub-categories, such as, sports games, massively multiplayer online role playing games (MMORPGs), racing games, and so on, are commonly identified segments which can be benchmarked with competitive analysis and product positioning (Rust et al. 2004). Most importantly, the labels of these categories are also those used by the consumers who play the games.

2.2 Experience

The essence of consumer behavior has been described as a choice between different product and service alternatives (Ajzen et al. 1980). This is based on the assumption that behavioral changes related to choice are often dominated by cognitive processes and systematic use of available information, even if people often strive to simplify their decision making (Howard et al. 1969). Nevertheless, consumer decisions are context dependent and subject to, for instance, the influence of product type and category (Zeithaml et al. 2006). Additionally, individual differences drive consumers to manage their deliberative processes differently, depending on many factors and situations (Foxall 2005). Decision making involves many environmental factors that lie outside the control of the individual. Foxall (2005) maintains that social, business, cultural and
Economical factors affect the consumers’ stimuli and attention. When information is received, it is recorded either on the short or long term memory, and processed depending on the consumer’s prior experiences, beliefs, attitudes, goals and other evaluation criteria.

Consumers’ choice criteria are influenced by prior knowledge and experience (Bettman et al. 1980). People with little prior knowledge and experience tend to simplify their product evaluation process and decision making. While they acknowledge the benefits of additional product information, the perceived high cost of information processing discourages search for and processing of more information. In contrast, people with high levels of prior knowledge face low search costs; yet, they tend to shortcut the search process, as they rely on previously acquired information. People with some prior knowledge have both the ability and motivation to process new information available to them. Prior experience shapes the decision process through other heuristic effects, as well. For example, consumers with high levels of experience tend to engage in brand comparisons, while less experienced consumers rely more on product attribute information (Bettman and Park 1980). Most importantly, an experienced user has different, typically higher, enjoyment related expectations than inexperienced user (Atkinson et al. 1997).

Experienced consumers are expected to be more confident sources of recommendations than inexperienced consumers. Word-of-mouth (WOM) recommendation has been depicted to be an effective method to influence consumers in their choice process. According to Brown and Reingen (1987), WOM—type of recommendation can be divided into two distinct sources. Firstly, strong tie sources are those that are socially relevant to the consumer and known personally (e.g. friends and family). Strong ties have been shown to be important at the micro level of referral behavior. Secondly, weak tie secondary sources are those seldom contacted acquaintances, or those not known personally at all, that have been found to play a crucial role in the flow of WOM information across groups (Brown et al. 1987 p. 360). Most importantly, recent research evidence suggests that a simple response to a question “How likely is it that you would recommend this product to a friend or colleague?” would actually reveal how loyal a consumer is to a specific product or a brand (Reichheld 2006). Even though our focus is on the strong tie elements, we acknowledge that the weak tie effects of larger social communities have a great importance in individual’s behavior (Granovetter 1973).

In essence, digital games are experience goods, the quality of which can be determined only through consumption (Bryce et al. 2006; Zeithaml et al. 2006). Information in different forms (e.g. demonstration versions) and from different sources (e.g. reviews on websites and WOM) helps the consumers in obtaining critical pre-purchase product information (Klein 1998). Intentional or not, these different sources of information act as recommendations which influence consumers’ product evaluation process positively or negatively (Smith et al. 2005).

3 PROPOSED RESEARCH MODEL AND HYPOTHESES

In this study, we propose that different digital games vary by their perceived hedonic (HED) and utilitarian (UT) value. Following that, we propose that HED is a better predictor of recommendation than UT.

We formulate our hypotheses as followed:

Hypothesis 1 (H1) = Consumer’s prior experience significantly influences his/her perceived level of both HED and UT value of digital games at the product subcategory level.

We report this by creating a scatter plot of the summated variables and comparing statistical differences between single, summated, and latent factor variables.
Hypothesis 2 (H2) = Consumer’s perceived HED value of digital games predicts better strong-tie recommendation than the UT value.

We measure the effect of HED and UT latent factor variables on recommendation with a multiple linear regression model: Recommendation = constant + HED + UT. Specifically, we are interested in the proportion that HED and UT explain recommendation (see Figure 2.).

![Figure 2. Research model.](image)

4 RESEARCH METHOD

According to Voss et al. (2003), the hedonistic and utilitarian constructs can be reliably observed by using five variables in both latent constructs. We adapted these variables for our survey questionnaire (see Table 1). The questionnaire item labels were translated from English to Finnish. Due to the translation issues (e.g., synonyms and overlapping terms), only four of the suggested five terms were used. Due to these issues, two attributes (Enjoyable and Useful) could not be measured as variables within their respective original constructs as they are used to define HED and UT constructs. Thus, they were observed separately as two HED/UT (single) variables. We used semantic differential scaling from -3 to +3 in questionnaire items, however, the final results were transformed to scale from 1 to 7 for easier comparability with prior research.

<table>
<thead>
<tr>
<th>Hedonic variables</th>
<th>Utilitarian variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not enjoyable</td>
<td>Enjoyable</td>
</tr>
<tr>
<td>Dull</td>
<td>Exciting</td>
</tr>
<tr>
<td>Not delightful</td>
<td>Delightful</td>
</tr>
<tr>
<td>Not thrilling</td>
<td>Thrilling</td>
</tr>
<tr>
<td>Not fun</td>
<td>Fun</td>
</tr>
</tbody>
</table>

Table 1. Utilitarian and hedonic variables used in questionnaire (adapted from Voss et al., 2003)

The questionnaire form was first commented and pre-tested by five colleagues and pilot users. A web server based application called Webropol (webropol.com) was used to create and conduct the survey. The respondents were students in a Finnish Business School, taking a course on “Personal Computing Skills” in September, 2008. The empirical set of data was processed using the SAS Enterprise Guide, version 4.1.

During the first two actual data collection sessions out of five in total, 71 respondents were asked to categorize the adjectives used in the questionnaire by their perceived meaning as utilitarian or
hedonic words. All used adjectives were correctly grouped under their respective, expected constructs.

After answering questions measuring background information, the subjects assessed digital games without any reference to any specific game brand or subcategory. For general digital games category, respondents were asked to answer to a question in which the level of HED (enjoyable) and UT (useful) was measured by using them as opposite terms.

Thereafter, respondents were asked to answer questions on 16 different digital game product categories (see Appendix 2). Users were asked to evaluate different categories such as, sports games or massively multiplayer on-line role playing games as a whole. Each questionnaire page with a product category started always with two to nine real digital game package cover pictures. The objective was to create better understanding of real-life products related to the evaluation of images rather than only by using text. All selected games were relatively well known and widely spread, and mostly published for consoles, handheld consoles and PC platforms.

After assessing the psychometric values for game categories, the respondents were asked about their experience during past twelve months in each category. The order of the variables was randomized, but they were in the same order at each product subcategory level. We deemed it unlikely that all respondents had either awareness or experience on every category. For better reliability, each psychometric questionnaire item also included an option to respond “I can not say”.

5 RESULTS

5.1 Descriptive statistics

There were 135 usable responses out of 136 in total (1 uncompleted form). Forty-eight percent of the respondents were female, and 52 % were male. The respondents were between 18-31 years, 20.4 years being the average age and 20 years the median. As many as 44 %of the respondents reported to be active game players, while the remaining 56 % had not played at all in the past 12 months. The average age of starting playing digital games was 7.6 years (Table 2.).

<table>
<thead>
<tr>
<th>Total (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondents (#)</td>
</tr>
<tr>
<td>Men / Women (#, %)</td>
</tr>
<tr>
<td>Age in years (average, median, min-max)</td>
</tr>
<tr>
<td>Age when first time played digital games (years, median, min-max)</td>
</tr>
<tr>
<td>Players vs. Non-players (#, %)</td>
</tr>
<tr>
<td>Average weekly playing time (average, median, range)</td>
</tr>
</tbody>
</table>

Table 2. Descriptive statistics on the respondents

There were only a few missing answers. Generally, with few exceptions, those who reported being experienced game players answered to all questions. There were only two categories which had relatively high amount of “I can not say” responses (professions and text based adventure). In all other categories, the number of "I can not say" -responses ranged from 3 to 28 inexperienced respondents.
5.2 Data analysis

To the general question about digital games HED/UT value level, respondents evaluated digital games to be more fun or entertaining than useful (in a scale of 1-5). Experienced respondents’ (n=59) average was 3.93, while it was 3.61 for the inexperienced (n=76). Furthermore, the difference between experienced and inexperienced respondents was statistically significant (t-value = 2.46, p < 0.05).

We first produced a correlation analysis and calculated the reliability score for both HED and UT variables in general digital game and subcategory levels. All the used variables fit well into their responding latent factor variables (Cronbach’s Alpha > 0.86). Then, we formed a single composite summated measure of both HED and UT by combining their respective variables (Hair et al. 1984).

To test our hypotheses, we first created a summated scatter plot that presents the perceived HED/UT value for each category and between experienced and inexperienced respondents (Figure 3) (for detailed statistical data, see Appendix 2 and 3).

Figure 1. HED / UT summated scatter plot
We then computed factor scores for both HED and UT in all subcategories by using principal component analysis without rotation. These latent variables were used to measure the effect of HED and UT on recommendation with a linear regression model.

The summated HED/UT scatter plot indicates that there is no game subcategory which would be high on only hedonic or on utilitarian value. The experienced respondents seem to be more coherent in their opinions, and there was less variation between the game subcategories for them than for the inexperienced respondents. The inexperienced assess mainly more HED/UT value for playing games generally, but at the product subcategory level their beliefs and attitudes become more negative, decreasing both HED and UT. In contrast, the experienced game players mainly assess the subcategories higher than the digital games in general. The results indicate that higher experience may lead to more optimistic evaluation.

The most notable difference (mean difference => 1.99 - 2.74) between experienced and inexperienced respondents in both HED and UT was in games which war and violence are the focus of the gameplay (FPS as first person shooters, war strategy and action adventure games). The greatest perceived hedonic value among experienced was assessed to party games with a significant difference to inexperienced game players. Singing, dancing and playing together is perceived as a hedonic act also in real-life and those not experienced are probably not interested in these acts in real-life, either.

The least difference in HED and UT was assessed to the general category of digital games and to exergames (games incorporating real physical exercise). Exergames is a new subcategory in which especially Nintendo (Wii) has been very active, promoting the console as a new way of experiencing digital games. It can also be stated, that this category is marketed to inexperienced consumers with a purpose of enlarging the market potential for game industry. These were followed by educative and platform games. Educative games are probably perceived to include beneficial outcomes for their players. In this research setting, the examples of platform games were well known game characters such as Mario and Sonic. The extensive brand building efforts by the brand owners and earlier experiences from the respondents’ childhood could have added to the positive image.

In four categories, there were not enough experienced respondents to make reliable comparisons. These were pet raising (targeted at children), professions (targeted at young girls), text based adventure and MMORPGs.

We then tested the interaction effect of HED/UT latent factor variables to recommendation (RECO) by using multiple linear regression model in all subcategories. Furthermore, the means procedure and t-tests were computed (see Appendix 3). Among the experienced respondents recommendation for all game categories was high (>4.2) as for inexperienced respondents it was relatively low (<3.0). In all subcategories differences were statistically significant (t-test between means, p<0.5). Generally, digital games were recommended based on their hedonic value, which is similar to the finding for the single variable (HED vs. UT) item. Further, the level of adjusted coefficient of determination, r², is notably higher in the subcategory level than in the general category of digital games. This indicates that HED and UT explain better RECO in subcategory level and respondents are more confident about their opinions.

The largest differences between experienced and inexperienced users can be found in games with war and violence (FPS, war strategy, action adventure). The highest recommendation would be given for party games among experienced and for educative and exergames among inexperienced consumers.
Our first hypothesis (H1) was supported. Experience significantly influences the perceived level of both hedonic and utilitarian values of digital games at the product subcategory level. The experienced have constantly higher perceived HED and UT for different subcategories.

For the second hypothesis (H2), we may conclude that even though recommendation can be explained by using hedonistic and utilitarian value, the main interacting variable varies between product categories. There is no general, systematic evidence that only HED would explain recommendation but that it depends on the subcategory. Hence, H2 was rejected.

6 SUMMARY AND CONCLUSIONS

In this study, we set out to investigate the hedonic and utilitarian values of digital games, as perceived by the consumers. Using Business School students as subjects, we conducted a survey to test our hypotheses.

Results of our empirical study demonstrate that, digital games, in all observed subcategories, provide more perceived hedonistic than utilitarian value. However, our analyses suggest that digital games are not only high on hedonic value, but that the level of perceived HED and UT depends on the user's gaming experience and the product subcategory. Hence, digital game evaluation and product positioning should be done at the product subcategory level, rather than generalizing all games being equal in their proposed outcome.

Secondly, we classified the respondents into experienced and inexperienced players. Experience was clearly found to be a differentiating factor for evaluation of the different outcomes of digital games. To accomplish these objectives we used psychometric measurement instruments, specifically hedonic and utilitarian values of information systems as an operational tool.

Statistically, HED and UT variables strongly correlate in all different kinds of data analysis and among different subcategories. This would suggest that perceived hedonic and utilitarian values are not separate constructs, but that, in the context of digital games, they are processed simultaneously in product evaluation situation.

The consumer’s perception about the digital game subcategory affects the beliefs about specific products within it. Digital games subcategories act as a more confident source of information for the consumers than the general category of digital games. Further, the results indicate that higher experience on games within a subcategory may lead to a more optimistic evaluation of HED and UT. This finding has clear implications for the practitioners in the digital game development business who intend to attract new players for their games.

There are some limitations in this study. First, the dualistic perspective in which complex sensory and emotional experiences are measured by using two quantitative constructs such as HED and UT within a survey may be misleading. HED and UT do not fully explain what is the true meaning of these games for the users and why certain game categories are preferred. Therefore, a more in-depth qualitative study by interviewing game players is needed to understand why these differences may occur. Another limitation is the use of students as our subjects. This sample did, however, work well to accomplish our objective of differentiating different types of game categories.

References


Proceedings ECIS 2009


Siwek, S.E. "Video Games in the 21st Century", Entertainment Software Association, p. 36.


**Appendix 2. HED/UT sample size, means, and statistical differences in means.**

<table>
<thead>
<tr>
<th>Category</th>
<th>HED - UT SSD</th>
<th>HED - UT Ex</th>
<th>HED - UT In</th>
<th>HED M</th>
<th>HED F*</th>
<th>HED Si*</th>
<th>HED M</th>
<th>HED F*</th>
<th>HED Si*</th>
<th>UT M</th>
<th>UT F*</th>
<th>UT Si*</th>
<th>UT M</th>
<th>UT F*</th>
<th>UT Si*</th>
<th>MD HED</th>
<th>MD UT</th>
<th>n Ex/In</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital games</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>5.26</td>
<td>4.51</td>
<td>4.73</td>
<td>4.08</td>
<td>0.75</td>
<td>0.65</td>
<td>59/73</td>
<td>59/70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FPS</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>5.44</td>
<td>2.70</td>
<td>4.92</td>
<td>2.50</td>
<td>2.74</td>
<td>2.42</td>
<td>50/75</td>
<td>50/73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>War strategy</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>5.37</td>
<td>2.95</td>
<td>5.14</td>
<td>3.11</td>
<td>2.42</td>
<td>2.03</td>
<td>47/78</td>
<td>47/76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action adventure</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>5.74</td>
<td>3.72</td>
<td>5.28</td>
<td>3.28</td>
<td>2.02</td>
<td>1.99</td>
<td>49/74</td>
<td>49/74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Party games</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>5.89</td>
<td>4.19</td>
<td>5.60</td>
<td>3.89</td>
<td>1.70</td>
<td>1.72</td>
<td>87/36</td>
<td>87/34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sports games</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>4.99</td>
<td>3.49</td>
<td>4.86</td>
<td>3.60</td>
<td>1.50</td>
<td>1.26</td>
<td>63/64</td>
<td>63/62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fighting</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>5.13</td>
<td>3.64</td>
<td>4.58</td>
<td>3.41</td>
<td>1.49</td>
<td>1.17</td>
<td>34/79</td>
<td>34/78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Racing games</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>5.33</td>
<td>3.99</td>
<td>5.19</td>
<td>3.82</td>
<td>1.34</td>
<td>1.37</td>
<td>62/61</td>
<td>62/60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excergames</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>5.50</td>
<td>4.44</td>
<td>5.25</td>
<td>4.47</td>
<td>1.06</td>
<td>0.79</td>
<td>38/77</td>
<td>38/8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Puzzle</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>4.77</td>
<td>3.71</td>
<td>5.15</td>
<td>4.22</td>
<td>1.06</td>
<td>0.94</td>
<td>65/53</td>
<td>65/52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educative</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>4.71</td>
<td>3.70</td>
<td>5.37</td>
<td>4.53</td>
<td>1.00</td>
<td>0.85</td>
<td>36/84</td>
<td>36/81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Platform</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>5.42</td>
<td>4.80</td>
<td>5.00</td>
<td>4.34</td>
<td>0.63</td>
<td>0.66</td>
<td>39/79</td>
<td>39/79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real life simulation</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>5.33</td>
<td>4.13</td>
<td>4.85</td>
<td>4.00</td>
<td>1.19</td>
<td>0.84</td>
<td>23/83</td>
<td>23/83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMORPG</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.87</td>
<td>2.97</td>
<td>-</td>
<td>-</td>
<td>16/95</td>
<td>16/91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pet raising</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.03</td>
<td>3.26</td>
<td>-</td>
<td>-</td>
<td>6/104</td>
<td>6/101</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professions</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.58</td>
<td>3.59</td>
<td>-</td>
<td>-</td>
<td>2/79</td>
<td>2/78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text based adventure</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.32</td>
<td>2.48</td>
<td>-</td>
<td>-</td>
<td>2/72</td>
<td>2/68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Statistically significant diff. p < 0.05, summated scale (S), factor scale (F), single variable (Si). - = Few observations.

** MD = Mean difference**

**Appendix 3. RECO sample size, mean differences, and results on linear regression (HED/UT).**

<table>
<thead>
<tr>
<th>Category</th>
<th>RECO Mean</th>
<th>MD Reco</th>
<th>n</th>
<th>r²</th>
<th>n</th>
<th>n M **</th>
<th>HED C ***</th>
<th>UT C ***</th>
<th>HED i %</th>
<th>UT i %</th>
<th>MIV ****</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EX IN</td>
<td>EX/IN</td>
<td>Adj.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital games</td>
<td>4.53</td>
<td>3.09</td>
<td>1.43</td>
<td>59/76</td>
<td>0.28</td>
<td>107</td>
<td>28</td>
<td>0.42</td>
<td>0.16</td>
<td>72</td>
<td>28</td>
</tr>
<tr>
<td>FPS</td>
<td>4.64</td>
<td>2.01</td>
<td>2.63</td>
<td>50/84</td>
<td>0.75</td>
<td>105</td>
<td>30</td>
<td>0.35</td>
<td>0.53</td>
<td>39</td>
<td>61</td>
</tr>
<tr>
<td>War strategy</td>
<td>4.91</td>
<td>2.46</td>
<td>2.46</td>
<td>47/87</td>
<td>0.72</td>
<td>107</td>
<td>28</td>
<td>0.56</td>
<td>0.30</td>
<td>65</td>
<td>35</td>
</tr>
<tr>
<td>Action adventure</td>
<td>5.12</td>
<td>2.87</td>
<td>2.26</td>
<td>49/83</td>
<td>0.66</td>
<td>107</td>
<td>28</td>
<td>0.32</td>
<td>0.51</td>
<td>39</td>
<td>61</td>
</tr>
<tr>
<td>Party games</td>
<td>5.63</td>
<td>3.64</td>
<td>1.98</td>
<td>88/45</td>
<td>0.78</td>
<td>109</td>
<td>26</td>
<td>0.74</td>
<td>0.15</td>
<td>83</td>
<td>17</td>
</tr>
<tr>
<td>Sports games</td>
<td>4.71</td>
<td>3.07</td>
<td>1.64</td>
<td>63/71</td>
<td>0.63</td>
<td>113</td>
<td>22</td>
<td>0.39</td>
<td>0.43</td>
<td>48</td>
<td>52</td>
</tr>
<tr>
<td>Fighting</td>
<td>4.29</td>
<td>3.00</td>
<td>1.29</td>
<td>34/99</td>
<td>0.70</td>
<td>99</td>
<td>36</td>
<td>0.48</td>
<td>0.37</td>
<td>57</td>
<td>43</td>
</tr>
<tr>
<td>Racing games</td>
<td>4.74</td>
<td>3.50</td>
<td>1.24</td>
<td>62/70</td>
<td>0.53</td>
<td>114</td>
<td>21</td>
<td>0.47</td>
<td>0.28</td>
<td>63</td>
<td>37</td>
</tr>
<tr>
<td>Excergames</td>
<td>4.95</td>
<td>4.08</td>
<td>0.86</td>
<td>38/95</td>
<td>0.58</td>
<td>104</td>
<td>31</td>
<td>0.68</td>
<td>0.09</td>
<td>88</td>
<td>12</td>
</tr>
<tr>
<td>Puzzle</td>
<td>4.68</td>
<td>3.71</td>
<td>0.97</td>
<td>65/65</td>
<td>0.48</td>
<td>113</td>
<td>22</td>
<td>0.29</td>
<td>0.44</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>Educative</td>
<td>5.06</td>
<td>4.12</td>
<td>0.93</td>
<td>36/99</td>
<td>0.50</td>
<td>110</td>
<td>25</td>
<td>0.07</td>
<td>0.66</td>
<td>9</td>
<td>91</td>
</tr>
<tr>
<td>Platform</td>
<td>4.78</td>
<td>3.68</td>
<td>1.09</td>
<td>40/92</td>
<td>0.57</td>
<td>114</td>
<td>21</td>
<td>0.48</td>
<td>0.30</td>
<td>62</td>
<td>38</td>
</tr>
<tr>
<td>Real life simulation</td>
<td>4.65</td>
<td>3.46</td>
<td>1.19</td>
<td>23/107</td>
<td>0.60</td>
<td>99</td>
<td>36</td>
<td>0.52</td>
<td>0.28</td>
<td>65</td>
<td>35</td>
</tr>
<tr>
<td>MMORPG</td>
<td>-</td>
<td>2.50</td>
<td>-</td>
<td>16/117</td>
<td>0.65</td>
<td>94</td>
<td>41</td>
<td>0.55</td>
<td>0.27</td>
<td>67</td>
<td>33</td>
</tr>
<tr>
<td>Pet raising</td>
<td>-</td>
<td>2.86</td>
<td>-</td>
<td>6/129</td>
<td>0.48</td>
<td>94</td>
<td>41</td>
<td>0.27</td>
<td>0.45</td>
<td>37</td>
<td>63</td>
</tr>
<tr>
<td>Professions</td>
<td>-</td>
<td>3.47</td>
<td>-</td>
<td>2/129</td>
<td>0.48</td>
<td>69</td>
<td>66</td>
<td>0.31</td>
<td>0.41</td>
<td>43</td>
<td>57</td>
</tr>
<tr>
<td>Text based adventure</td>
<td>-</td>
<td>2.86</td>
<td>-</td>
<td>2/130</td>
<td>0.51</td>
<td>62</td>
<td>73</td>
<td>0.38</td>
<td>0.36</td>
<td>51</td>
<td>49</td>
</tr>
</tbody>
</table>

* All linear regression models and differences (means t-test) were statistically significant at p<0.05.
** Non-used observations.
*** Variable coefficient, **bold** = statistically insignificant variable.
**** MIV = Main interacting variable.
The Two Cultures and the Internet Revolution

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0298.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Culture fit / differences / heritage/ intelligence / issues / theory / values, Alignment, Socio-technical theory, Information literacy</td>
</tr>
</tbody>
</table>
THE TWO CULTURES AND THE INTERNET REVOLUTION

Abstract

This document reinterprets C. P. Snow’s famous “Two Cultures” (the so-called “literary elite” and scientists) lecture of 1959, in light of advances in information systems in the past fifty years. While Snow referred to specific groups, his analysis is generalizable: cultural groups differentiate through lack of communication. Here Snow’s analysis and advice are applied to a different pair of “cultures” (IT purveyors and IT users) as an example of his general principles. At a time of great unease about terrorism in the face of apparently relentless technological advance—analogous to Snow’s speech at the height of the Cold War—and also during a time of (then) apparently dramatic technological advance, the lessons Snow derived can now apply usefully to today’s specific “two cultures” case.
1 MID-CENTURY ANGST

The year 2009 is the 50th anniversary of Sir Charles Percy Snow’s now-classic “Two Cultures” Rede Lecture at Cambridge University (Snow, 1959, 1971). This lecture seemed, at the time, to be a polemic intended to inflame and annoy and yet is now clearly typical of the mid-20th-century essay attempting to find order in the chaos created by conflicts between science and the arts (Boytinck, 1980; Davis, 1965; Glass, 1959; Halperin, 1983). This current 2009 essay seeks to find lessons in Snow’s lecture for practitioners of the “science” of information systems; position IS intellectually on a far broader canvas than is usual, and contribute to critical thought concerning what IS is really about. Using Snow as a springboard for insight, we come to the conclusion that users can and should play a bigger, more pivotal role in determining use and ultimately the value of information systems.

Here is the plan of the paper: to recast Snow’s “two-culture” hypothesis as a general principle by revisiting his original speech and then to apply that general principle to the current interaction of another “two cultures”: purveyors (ITP: developers, marketers, and researchers of information technology) and users (ITU: those who employ IT for business and daily life). It will focus on the following questions:

- What is the essential argument that Snow made about the “split into two polar groups,” in his words, of Western society, independent of the specific groups he identified?
- Which of Snow’s statements concerning the “gap” he pronounced as existing between the two cultures he cited in 1959 are true of today’s ITP and ITU “cultures”?
- Are the causes of today’s cultural gaps identical, similar, or analogous to those Snow decried in 1959? Are they in any sense predictable and/or controllable?
- Are the lessons Snow derived in 1959 applicable today to the “user-IT” gap?

There is a caveat. When Snow spoke about the “scientific revolution” in his speech, he certainly did not mean the Copernican revolution, the rise of the scientific method and the development of scientific thinking. It is the Industrial Revolution to which he is really referring. He said in his lecture that the Scientific Revolution “comes from the application of real science to industry, no longer hit and miss, no longer the ideas of odd ‘inventors’, but the real stuff.” Similarly, when we refer here to the “Internet Revolution” we are not necessarily referring to the immense change in focus that the Internet is bringing about, including virtual communities, globalization, and vastly distributed computing. Instead, we are thinking about the application of these ideas to everyday life. To rephrase Snow’s words for today: “The Internet Revolution comes from the application of real global, networked information to industry, no longer hit and miss, no longer the ideas of isolated ‘technologists’, but the real stuff.”

The middle of the twentieth century was a time of great re-thinking. Not only had a great war devastated much of the heart of modern civilization (east and west), but atomic threats to the survival of the human race were made palpable to the public through modern realistic media such as television. Most of what passed for history during the second half of the twentieth century was a working out of the great clashes of the first half: the end of empires and colonialization, the modernization of societies, the shifting of labor from the farm to the factory and the office, the movement of people from the countryside to cities, an explosion in literacy and democracy. In every sense, the 1950s were a time of thinking about and exploring the implications of the great dichotomies of the previous half century: freedom vs. totalitarianism, ignorance vs. knowledge, privilege vs. participation, fear vs. security, workers vs. the establishment, gender roles, and so on.

In this intellectual environment in 1959, the British novelist-scientist-politician Charles Percy Snow, not
yet Baron Snow of Leicester, delivered the annual Rede Lecture at Cambridge University. The series provides a forum for noted individuals to discuss the issues of the day. For Snow, delivering the Rede lecture was not only a personal triumph for the scientist-turned-author-turned-bureaucrat, but also a chance to deliver a personal statement very much in tune with the mid-century spirit of the times.

Snow’s career embodied many of the dichotomies that were the subtext for the mid-century angst expressed on both sides of the Atlantic and elsewhere. In addition, he brought together the disparate intellectual threads found in the range of speakers in the Rede Lectures. The son of a working-class Leicester family, he became a life peer; an outsider in an “old-school” British Establishment society. He created a new kind of literature documenting his own kind – the scientist-bureaucrat. He was appointed to position of parliamentary secretary to the first Minister of Technology in 1964 and served in this position two years (the ministry survived on its own only four more years, until 1970).

Snow was thus in a unique position to see all sides of technological progress, or lack thereof, from both the scientific as well as the high-level government viewpoint. Sir Charles delivered a lecture that was a synthesis of a number of discourses on the topic of “gaps” or “divides” and how to bridge them or narrow them. True to the mid-century ethos, he dwelt on these, among others, distinctions: the West vs. the East, science vs. the arts, pure science vs. applied science, literary intellectuals vs. scientific intellectuals, the wealthy countries vs. the poor countries, industry vs. art, the political left vs. the political right, among others.

Snow’s conviction in 1959, in the face of real danger of all-out nuclear warfare directed by governments that didn’t really understand the science, was that one very useful way to defuse the situation was through technology transfer. Not, of course, nuclear technology transfer, but scientific knowledge transfer. The mid-century nuclear angst passed with the fall of the Berlin Wall. It is questionable whether an analogous information wall’s falling (i.e., the advent of the Internet) relieves the early 21st-century information angst.

Snow made three major points in his lecture, one empirical and subject to test, one controversial and highly personal, and one universal and largely ignored. The empirical statement concerned what he termed “the two cultures”, namely that scientists and the “literary intellectuals” (i.e., British writers, critics and university professors) operated in completely separate, mutually non-communicating cultures (one major problem with Snow’s lecture is that he provided no consistent, acceptable, reliable definition of “culture”), each with its own language, rites, rituals, and values. This point was uncontroversial in its essence and not novel (for example, Matthew Arnold delivered a Rede lecture on the same topic in the 1880s). However, after the lecture, this point was misunderstood by many and used as a smokescreen for those wanting to avoid discussing the second, more controversial, point.

This second point was that the culture gap was a threat to civilization and that it was the job of education to repair this. The point became controversial because he chose to criticize the very social system that enabled the rise of Baron Snow by implicitly derided a seven-hundred-year-old educational system. Less controversial in America, Snow’s second point was welcomed with open arms by educationalists eager to reform a system always ready for reform. But in Britain, the debate focused instead on Snow’s literary “credentials” as a member of the literary intelligentsia (Leavis, 1963).

Point number three was that the cultural and educational gaps aid and abet a third gap, that between the rich and the poor worldwide, and that this gap would respond positively to technological knowledge transfer. Of course Snow didn’t invent this idea; the World Bank came into existence in 1944 for reconstruction of war-torn Europe. Snow’s idea, of a kind of technological Peace Corps (to use an American term from the 1960s), was well intentioned and useful, but generally ignored. It’s an idea that
needs to be reexamined.

In summary, Snow could point out in 1959 that the world was dangerously divided in multiple ways around the role of science amidst the threat of atomic annihilation, driven by economic inequalities. In 2009 the same is still true, but the pivot is information. If Snow’s thoughts could be brought up to date and put into information terms, what advice would he give now analogous to what he suggested in 1959? Each of the ideas in his speech is worthy of revisiting, but it is to his first two that we turn here, mapping his thoughts from “science” to “information.” Sections two through four of this essay refer to his first idea, that of the “two cultures” gap and its implications. Sections five and six address ways of reducing this gap.

A note is in order on the use of the term “culture.” Snow did not give it much thought in 1959. At times he seemed to be referring to “high” culture (i.e., literature and the fine arts) and at other times merely to group membership. As we will see later, “shared values” may serve better as a distinguishing characteristic, although Snow did not dwell on this concept in the Rede lecture, but of course a great deal of thinking has been directed towards this term in the past 50 years. Leidner and Kayworth (2006) have provided the seminal work here. They chose to conceptualize this troublesome idea of “culture” as “shared values”. They would fit Snow’s thoughts -- and our recasting of them in modern terms -- into their theory of IT-culture conflict as contribution conflict arising from mismatches between user group values and IT values. While this does not completely capture the very broad and inexact application of the term “culture” that Snow intended, it is useful for our discussion below.

2 TWO POLAR GROUPS

The Rede Lecture gave rise to the “Two Cultures Hypothesis.” The largest part of the speech presented evidence for a gap between two important groups in mid-century British intellectual life: scientists and the “literary elite.” Little of his speech was true science subject to hypothesis testing. To refer to his idea as a hypothesis stretches the metaphor. Nonetheless, let us assume that the gist of what he was referring to was true in 1959. What is the essential argument that Snow made about the “split into two polar groups,” in his words, of Western society?

C. P. Snow referred specifically to “literary elite” (sometimes he refers to them as “literary intellectuals” and other times merely as “intellectuals”) and “scientists.” It is clear that the literary intellectuals he referred to were writers, essayists, Oxbridge dons, and to some extent the journalists who make the intellectual basis available. By extension, he was referring to the British political and social elite with this term. Initially – and increasingly – Snow’s “literary intellectuals” became identified with “the arts” and numerous essays appeared to debate the “science vs. the arts controversy” (Boytinck, 1980).

By “scientist”, Snow meant practicing bench scientists, individuals who derive and test scientific hypotheses. Snow was not comfortable with including technologists and engineers in this group and it is doubtful that he would have had as much to say about applied scientists, social scientists, and computer scientists, either. On occasion he does put engineers and scientists together, but generally he makes a strong distinction between pure scientists and those who are applying scientific principles.

Why were and are these two groups important? One might, with the comfortable space of fifty intervening years, conclude that Snow’s polemic was personal. Perhaps the only real reason he felt the need to draw attention to the gap between these two groups was that as a novelist and a physicist Snow felt drawn to both camps. But if we follow O’Hear’s (n.d.) argument, these are two essential groups in a society such as ours as they represent ways of apprehending the world through either a value-tinged lens or a value-free
lens. O’Hear said

To put all this another way, science aims at an observer-independent account of the world, transcending human meaning, culture and ideology. Its success derives from its success in approximating to this aim, for it is in so far as we go beyond looking at the natural world in terms of its first meanings for us that we are able to penetrate further its causally essential core, and so become rather more adept at manipulating and directing it than those who remain at the level of first impassions. The lesson of post-Galilean science is that there is no reason to suppose that the effects and processes we identify in our first transactions with nature will turn out to be those which are fundamental from a causal point of view.

Hence, by working in complementary ways, science and the arts put the value basis of our living in stark relief. By transcending “human meaning”, science, in O’Hear’s words, shows us where values are active. Without science, everything is “impassioned”. Since we cannot reliably describe, define, and delimit the passions and come to understand everything, the literary elite need science.

Similarly, science also needs the literary elite. O’Hear says that “…an exclusive concentration on scientific modes of thought can affect the way in which judgments of value are made. In particular, it can lead to an importation of quantitative considerations, and a tendency to see social and moral problems in terms of hygiene and environmental manipulation.” There are two complementary ways of apprehending reality. This complementarity is a characteristic, too, of the ITU/ITP “gap”.

### 3 TODAY’S “TWO CULTURES” AND TODAY’S “GAPS”

Let us now shift the emphasis and focus from “Would Snow be right today?” to “Is there something happening today that is like what Snow was describing both in extent as well as impact? ITP and ITU communities share characteristics that separated Snow’s original “two cultures.” ITP and ITU speak different languages, have experienced different education; have paths that, if they ever intersect, cross only at a desktop computer; and generally appear ignorant of each other’s ideas and values.

The “gap” between these two groups is legendary. For much of the past fifty years, there has been a monumental struggle within business organizations concerning information tools and their purveyors. Beginning in the 1960s with the concept of “resistance” and moving through the “end user revolution” and “IT/business alignment” to today’s thoughts about “IT governance”, the struggle over who gets to control IT investment and deployment has endured through multiple generations of software, hardware and Microsoft products. It does not seem to go away, whatever the current manifestation of the presenting problem. The underlying “cultural” gap may be the ultimate cause. Complaints of “The system doesn’t work!” “The help desk is unresponsive”, “The IT guys are arrogant and speak in computerese”, and “IT is too expensive” come from the ITU side. “IT doesn’t sit at the ‘table’”, “We are seen as plumbers, not as contributors to strategy”, “I can’t get my budgets approved” and “It’s impossible to keep talented staff in this environment” come from the ITP side. This certainly indicates a gap. Is it the kind of gap that Snow referred to in 1959? Is it as important? Is the “ITP/ITU gap” the “two cultures gap” of 2009?

One way to approach this question is to think about which of Snow’s statements concerning the science/arts “gap” in 1959 are true today of ITP and ITU. Are ITU and ITP “cultures” in Snow’s sense? Do they have the kind of alienation Snow alluded to in 1959? Are the predictions the same? How can Snow’s hypotheses be understood in terms of today’s society and its information issues?

Clearly there are important differences. Science is a system of procedures and philosophies with a long
history and the goal of understanding the physical universe. The liberal arts represent the accumulated total of human experience and thought on moral issues. On the other hand, the activities and the body of thought of ITP go back at most about sixty years. This body of thought is generally limited to two major areas: business (or, more broadly, administration) and engineering (including software engineering and computer science). The comparison might seem stretched and nebulous. Unlike science, there is no real canon of IT practice; no ethical history; no historical struggle against, say, established teachings; and no systematic methodology comparable to the scientific method. Unlike the practitioners of the liberal arts, the users of IT (ITU) are not a group of intellectuals focusing on the fiber and core of civilization; they aren’t homogeneous and dedicated to increasing our understanding of issues unrelated to profit or proper administration; they don’t wield “power” in any arena and are, in many cases, totally powerless, at least to get their IT “working properly.”

Yet these differences do not dispel the comparison. Because IT is now critical to the survival of culture (which has come to depend on digital information and the World Wide Web to keep commerce and culture alive) and because IT users are in fact the drivers of that culture, Snow’s comparisons are particularly apt. Here are some from the original lecture, slightly modified to reflect the new groups:

[T]he intellectual life of the whole of western society is increasingly being split into two polar groups. When I say the intellectual life, I mean to include also a large part of our practical life…. Two polar groups: at one pole we have the [information users]…. at the other [information purveyors]…. Their attitudes are so different that, even on the level of emotion, they can’t find much common ground. On each side … some of [this prejudice] which is not entirely baseless. It is all destructive. Much of it rests on misinterpretations which are dangerous. [Emphasis mine]

What about the relationship between these two “cultures”? Surely IT users use IT in ways that the “intellectual elite” never “used” science in 1959? IT users depend on IT more directly and intensely than the intellectual elite of 1959 depended on science. Innovative IT users sometimes develop their own systems and applications using common packages such as spreadsheets or databases. And it is unlikely that the intellectual elite of 1959 developed scientific hypotheses which they then tested in experiments. Is this comparison strained, then?

Yes and no. The alienation Snow spoke of was related mostly to a lack of communication, the lack of a common language, differences in values, and a large “gap” in terms of attitudes and values. In Snow’s own words:

The two cultures were already dangerously separate sixty years ago….Thirty years ago the cultures had long ceased to speak to each other: but at least they managed a kind of frozen smile across the gulf. Now the politeness has gone…. It may well be that this process has gone too far to be reversible.... Closing the gap between our cultures is a necessity in the most abstract intellectual sense, as well as in the most practical. When those two senses have grown apart, then no society is going to be able to think with wisdom.

Today, this alienation takes on a more modern appearance if we stop thinking about science as an esoteric practice completely out of reach of the literary elite and instead view it more as a different way of understanding the world. Of course “the world” in Snow’s terms was restricted to the physical universe – and he focused strongly on physics as the exemplar of science. A C. P. Snow of 2009 would have to recognize social science and administrative science as legitimate scientific ventures, whose purpose is to understand in a systematic way a limited set of phenomena, some of which take place in the physical
universe, some of which take place within derived social – but still physical – domains and some of which involve virtual domains such as money and influence. Yet, within these domains, the goal is still understanding. But the goal of the literary elite is also understanding -- understanding the human heart, in its broadest (and least medical) sense: human endeavor, aspiration, love, morality, history and conflict. It was just a different way of understanding from that of science. In both cases, there is an “out there” that needs to be understood by people “in here.” Science and the arts are separated most fundamentally by their ways of apprehending the world. That this understanding is *supposed* to be objective for science and *can be* comfortably subjective for the arts is important for judging the aptness of our comparison of ITP and ITU as cultures.

What separates IT purveyors and IT users to generate a gap comparable to that between science and the arts? The answer must be that information is the “object” of the purveyors and the “subject” of the users. The purveyors value information as an object, one that must be handled systematically, carefully, and with great respect for its beauty and brittleness. Most IT users – especially business users -- value information because of what they do with it. The information describes the subject of their work lives so information is respected respect for what it means, not what it is. Information is, ironically, a work of art to the purveyors, and a means to an end to the users. Information *represents* power, confidence, and influence to the users. Information *is* none of this to the purveyors; it represents nothing, but is itself, the thing. And this difference is precisely the same as that dividing science and the arts.

Are the two IT groups “cultures” in the same sense in which Snow applied the term? Snow was not an anthropologist but he was using the term “culture” in its usual (dictionary) sense: a particular group of people and their ideas, which are passed along or transferred to successive generations. Both scientists and literary intellectuals do indeed have sets of ideas that they pass along, through education and the media to which they contribute and subscribe, to successive “generations”. Within their own ambit, both science as well as the literary and intellectual arts do ascribe ownership to sets of ideas (sometimes referred to as “schools”), provide titles, and enforce customs and regulations. It was this restricted sense of “culture” that made it easy for Snow to identify these two “cultures” and their alienation, since they did not share “a language” between them, but only shared a language within a group. He said, “[A]fter a few thousand Atlantic miles, one found Greenwich Village talking precisely the same language as Chelsea, and both having about as much communication with M.I.T. as though the scientists spoke nothing but Tibetan.” Thus knowledge and language served to divide cultures in Snow’s mind, although he also referred to shared values (scientists were, to Snow, what we would refer to today as “liberals”; the literary elite, “conservatives”) but treated differing values more as the result of differing cultures than the cause, unlike Leidner and Kayworth (2006).

In the way Snow used the term “culture”, it is difficult to label ITU a homogeneous “culture”. But clearly “users”, despite the enormous difference in their jobs, goals, and work processes, share similar attitudes and values, at least with respect to information. Perhaps this sharing stems from a common feeling of helplessness at the end of the phone line to the help desk. Or maybe these attitudes stem from accumulated negative emotions through years of experiencing unresponsiveness or even arrogance of the ITP. Regardless of the source, evidence of the cultural “values” of the ITU are mostly evidenced by resistance, coping strategies, and work-around activities (Jasperson, Carter and Zmud, 2005; Lapointe and Rivard, 2005; Beaudry and Pinsonneault, 2005). In general, their experiences are at best neutral. And these experiences and values are definitely “passed along” through on-the-job training, office “lore” and gossip. This is alluded to by Leidner and Kayworth (2006) in their listing of “IT values”. Their contention is that a mismatch between IT values and “group member values” (i.e., user values) results in “contribution conflict” a mutual misunderstanding of what IT is “about.”
ITP “culture” is easier to describe. Because IT education is generally highly circumscribed, the IT world view is relatively homogeneous, centered on the technology itself and the supporting platforms such as application environments, operating systems, physical equipment, protocols, etc. The ITP view of ITU is that ITU users are the uncontrollable variable at the other side of the “interface”. This is not to deny that there are sub disciplines in the ITP universe that focus on users. We have multiple theories of technology adoption, end-user system involvement, and human-computer interaction. The field of “management information systems” promotes the idea that information systems are essential in business and work towards bridging the IT-user “gap”. Yet there are few useful theories of the role of the user in system use. The literature of user involvement generally focuses on users as objects of study (during the early stages of system development) or as difficult-to-handle interfacing systems (looking at the diffusion or adoption of innovation and the “resistance” literature). There is almost no post-implementation literature looking at user innovation, user stewardship or ownership, or user responsibility. Just as scientists do not really worry daily about whether non-scientists will actually apply the principles upon which gunpowder functions as an explosive to build and operate lethal weapons, so, too, are members of the ITP largely unconcerned about what users actually do with information systems. The values of the ITP rarely extend to any aspect of the deployment of systems. And these values and experiences are also passed “along” through on-the-job training, office “lore” and gossip.

4 THE CAUSES OF THE “GAPS”

If we have two IT cultures with a relationship strongly analogous to those Snow described, can we also ascribe this relationship to causes similar to those Snow referred to? Snow strongly implicated specialized education as the major reason for the gap between his two 1959 “cultures”. ITU and ITP also have differing educations. Just how different are these educations? And is it the education that is the cause, or is it the effect? In 1959, Snow referred extensively to an educational difference as “specialization”. British university education in the 1950s was highly specialized. Even today, students enter these universities directly into narrow specializations in science, literature, and other fields. In America, most universities and even most institutes of technology require a one or two years of “liberal arts” education in any degree program. Few students enter a “major” in their first year of university training. Things are a bit different in technical schools, where training is specifically in the discipline.

However, one important difference between the “two cultures” of Snow’s 1959 and the “two cultures” of 2009 is that science education was then (and remains now) a purely academic discipline, whereas IT education in the US can be obtained in a variety of ways. Typically any bachelor’s or first university-level degree will have a large liberal arts component. For members of the ITP, understanding of the ITU will come from daily life, for the most part, rather than from academic training. Few offerings to members of the ITP contain any content oriented towards understanding the role, impact, function, responsibility, or worth of information and information technology in general society, even in liberal arts course. Courses that may be required of some members of the ITP may examine “social issues” or “ethics”, but unless the lessons are reinforced through positive and clear examples, they are hardly going to survive as important bridges after this training.

For members of the ITU, any vague understanding of the ITP can stem from modern life, with its myriad computer-based, -enabled, and –enhanced applications. There are the ubiquitous “computer literacy” courses offered at universities, colleges, community and seniors' centers, voluntary organizations, service clubs and prisons; all provide some introduction to one or more applications under the heading of “literacy”. Whether or not this kind of literacy is really anything like reading is controversial and is very much related to the “two cultures” challenge, if only because building the bridge across the gap from only
one side is dangerous engineering. Just having users understand “how computers work” is not sufficient to bridge the gap.

Secondly, beyond the lack of IT cross-training is a lack of shared positive experiences. A typical member of the ITU will experience information systems in two different ways: “I use an application and it’s great; there are no problems” or “What a piece of garbage!” In the first case, there is little reason to explore the other “culture” – rarely are members of the ITU asked to contribute in meaningful ways to subsequent releases of systems that already work well. In the second case, interaction between ITU and ITP people generally suffers from all the communication problems noted by Snow, exacerbated by the pressure of modern life to get things done fast using information-based, enabled, and –enhanced system. When systems don’t seem to work and when users have little recourse to getting them fixed, users manifest anger and resentment. Often individuals in the ITU try to become knowledgeable about IT either in self-defense, rarely out of interest. They become the office “computer guy” even when they aren’t in any way members of the ITP fraternity. This might create more positive experiences for others in the office, but it doesn’t build shared positive experiences.

A third cause of the gap, especially in the business arena, is the development of IT ghettos, areas of the organization into which members of the ITP are relegated and isolated from the rest of the organization. While most firms would deny this reality, the numbers clearly tell the true story. Few members of the ITP ever become CEOs (Chief Executive Officers) of any non-tech firm. Few CEOs of modern companies have any vocational background in IT. Although there is some disagreement over the statistics, fewer than half of CIOs (Chief Information Officers) or individuals having a CIO-like position have a meaningful role in determining corporate policy or strategy. “Having a seat at the table” is the cause célèbre of the IT trade magazines. This seating varies depending on the industry, with information-rich industries such as banking and entertainment promoting CIOs to “table” status but most primary and heavy industries relegating the CIOs to reporting through Chief Financial Officers. The concomitant loss of prestige and benefits to the members of the ITP because of the relatively low status of their primary advocates is harmful enough. But by denying positive and regular interaction with CIOs leaves CEOs impoverished and information-disempowered, at the mercy of Boards of Directors and vendors when it comes to making decisions about IT. Even worse than potentially negative economic outcomes is the lack of contact itself, which leads to disinterest, or at least further isolation. This can only hurt IT governance, as an example of one critical aspect of IT. There is no doubt that most governance schemes, such as ITIL, focus on procedures rather than governance. This may be because the organization itself cannot be trusted to “do the right thing” with IT, because of a history of poor communication focused on problems, a lack of a common language and values, and totally different world views. This is the legacy of “two cultures”, exactly paralleling Snow’s description from 1959. Snow’s reference to Tibetan is particularly apt, since (1) few people speak Tibetan, (2) Tibet is a long, long way from the West, and (3) Tibet no longer exists as a separate political entity. Each of these is a quality of the ITP, at least in business organizations.

The underlying causes can be addressed. It is possible for an IT person to interact meaningfully with users and it is possible for a user to argue for IT. However, it is unlikely to happen without a major shift in business values. For instance, for two decades, alignment has been either the top issue for CIOs or in second place. Aligning IT to corporate strategy is always on CIOs’ minds, it seems. This lack of alignment is, of course, simply a symptom of the “two cultures.” Generally, the emphasis is on “aligning IT to the firm” rather than the other way around. There is a large literature on this topic, too great to be reviewed here. There is, however, no literature on the opposite: aligning the firm to IT. And there is only a very small body of thinking about mutual alignment (Luftman, 2000, 2003, 2007). Keeping the alignment “problem” in mind, let’s turn, once again, to Snow’s lecture to see how these two cultures can be brought together, perhaps through lessons he pointed out half a century ago.
5 SNOW'S LESSONS

Snow’s diagnoses and prescriptions may have some value in 2009. In this section we will review several of them and see how the lessons can be applied. In essence they boiled down to providing more science education in the schools. But given how little difference such education has made in ameliorating the gap between his “two cultures” since 1959, one may legitimately wonder about the effects of such education on the ITU-ITP “gap”.

First, here’s what Snow prescribed in order to reduce or ameliorate the “two culture” gap of 1959:

[Twentieth-century science] has got to be assimilated along with, and as part and parcel of, the whole of our mental experience, and used as naturally as the rest. There is only one way out of all this: it is, of course, by rethinking our education. This [list], or something like [it], is the specification for the scientific revolution. First of all as many alpha plus scientists as the country can throw up. Second, a much larger stratum of alpha professionals -- these are the people who are going to do the supporting research, the high class design and development. Third, another stratum, educated to about the level of Part I of the Natural Sciences or Mechanical Sciences Tripos, or perhaps slightly below that. Some of these will do the secondary technical jobs, but some will take major responsibility, particularly in the human jobs. Fourthly politicians, administrators, an entire community, who know enough science to have a sense of what the scientists are talking about.

Snow advocated an increase in the number of scientists. But he also spoke strongly in favor of what might be called the “users-of-science community” -- politicians and administrators -- “an entire community” learning enough about science “to have a sense of what the scientists are talking about.” And by this, he did not mean understanding the basis of particle physics. He meant enough understanding of the importance of particle physics to be able to make critical judgments. That is what politicians, government officials, civil servants and business leaders do: they make decisions among courses of action. Snow, naturally enough as a scientist-politician (and a scientist-administrator, too, during WWII), knew that these decision makers were influential enough and had enough clout to move the rest of society. Snow proposed educational programs both for scientists and for the “literary elite.”

Was this effective? Unfortunately we have no way of knowing, because the intervening years were not kind to Snow’s ideas as scientific hypotheses. No one did the experiment. Instead, the enormous ramp up in technological and applied-science development has culminated in a technology-oriented society, at least in most of the West. This seems to have completely obviated the apparent need to understand science. America, for example, is a society that literally runs on technology. For example, the Internet is rapidly replacing the shopping mall as the price of fuel rises. America looks to technology for solutions to almost all social problems (for example, tethering of convicted criminals instead of imprisonment, high-tech medical care instead of nursing, e-learning instead of wide-spread education, for example). Yet decreasing numbers of Americans enter science and engineering programs each year.

Snow was also ambivalent about industry and technology:

And we know almost nothing about [the industrial society of electronics, atomic energy, automation. H]ighly educated members of the nonscientific culture couldn't cope with the simplest concepts of pure science: it is unexpected, but they would be even less happy with applied science…. Pure scientists and engineers often totally misunderstand each other….
It is simply that technology is rather easy. Or more exactly, technology is the branch of human experience that people can learn with predictable results.

Yet even this “easy” pursuit, which pure scientists misunderstand as much as highly educated members of the nonscientific culture, is unpopular with members of the ITU. Here in America, for example, the management of ITU “culture” is increasingly content to outsource its information needs to India and to populate their schools of engineering and science with students from India and China, countries that Snow ironically had much to say about in the third part of his lecture, concerning the developing world.

So perhaps “education” is not the answer. Instead, consider the underlying message of Snow’s argument in 1959. That message is that even a non-scientist who doesn’t understand science (in depth) should understand and be responsible for the implications of his or her use of science. That is, members of the “literary elite” (which correspond to our ITU today) should at least appreciate the importance of science in contributing to the basis of their own activities. This is responsibility.

To this end, Snow proposed a sort of scientific (and to an extent technological) “Peace Corps” to go to what we now call the “developing world” and bring it into the modern “scientific” era. His proposal was to bring the vehicle of science -- the scientific revolution -- to the rest of the world:

The West has got to help in this transformation [from half rich and half poor]. The scientific revolution on the world-scale needs, first and foremost, capital …including capital machinery. The second requirement is …trained scientists and engineers adaptable enough to devote themselves to a foreign country’s industrialization for at least ten years out of their lives. [T]he third essential of the scientific revolution … is an educational programme …With scientific teachers from this country and the United States, and … teachers of English, other poor countries could do the same in twenty.

In effect, Snow implied that while the individuals who might bring the “scientific revolution” (read here today “the Internet revolution”) to the developing world would be “scientists and engineers”, such a project would have necessitated the intellectual elite taking “stewardship” of science and technology. It is this stewardship idea to which we now turn.

“Stewardship” means the sense of responsibility that an individual has for the proper use and sustaining of a resource. We speak, informally, of being “stewards of the earth” when we refer to environmental responsibility. Similarly, it is possible to speak about “information” or “application” or even “system” stewardship in the same vein. Such IT-oriented stewardship builds a sense of responsibility and respect for information, information tools and, ultimately, information purveyors (i. e., members of the ITP) among the users (i. e., members of the ITU). Typically users expect the corporate IT unit to be responsible for designing, developing, distributing, and maintaining systems. For the most part, this responsibility rests well with members of the ITP. However, again typically, IT users also expect members of the ITP to be responsible for the use of the applications. The users certainly don’t feel this responsibility. For example, few firms have active programs to reward innovation in IT use. Even fewer firms consider IT users as potential initiators and managers of IT development projects even when the users are paying for these projects. It is doubtful whether firms have sufficient maturity to understand the concept of application stewardship, so ingrained is the idea that an information tool is something that only the ITP should be responsible for. Yet the use of such tools logically cannot be solely an ITP responsibility; it is also an ITU responsibility. Programs encouraging stewardship are important. Successful programs of stewardship remove both major causes of the “two cultures” gap and thus ameliorate the harmful implications that Snow dreaded fifty years ago. The “two cultures” must explore each other’s workplace reality. This is the
best reading of Snow's mid-twentieth-century advice for today's two IT cultures.

6 CONCLUSION

The gap described between ITP and ITU in 2009 is analogous to the “Two Cultures” gap of C. P. Snow's 1959. The pairs of cultures (ITP/ITU in 2009 and “scientists”/“literary elite” in 1959) differ, of course, but the causes of the gaps are similar. The effects of the gap are analogous, too, and probably just as threatening. For example, did the ability to create and track mortgage-backed securities enable the recent large-scale meltdown in worldwide financial markets? Information systems provide the ability not only to create a large range of unregulated derivatives of questionable provenance, but certainly to manage and make marketable these virtual “securities.” This is not the atomic Armageddon of Snow's subtext in 1959, but it could have analogous destructiveness if the economic system of the West (and Russia and China) comes tumbling down. One way forward may be to develop stewardship programs among users to increase user responsibility. Stewardship programs have an educational component as well as an experiential one. Not only will members of the ITU and ITP learn more about each other's “cultures”, but by having common and positive experiences, they will come to value each other's way of understanding the world. Each employs information in a unique, but complementary way.

REFERENCES


O’Hear, A. (n.d.). *Two Cultures’ Revisited.*


OPPORTUNITIES AND CHALLENGES OF MOBILE PERSONALIZATION: AN EXPLORATORY STUDY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0325.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>M(obile)-banking / M-commerce, E-commerce (B2B / B2C / B2G / G2C), Human computer interaction (HCI), Consumer behavior / choice / demand / empowerment / reviews / consumerism</td>
</tr>
</tbody>
</table>
OPPORTUNITIES AND CHALLENGES OF MOBILE PERSONALIZATION: AN EXPLORATORY STUDY

Ho, Shuk Ying, The Australian National University, Canberra, 0200, Australia, susanna.ho@anu.edu.au

Abstract

In the decade since the creation of wireless handheld devices, mobile commerce (m-commerce) has become a ubiquitous channel for accessing information and conducting business. Mobile users can now access information anytime and anywhere. Mobile advertising, retailing, and gambling are popular, and gradually the competition among mobile services providers turns fierce. Hence, some services providers adopt personalization technologies to customize content for their users. This paper explores the opportunities and challenges of the use of personalization technologies in m-commerce. Although the effectiveness of personalization on the web is well-examined, there is little work on personalization in mobile services. The debates regarding the effectiveness of personalization and technological limitations and privacy concerns motivate us to conduct focus groups with mobile users, and explore the opportunities and challenges of personalized mobile services. The focus groups findings illustrate that mobile users are very concerned about their privacy and spam. We then extract a list of personal information from the focus groups findings. This list of personal information is highly related to mobile users’ privacy concerns. We conduct an online survey to gain a better understanding of which piece of information mobile users are more willing to share with services providers and we perform a multi-dimensional scale analysis.

Keywords: Personalization, Mobile Commerce, Expectation, Focus Groups, Multi-Dimensional Scale Analysis
1 INTRODUCTION

1.1 Personalization in Mobile Commerce

In the decade since its creation, mobile commerce (m-commerce) has become a ubiquitous channel for accessing information and conducting business (Allen 2003; Tarasewich et al. 2002). M-commerce is the buying and selling of products and services through wireless handheld devices (Hong and Tam 2006). These devices range from small handsets, such as cell phones and personal digital assistants (PDA), to wireless laptop computers. Apart from retailing, there are applications on mobile advertising and mobile gambling. According to Wu and Wang (2005), only 16% of half a billion Internet users were global wireless Internet users in 2001; however, the percentage jumped up to 57% of 1.5 billion Internet users in 2007. A telecommunications analyst, Paul Budde, estimated that 300 million text messages are sent by 11.5 million mobile phones over the three Australian mobile networks each month. Undoubtedly, m-commerce is gaining significance.

The proliferation of wireless-based applications has turned these devices into an essential “touch point” of corporate services (Gopal and Tripathi 2006; Stafford and Gillenson 2003). With recent advances in wireless location technology, such as Global Positioning Systems (GPS), mobile operators can identify the location of a wireless device within several metres. Prior research asserts that with such a location awareness capability, this mobile channel provides firms with more valuable promotions and sales opportunities than the web channel does (Balasubramanian et al. 2002).

In leveraging the strengths of m-commerce, the question that lurks in the back of mobile operators’ minds is: what type of location-based content should be delivered? With increasing customer expectations, presumably, one-message-fits-all is not a strategic move. Customers prefer customized, or even personalized, services (Blom and Monk 2003). Tailoring the content in relation to users’ locations (and even their preferences, shopping goals and context) is necessary in order to realize business opportunities in m-commerce. Thus, personalization is the answer for mobile operators.

Personalization is to tailor content to the needs of individual consumers. The technology enabler, a personalization agent, is a collection of software modules that deploys tools to collect and analyze customers’ behavior and their purchase transactions. These modules, including data mining, collaborative technology, click stream analysis components, and pattern recognition, allow real-time detection of user behavior and manipulation of web content (Ardissono et al. 2002). The agent is a context-aware application designed to deliver targeted promotions to online users about the products they like and protect them from information overload. Providing responsive and high quality services is a key factor for firms to achieve a sustainable competitive advantage. The goal of personalization is to leverage the above technologies to provide the right content in the right format to the right person at the right time in the right location. That said, not only are firms now able to retrieve the profile of a user when the user accesses a service, they are also able to change the content and its format adaptively according to the timing and the context of interaction (Fan and Poole 2006).

On the web channel, web personalization is found to be a very useful approach to facilitate customers’ decision making (Tam and Ho 2006). If personalization is applied in m-commerce, many business opportunities are available. One popular application is mobile advertising. Mobile services providers use a short message service (SMS) to send messages to their subscribers. The message content can be an advertisement or a cash coupon. Consider the following scenario: mobile operators learn a user’s preferences with regular surveys, and cooperate with different business partners to offer trade

---

promotions. With GPS, mobile operators detect the location of a cell phone user. This user is fond of Chinese food (i.e. user’s preference). It is about lunch time (i.e. time and goal), and she is now near Chinatown (i.e. location). If the operators send her a mobile coupon (e.g. 10% discount in The Dragon Restaurant) in the form of SMS and she can enjoy the discount by simply showing this SMS to the waiter, then this can greatly increase the chance of her having lunch there.

Location-based services can be classified into various categories (Rao and Minakakis 2003). The first category is “where am I?” services. They relate to information about users’ locations and navigations. After detecting where service clients are, service providers can send information such as maps, driving directions and yellow-page listings to service clients. The “where am I?” services are also adopted by drivers. The second category is point-of-need information delivery. Service providers send product information and promotions to service clients. To personalize the information, the providers not only need to know the clients’ location, but also access their preference profiles. The example of The Dragon Restaurant belongs to this category. The third category is industrial/corporate applications, which are business-to-business oriented. Barcode scanners become the technology enablers for firms to track materials and product movements. Location-based services can be incorporated in the supply chain management and asset management, and location-specific task checklist and real time reports can be sent to workers. Table 1 presents a summary. The first and the second categories are the focus of this research.

<table>
<thead>
<tr>
<th>Category of Location-Based Services</th>
<th>Business Model</th>
<th>Information Used by Service Provider</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Where am I?”</td>
<td>B2C</td>
<td>Client’s location</td>
<td>Maps, navigation systems</td>
</tr>
<tr>
<td>Point of need info delivery</td>
<td>B2C</td>
<td>Client’s location &amp; preferences</td>
<td>Mobile coupons, shopping recommenders</td>
</tr>
<tr>
<td>Industrial / Corporate applications</td>
<td>B2B</td>
<td>Material’s location &amp; operational resource networks</td>
<td>Supply chain management, asset management systems</td>
</tr>
</tbody>
</table>

Table 1. Summary of Three Categories of Location-Based Services

1.2 Motivations and Research Questions

Mobile services providers adopt personalization technologies with the intention to better communicate with their users and to generate more business opportunities. The impact of personalization in m-commerce is under-investigated. Generally speaking, we believe personalization can increase the value of an organization by focusing on customer intimacy, but at the same time, it leads to users’ concerns.

With the gaining popularity of wireless handheld devices, this paper examines the use of personalization technologies in m-commerce. We are going to address the following research questions:

1. What business opportunities are offered by personalization in m-commerce?
2. What are major challenges to personalized mobile services?

The paper is organized as follows: In Section 2, we review the literature of personalization. Section 3 outlines the methodology, and Section 4 presents the focus groups findings. With the insights brought by the focus groups, a survey was conducted. Section 5 presents the survey and the findings. Section 6 discusses the findings and the last section concludes the paper.

2 LITERATURE REVIEW

The use of personalization in web-based applications has been widely examined. Several authors have found evidence for the effective influence of personalization on users’ decision making (Kumar and
Benbasat 2006; Tam and Ho 2006), as well as for perception of e-services (Nysveen and Pedersen 2004). By providing individualized content, offers and services, personalization eliminates aimless surfing activities (Shahabi and Banaei-Kashani 2003), eases business-to-consumer interaction (Ardissono et al. 2002) and increases users’ satisfaction (Liang et al. 2007). Also, personalizing content empowers firms to deliver customer value and to achieve profitable growth. It is reported that online retailers using personalization technology have had significant revenue increases.

While much prior work examines web personalization, the current work focuses on personalization in m-commerce. Personalization in m-commerce refers to the process of adapting content with the dual objectives of serving users’ needs as well as maximizing business opportunities. In general, it can take three forms:

(1) User-driven personalization – A user specifies in advance the desired layout and content that match his/her interests and preferences. For example, an investor would like content on a financial portal tailored to his own investment portfolio. The goal is to provide tools and options for users to customize the available information and the presentation format.

(2) Transaction-driven personalization – Similar to the previous case, customized layout and content are generated by the online merchant. In the transaction-driven context, personalization is driven by previous transactions rather than specified by the user in advance (Al-Natour et al. 2006; McGinty and Smyth 2006; Wei et al. 2008). For example, page organization and information indexing of mobile sites is automatically adjusted for different users depending on their preferences as inferred from previous transactions.

(3) Context-driven personalization – In this case, a very adaptive mechanism is employed to customize content and layout for each individual user. The mechanism is sensitive to the context of the interaction and users’ location, and adapts to the context in real-time. Advances in click analysis and text mining have made it possible to understand the context and to infer the processing objective of the user in real time. For example, context-sensitive persuasion messages in the form of cash coupons are dynamically generated and sent to users via SMS to realize cross-selling and up-selling opportunities.

Figure 1 summarizes different types of personalization technologies. The size of the bubbles reflects the usage in B2C commerce. While transaction-driven personalization is common at the current moment, we expect widespread adoption of context-driven personalization in the future, as mobile services providers can harness the power of GPS technologies, and take users’ location and contexts into consideration.

But still, there remains skepticism on the prospects of personalization. Several personalization initiatives have failed without generating any benefits to the adopting firms. One main reason for such failures is inappropriate resource allocation. Chellappa and Sin (2005) claim that investments in online personalization services may be severely undermined if online users do not use these services due to privacy concerns. Nantel and Sekhavat (2008) suggested that the effects of SMS advertising depend...
very much on the languages used in SMS and the receivers. This motivates us to conduct focus groups with genuine mobile users to explore business opportunities and possible challenges.

3 METHODOLOGY

We conducted a number of focus groups to explore mobile users’ concerns and expectations of personalized mobile services. They were dual-moderator focus groups. That is, there were two moderators. One moderator ensured the session progressed smoothly, while another ensured that all the topics were covered. The reason for adopting this approach is that focus groups are likely to bring to light the collective views of mobile users and are effective in producing richer data because of interaction (Bryman 2001; Cooper and Schindler 2008 p. 177).

3.1 Participants and Research Design

We conducted six focus group sessions comprising 48 students from a university in Australia. We considered students to be appropriate participants, because they were typical mobile services users. The sessions were held in the university during the second week of October 2007. Sessions were recorded and transcribed to produce qualitative data for subsequent analysis. In this study, we followed the general guidelines of focus group research closely (Fern 2001). Each session comprised four to six students.

The participants we selected were all mobile phone users. There were 29 females and 19 males. Their average age was 21. On average, they had subscribed to mobile services for at least five years. Thus, they had reasonable knowledge of mobile services. A moderator who had performed a similar role in previous research ran the focus groups. The main role of the moderator was to facilitate useful and relevant discussions among group members. We encouraged participants to freely express their views since their identities would remain anonymous throughout the data collection and analysis processes. The questions were open-ended. The focus group questions included “What is mobile-commerce?”, “In your opinion, what are personalized mobile services?”, “Why do you use personalized mobile services?”, and “What possible reasons may discourage you to use personalized mobile services?”. For example, we asked students their perceptions of mobile services, whether they would use personalized mobile services, and their concerns. Each focus group lasted approximately 1.5 hours.

3.2 Data Coding and Analysis

After the data were recorded, they were transcribed immediately to allow for coding and analysis. The data analysis involved rigorous examination of extensive focus groups transcripts. We compared, conceptualized and categorized our data. Coding categories reflected the interpretations that focus group participants formed about m-commerce. We eventually arrived at ten fundamental attributes. The data suggested most participants we talked to were curious about and had high expectations for personalized mobile services, however, they expressed strong privacy and security concerns. The concepts and patterns observed were linked to the IS literature. We endeavored to provide a data-theory link (Klein and Myers 1999) and attempted to understand participants’ perceptions of personalized mobile services that are associated with the theory. In order to reduce researchers’ bias and also to validate that no important attributes had been missed in the result summaries, a colleague was asked to comment on the analysis of the data. This allowed the incorporation of two different perspectives for our results and minimized the potential that we might have overlooked something important. The role of this colleague was to “bring a different and possibly more objective eye to the evidence” (Eisenhardt 1989 p. 538). The information this colleague received did not include the field researcher’s list of attributes. We went through the focus group transcripts several times and moved back and forth between the data, IS literature and the concepts emerged.
4 FINDINGS

In the following, we will present the conceptualization and categorization of the data in our final round. There are ten major categories.

4.1 Technological Issues

This category focuses on the technological aspects of personalized mobile services. It covers performance of personalization and mobile interfaces.

4.1.1 Location-Related Functions

Most focus group participants suggested that the major use of personalization was to meet their instant needs. Personalized mobile services act as a recommender, and this aligns with personalization literature (e.g. Gretzel and Fesenmaier 2006). Weather report, stock prices, restaurant recommendation and movie listings are typical examples of instant services. For instance, one participant used the mobile devices to “check the weather forecast”. Another participant mentioned, “I got an SMS from restaurants, stock exchanges and other service providers. Sometimes reading them is quite interesting and it’s always good to get some valuable information for free.” However, the participant highlighted that, “I don’t want them to send me junk SMS everyday.” Mobile services are useful for impulse purchase as well. One participant said, “when you go to the starbucks in the US, and if you like the music playing there, you can download it to your mobile phone”. Participants also tended to focus on hedonic activities, such as “discount sales” and “social events”.

Traffic updates are considered to be very useful mobile services. For instance, one participant mentioned that “maybe personalized services can be location based in such a way that tells us where the nearest petrol station is and also list out the one that’s offering the best price”. Another said that “I use it for traffic updates back home in Malaysia and it is a very helpful service”. The main advantage of using location-based mobile services was “convenience”. One participant said, “If I get the level of personalization it’ll be good and it’ll save me a lot of time.”

4.1.2 Personality-Related Functions

Focus group participants gave suggestions on location-based mobile services. For instance, a few participants suggested personalized mobile services could take more variables on customers’ characteristics, in particular personality traits, into consideration. One said, “SMS advertising should be related to my character and it should be things that match my personality, but in order for them to know that, we must provide them with that information”. And another participant mentioned that “I know some other websites that take personalization to the next level and actually study your personality and base their advertisements on your past preferences”. This echoes the findings from the work by Moon (2002) who empirically showed that a match between users’ personality and personalization strategies improves the effectiveness of the strategies.

participants pointed out that emotions and moods are important variables for individuals to make decisions on purchases. One participant said, “Maybe if they sent me a listing of movies that is in some language I don’t understand. But ultimately, it depends on my mood if I want to watch it or not.” This indicates that users’ preferences and peripheral user parameters (e.g. spoken language) may not be the only factor to be considered in personalization, and users’ moods and emotions, which are fuzzier, are useful in personalization. However, some participants thought that it was not easy to detect individuals’ emotions and moods with mobile devices.
4.1.3 Technological Limitations

(a) Feasibility of Predicting What Users Want

Although there is a significant amount of “hype” about what sort of personalized mobile services will be useful, most participants were not optimistic about the feasibility of location-based personalization. Some participants did “not like to be categorized and put into a frame”. Some emphasized that individuals’ preferences were changing constantly, and thus, hard to predict. For instance, one participant said, “I think that personalization takes too much time and my preferences can change”, and another participant mentioned that “some people don’t exactly know what they want or what they want to do. So they cannot give accurate information”.

(b) Small Screens and Limited Display

One major limitation of wireless handheld devices was small screens. One participant said that “the screens are too small to surf the internet”. Another participant mentioned that “some websites are not designed to fit on the mobile screen even if it’s a PDA”. This aligns with the suggestions by a participant, who said that “I don’t like phone advertising mainly because it’s so small and I prefer pictures, color and motion. SMS is so boring. Advertising is creative and so should be the advertisements”. Although the screens are small, participants expected more information to be packed on the screen. For example, one participant expected a rating system. That is, there are “a green-light or a red-light to show the importance [of the recommendations]”, and this type of color code is expected to “be standardized among the service providers”. This dilemma aligns with an IS research stream which explores increasing the usability of small screens (e.g. Kumar et al. 2004; Lee and Benbasat 2004).

(c) Security Issues

The increasing attacks on mobile devices by malicious hackers heighten individuals’ awareness of security. Individuals are concerned about how data are stored at mobile devices and how data are transferred across the wireless channels. Participants were concerned about “identity theft”, which they considered to be “a major issue right now” and “viruses that [might] be downloaded into the mobile phone”. One participant questioned, “I would like to know that all the information that the service providers have about me, is it safe?”

Security concerns discouraged individuals from using mobile services in general. For instance, one participant said, “Occasionally when I want to share a picture I send an MMS or sometime use the Bluetooth transfer. But I keep it off most of the time because hackers are always on the look out. So I keep it off for security reasons.” Another participant mentioned that “I wouldn’t use services like gambling because I don’t know what the source of the message is. One participant said, “there is always an issue if these service providers sell your information to third parties. It is illegal but it still does happen on a very large scale. Sometimes employees from within the company who have access to all this information can sell it illegally.”

People have serious concerns about the security of mobile devices, because mobile devices play a central role in daily life. One participant said that, “if someone hacks into my mobile phone or if I lose my mobile phone, they get details about my entire life.”

4.2 Social Issues

4.2.1 Opt-Out Options

Participants showed a considerable amount of concern about how these personalized mobile services should be regulated. For instance, giving users an option to opt-out is important. One participant questioned, “Will I be able to de-activate the service?” Since mobile communication plays a
significant role in our daily lives, it is not easy for customers to switch off the phone to stay away from spam. Hence, participants highlighted that what users could do is to “just turn off the service not the phone”.

4.2.2 Privacy Concerns

Individuals are concerned about their privacy. Most participants did not want to be tracked, and considered that personalization technologies may “abuse” their personal data. This aligns with the findings from prior IS work (e.g. Awad and Krishnan 2006). Some participants suggested that “the thought of the service provider being able to track your exact location is kind of scary.” They preferred simpler services to make their lives private. One said, “I think I’d rather get an incomplete service and keep my life private.” Another mentioned that “I think services should be personalized to some degree but it should not go over the limit where privacy issues are violated. Maybe it should just stay as it is right now. Just give me the information and I will decide what I want at that time.” Some even questioned, “Wouldn’t that be a bit dodgy if they knew where you are at all times?” Another said, “If that got really specific and if it could actually predict exactly what I want and what I like, I think it’ll be kind of spooky”.

4.2.3 Spam

Users dislike spam. Some participants got annoyed “when getting five to six of the same messages”. Some participants indicated that they simply “ignore all those kinds of SMS advertisements”. One participant said, “I heard that they get your number and send start sending you spam.” Users can subscribe to personalized mobile services free of charge; however, if they find that there is spam, users cannot stop the spam unless they pay a fee to the services providers. Hence, users are worried about the “hidden charge”.

4.3 Economical Issues

Price is a major concern. The participants tended not to use “mobile services that cost extra money”. One participant thought that personalized mobile services might incur extra costs, and hence, “it’s a bit too expensive”. Other participants said that personalized mobile services are “cool but expensive”. Some participants said that “if the price wasn’t a major factor, I think it definitely draws attention”. Users are uncertain of the prices, and their worry about being over-charged is another reason for them to reject mobile services. One participant said, “the most important thing is I don’t know how much it costs to send the message.”

If the services are provided by the Internet and by mobile services, the services on the Internet are usually free of charge. Hence, the Internet is a major threat to personalized mobile services. One participant said that “I remember when the soccer world cup was on, you could get the highlights each day on the mobile phone and that cost $60 a month which is quite expensive. You can do the same thing online for free.” Participants tended to “subscribe to the advertisements on a website from a company”, rather than receive personalized mobile messages. Another participant said, “Well, Google maps is slower compared to the GPS. But if it is a cheaper way then I’ll use the cheaper way”.

4.4 Firm-Customer Relational Issues

Participants did not trust mobile services providers, and they thought that personalized content presented by mobile services provider is “not neutral”. They considered “profit-making” to be the major objective of mobile services providers. One participant said, “I don’t know if I can ever trust a service provider because at the end all that they want is profit. Maybe if all this was supervised by the government, then there would be some sense of security”. Most participants considered that mobile services providers might bias the information. One participant said, “I am mainly concerned if the
personalized information I receive is actually neutral and not biased by the advertiser”. Another said, “I think that some advertisers make things that you don’t like sound like things that you do like.” This aligns with the findings that perceived “honesty” of services is crucial (Priester and Petty 1995).

One participant pointed out some tricks played by mobile service providers. He said, “when I’m at home, I use the free SMS service from a website. But there are some dodgy ones that keep sending you advertisements once you sign up with them. And to stop the advertisements, you need to send them an SMS which costs you $10.” These tricks, undoubtedly, arouse the concerns of individuals. These interesting findings echo prior research (e.g. Komiak et al. 2005; Komiak and Benbasat 2006) who emphasized the role of trust in a personalization agent in B2C commerce.

Some participants suggested that a contract on how mobile service providers use customers’ data is important. This helps customers develop a trusting relationship with mobile service providers. For instance, one participant said, “Well, I guess there would be some contract which says that they wouldn’t use the information for any other purposes.” Another participant pointed out, “I think companies must sign a confidentiality contract with the customers so that customers can trust the service provider. And once the contract is signed, the customer can give all the information that the provider requires.” But a participant also suggested that customers had the responsibility to “read the terms and conditions when they sign up for a service.”

5 A SURVEY

The findings of the focus groups illustrated that individuals were concerned about their privacy. Some of them were unwilling to be profiled and tracked. This motivates us to conduct a survey to gain a better understanding of what personal data they are willing to share with mobile service providers in return for personalized mobile services.

A self-administered online survey was conducted. Apart from demographic questions, the questionnaire consisted of 26 questions referring to 26 personal information attributes. We extracted these 26 information attributes2 based on the focus group findings. Subjects were asked to assess seven aspects of information nature. For each piece of information, we provided the participants with seven statements, and asked them to evaluate each of the 26 attributes on a 7-point Likert scale (1 = strongly disagree; 7 = strongly agree). The seven statements were: (1) this information is necessary for mobile personalization; (2) this information can improve the personalization quality, (3) this information reflects the uniqueness of a customer, (4) it is necessary for mobile services providers to get an updated information from customers regularly; (5) sharing this information will increase mobile users’ risks, (6) sharing this information will infringe on customers’ privacy; and (7) customers are willing to share this information with mobile services providers. The time required to complete the whole questionnaire was 20 minutes.

We posted the survey in five online forums for three weeks. The participation was voluntary and anonymous. We received 81 responses. There were 46 females and 35 males. Their average age was 29. We calculated the average score for each aspect of each personal information attribute. With the average score, we conducted a multi-dimensional scale (MDS) plot (Figure 2).

---

2 The 26 attributes are name, nickname, age, gender, date of birth, monthly salary, current location, preferred spoken/written language, preferences for food, preferences for music, previous purchases, home address, home activities, working address, working time, professional field, countries you plan to travel to, hobby, identity document details, financial details, income, leisure time, best friend’s name, best friends’ dates of birth, family members’ names, and family members’ dates of birth.
We arrived at a 2-dimensional plot. By analyzing the attribute pattern, dimension 1 (i.e. x-axis) referred to the importance of the attributes. For instance, hobby (d) and previous purchases (h) were found to be important attributes for mobile personalization. Family members’ names (m), the best friend’s name (l) and working time (k) were found to be less important in the personalization process. The x-axis dimension gives insights to mobile services providers on what questions they should ask their customer to collect rich and useful information for personalization.

Dimension 2 (i.e. y-axis) referred to the inverse of risk and privacy concerns. Sharing users’ attributes, such as education background (c) and preference for food (v), with mobile services providers were found to be less “risky”. There was a higher concern if users were asked to share information such as current location (f) and family members’ names (m). The y-axis dimension tells mobile services providers that mobile customers may be hesitated to answer some questions from the providers.

The interesting part is the top-right quadrant. Attributes falling into this quadrant are useful for personalization, but users tend not to answer questions on these attributes because of privacy concerns. Mobile services providers can explore personalization methods, e.g. collaborative filtering or data mining, to draw an inference on these attributes, rather than to ask users directly.

6 DISCUSSION

While prior research has shown that personalization is effective in influencing users’ information processing in the Internet, our understanding of its effectiveness in m-commerce is far from conclusive. This research aims to bridge the gap between the potential growth of mobile personalization and the lack of understanding of mobile users’ expectations. No prior IS research, of which the investigators are aware, examines this issue. This research conducts focus groups with mobile users, and our results can also provide implications to firms on how to leverage mobile personalized services to improve their service quality, resulting in higher customer satisfaction.

First, it is a pioneering effort to examine the location dimension in personalization research. Personalization of IT services intends to provide the right content in the right format to the right person at the right time in the right location. Recent research on personalization has focused on content personalization (e.g. Tam and Ho 2006) and adaptive interfaces (e.g. Billsus et al. 2002). Little work has been done on the location dimension. Since the success of personalization in m-commerce hinges on an understanding of the context of interaction which is location sensitive, it is imperative to understand whether individuals are willing to share location-related information with mobile services providers and what potential location-based personalized services can be offered. Together these dimensions (user preferences, content, layout, and location) capture many of the functionalities of
personalization and user characteristics at a particular instant. This represents a formal characterization of the notion of personalization.

Second, while personalization has been shown to be effective in influencing user behavior, its use should be balanced by taking a proactive approach to protecting data privacy on the user side. Our findings provide a better picture of users’ privacy concerns. Prior IS work (e.g., Awad and Krishnan 2006; Chellappa and Sin 2005) confirmed that users are concerned about their privacy being compromised by the personalization process. However, there is little information which mobile users are willing or unwilling to share with the services providers. Our work fills the gap by identifying 26 pieces of information that mobile users are concerned about. We conducted a survey and asked mobile users to comment on these 26 pieces of information. With an MDS plot, we identified some patterns. Mobile services providers should ensure fair access to individuals’ personally identifiable information and should provide a mechanism for users to change inaccurate or unauthorized personal information conveniently and quickly. Acceptable criteria for “opt-in” and “opt-out” should be provided as consumer privacy options. Our MDS plots provide insights for mobile services providers on mobile users’ focus of the mechanism implementation.

7 CONCLUSION

The role of personalization in m-commerce is gaining significance. Conducting focus groups with mobile users allowed us to gain a deeper understanding of their expectations and concerns for personalized mobile services. On one hand, users look for accurate, high-quality personalized services. On the other hand, users are concerned about security, privacy and spam. Based on the focus groups results, we conducted a survey and asked mobile users which personal information they are willing to share with mobile services providers. Our results provide mobile services providers with valuable knowledge on how to strike a balance between effectively leveraging location-detection features and easing users’ privacy concerns.

References


THE IMPACT OF CULTURE ON MOBILE PHONE PURCHASING:
A COMPARISON BETWEEN THAI AND BRITISH CONSUMERS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0515.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Culture fit / differences / heritage / intelligence / issues / theory / values, Cross-cultural issues, Customer satisfaction / service, Mobile communications</td>
</tr>
</tbody>
</table>
THE IMPACT OF CULTURE ON MOBILE PHONE PURCHASING: A COMPARISON BETWEEN THAI AND BRITISH CONSUMERS

Srikes, Monthathip, School of Management, University of Surrey, Guildford, UK, 
m.srikes@surrey.ac.uk
Louvieris, Panos, School of Management, University of Surrey, Guildford, UK, 
panos.louvieris@surrey.ac.uk
Collins, Catherine, School of Management, University of Surrey, Guildford, UK, 
c.collins@surrey.ac.uk

Abstract

The aim of this research is to determine and evaluate whether differences in national culture impact on the buyer behaviour of Thai and British consumers when purchasing a mobile phone. Furthermore, the research was conducted in order to identify and compare key cultural attributes that influence mobile phone purchasing between Thai and British consumers. An empirical study was based on the concept of Hofstede’s dimension of Individualism /Collectivism and Power Distance and Schwartz’s values dimension of Power, Achievement, Hedonism and Self-Direction. The data was collected from 140 questionnaires using students at the University of Surrey. The findings indicated that there is a significant difference between Thai and British consumers in terms of mobile phone purchasing behaviour as far as Hofstede’s cultural dimensions and Schwartz’s cultural values are concerned. The findings also recommend that managers in mobile phone organizations should be concerned with the cultural dynamics of consumers as part of their ongoing re-segmentation, communication and promotion strategies within their overall marketing strategies. Additionally, the cultural factors will assist managers to guide the specifications required for the development of online customer decision support systems.

Keywords: Impact of Culture, Mobile Phone, Culture

1 INTRODUCTION AND BACKGROUND

Mobile phones have become an integral part of human daily life and personal communication across the globe. By the end of 2007, there were approximately 3.3 billion mobile phone users worldwide which is equivalent to a penetration rate of 49% of the last year (International Telecommunication Union, 2008). Thailand and the UK were chosen for this research for several reasons. According to the Ministry of Information and Communication Technology (2008), the mobile phone penetration rate in Thailand increased its growth rate to over 50% in 2006 in comparison to 35% in 2005. In the UK, Ofcom (2007) reported that the mobile phone market grew by 41.3% between 2003 and 2007. Within this competitive market, it is essential for mobile phone companies to better understand purchasing behaviour to enable them to acquire new customers and retain existing ones. Blackwell et al., (2001) demonstrated that culture has a profound influence on ‘how’ and ‘why’ consumers purchase a range of products and services. Furthermore, Foxall et al., (1994) stated that the consumer’s motivation of product and service choices as well as lifestyle could be shaped by cultural dimensions. As a consequence, culture can influence an individual’s interaction with a product and ultimately the purchase. The cross – cultural comparison of mobile phone purchasing behaviour between the Thai
and UK markets will provide an insight to the overall East and West cultural divide. Therefore, it is imperative that cultural attributes need to be taken into consideration for marketing managers when investigating mobile phone purchasing behaviour.

The objectives of this research were to: (i) Determine and evaluate whether differences in national culture impact on the buyer behaviour of Thai and British consumers when purchasing a mobile phone; (ii) Identify and compare key cultural attributes that influence mobile phone purchasing between Thai and British consumers; (iii) Develop a framework for determining the cultural information requirements of a customer decision support system that will assist a marketing manager when addressing a culture sensitive market place; and (iii) Evaluate the management implications of the above objectives. In the following sections a definition and brief review of culture is presented. Secondly, Hofstede's cultural dimensions and Schwartz’s cultural values are reviewed and their relevance to the research is presented herein.

2 MOBILE PHONES

Srivastava (2005) stated that the mobile phone has shifted from being a ‘technological object’ to a key ‘social object’ as communication with others is the main purpose for mobile phone purchasing. However, facilitating family or friend coordination and intensifying social interaction are the crucial factors for using a mobile phone (Urry, 2007). According to Castells et al. (2007, p.85), “obtaining a mobile phone is a milestone that indicates success, not only financially but also culturally in term of the integration within society”. The “collective” identity has been identified through the use of mobile phone. Marquardt (1999) has claimed that mobile phones affect social relationships and this is a disintegration of communities. Mobile phone usage has resulted in greater electronic interactions between friends and family at the expense of face to face interaction which have been dramatically reduced. Consequently, it could be proposed that mobile phones are changing individual cultural norms and values (Rauch, 2005).

3 LITERATURE REVIEW

3.1 The concept of culture

There are numerous definitions of culture but for the purpose of this paper, culture is identified as the “collective mental programming” of people in an environment (Hofstede, 1980). Hall (1976) stated that culture is not genetically inherited, and cannot exist on its own, but is always shared by members of a society and is identified as a societal level construct. However, it certainly has implications for individual behaviour (Hofstede, 2001). Hofstede (1997) has also stated that culture influences an individual’s behaviour through the manifestations of values, heroes, rituals and symbols. Hence, an individual’s behavior is a result of that individual’s cultural value system for a particular context which are changed and developed over time (Luna and Gupta, 2001).

3.2 Hofstede’s Cultural Dimensions and Schwartz’s Cultural Values

The description and analysis of Hofstede’s and Schwartz’s theory are provided in this section. Hofstede (2001) stated that the cultural values research conducted by Schwartz (1994) was more appropriate to use in a cross cultural research project. However, Smith et al., (2002) argue that Hofstede’s cultural dimension remains the significant framework to apply in international research. Therefore, a study that combined Hofstede’s dimension and Schwartz’s theory was used in order to be more valid.
3.2.1 Hofstede Cultural Dimensions

In this research, the “power distance” and “individualism” dimensions are considered. As these two dimensions have been found to be valid across several other studies (Blodgett et al., 2001; Gregory and Munch 1997). Hofstede (2001) found that “individualism” and “power distance” where the two main attributes that characterized the difference between Thai and UK cultures. “High power distance” cultures (Thai) tend to be “low on individualism”, whereas “low power distance” societies (UK) tend to be “high on individualism”.

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power Distance</td>
<td>The degree of inequality among people within a society.</td>
</tr>
<tr>
<td>Uncertainty Avoidance</td>
<td>The member of a cultural feel endangered by uncertain, ambiguous, risk or undefined situations</td>
</tr>
<tr>
<td>Masculinity vs. Femininity</td>
<td>The sex role characteristics or attitude or norm or perception</td>
</tr>
<tr>
<td>Long – Term Orientation</td>
<td>The extent to which a society exhibits a pragmatic, future-oriented perspective rather than a conventional historic or short-term perspective</td>
</tr>
</tbody>
</table>

*Table 1. Hofstede Cultural Dimensions*

**Power Distance**

“Power Distance” is the extent to which people accept that power is distributed unequally, and is related to conservatism and maintaining status (Yeniyurt and Townsend, 2003). In high power distance societies, the individuals are associated with acceptance of one’s positioning society (Hofstede, 1984) and a social value exists where everyone has his or her own rightful place in the society (Morsini, 1998). On the contrary, the low power distance societies are much more concerned about society values independence and competition. The powerful members seek to look less powerful and they believed that there are an equal rights and opportunities for everyone (Greg et al., 1995). The research form Hofstede (1980) has illustrated that Thai societies are considered as a “high power distance” society which is relatively high in comparison to the UK society. On the other hand, the UK is implied as “low power distance”. Thus, relative to the UK; Thai culture is more acceptable of societal inequities.

**Individualism / Collectivism**

According to Hofstede (2001), “Individualism” refers to the society where the ties between individuals are very loose. In contrast, “collectivism” is defined as a society where individuals are integrated into strong and cohesive in-groups. In individualist cultures, people tend to be motivated by their own preferences, needs and rights in order to achieve their personal goals (Lee and Kacen, 2008). On the side of “collectivist culture”, societies have a significant attitude toward building long-term relationships and the role of trust. Members of societies are often motivated by duties and norms of societies (Usunier, 2000). Triandis (2004) also demonstrated that collectivist societies” are more concerned with ‘interpersonal relationship’ than an individualist culture. With reference to Hofstede’s work (1980), the UK scored high in individualism which is relatively high compared to Thai society. Thus, it can be implied that the UK societies can be defined as “Individualism”, whereas Thai societies are considered as “Collectivism” ones. Hence, it is possible that customer social values and reference groups have a greater degree of influence in Thai societies than the UK society when purchasing a mobile phone and this is a point that deserves further investigation.
3.2.2 Schwartz’s Cultural Values

Schwartz’s value theory is primarily concerned with the basic values of individual recognized across culture (Schwartz, 1992). There are ten key cultural values that were defined by the motivational goal it serves namely; Power, Achievement, Hedonism, Stimulation, Self-Direction, Universalism, Benevolence, Tradition, Conformity and Security. In accordance with Lee et al., (2002) and Ros et al., (1999), a more parsimonious of Schwartz’s cultural value are employed which fit in well with the social used mobile phone (See Table 2).

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power</td>
<td>Status and prestige, control or dominance over people and resources</td>
</tr>
<tr>
<td></td>
<td>(authority, social power, wealth, preserving the public image)</td>
</tr>
<tr>
<td>Achievement</td>
<td>Personal success through demonstrated competence according to social standards</td>
</tr>
<tr>
<td></td>
<td>(sense of accomplishment, successful, ambitious, capable)</td>
</tr>
<tr>
<td>Hedonism</td>
<td>Pleasure or sensuous gratification for one self /Self –Satisfaction</td>
</tr>
<tr>
<td>Self-Direction</td>
<td>Independent through and action –choosing, creating and exploring</td>
</tr>
<tr>
<td></td>
<td>(creativity, independent, imaginative, intellectual, logical)</td>
</tr>
</tbody>
</table>


Table 2. Schwartz’s cultural values

3.3 Limitations of Hofstede’s Cultural Dimensions and Schwartz’s Cultural Values

There are many researchers who have argued that Hofstede’s work may not be absolutely correct (Koveo and Tang, 2008; Smith et al., 2002 and Shenkar, 2001). First of all, Hofstede’s framework has become outdated as the data is forty years old. Kirkman et al., (2006) stated that Hofstede’s dimensions fails to capture the change of culture over time in which viewed culture as “static”. Indeed, Hofstede’s theory appears to perceive time as linear and ignores the profound influence of the substantial modernization such as travel, media and technology (Usunier and Lee, 2005). Also, Hofstede’s principle assigns to be “standard theory resistance when new work on culture distinction is substantiated (Kock et al., 2008). Additionally, Yoo et al., (2002) noted that Hofstede’s model lacks individual level analysis and maintain that Hofstede’s scores were calculated by total scores of the country thus, ignoring the individual difference within cultures. Finally, Hofsteded’s theory is assumed to be homogeneous and devoid of subcultures as the data was collected using only a single organization (David et al., 2008). However, as most of the dimensions are independent, Beckmann et al (2007) argued that Hofstede’s dimensions led to useful explanations of cross-cultural differences in consumer behaviour.

Schwartz’s Cultural Values

There are significant drawbacks associated with Schwartz’s cultural values that should be considered. The first limitation is caused by the obsolete information as the data was gathered fourteen years ago (Schwartz, 1999). Secondly, Schwartz’s cultural values scale focused on the research of an initial group set of basic human values (Schwartz 1994, 1999), which ignored the fact that individual actions are complex, reflexive and contingent on the context (Burroughs and Rindfleisch, 2002) which may create value conflicts. Nevertheless, from a psychology perspective, they have been shown to be valid and sufficient (Marcus and Baumgartner, 2004).

3.4 Buyer Behaviour Factors

Research conducted by Kimberly et al., (1995) using the Hofstede theory, indicated that cultural differences are seen as especially important for consumers’ choice of products and services. Roth (1995) discovered that the services that place emphasis on variety and hedonistic experiences can
generate value to an individualist society. Also, Strabub et al., (1997) found that high power distance and collectivist societies would reject the communication media which do not support the social pressures. According to Heine and Lehman (1997), self concept of independent corresponds to the cultural concept of individualism, whereby people express themselves as inherently separate and distinct. Conversely, the cultural concept of collectivism is related with the interdependent self concept which concerned on contextual, relational, and socially situated. Further, Lee and Kacen (2008) discovered that subjective cultures tend to influence the buying intention of consumers. The study from Choi and Geistfeld (2004) showed that functionality design, feature images and brand images are highly positive correlated with cultural characteristics of the users. Whilst, Page (2005) stated that promotional appeals have played an important role for international business practice.

4 HYPOTHESIZED MODEL

The main variables have been identified and are presented as a hypothesized model in Figure 1. The research has investigated the strong correlation between cultural attributes and buyer behaviour attributes.

![Hypothesized Model of the main variables of cultural attributes & buyer behaviour](image)

Figure 1. Hypothesized Model of the main variables of cultural attributes & buyer behaviour

5 RESEARCH METHODOLOGY

In order to achieve the objectives, the hypothesis formulation, data collection method, sampling method, sample size, the questionnaire design, pilot study, reliability and validity are presented below.

5.1 Hypothesis Formulation

Based on the research objectives and literature review, the hypotheses are stated below:

H a: There is no difference between Thai and British consumers when purchasing a mobile phone.

H b: There is no difference in Hofstede’s cultural dimension of Individualism /Collectivism between Thai and British consumers.

H c. There is no difference in Hofstede’s cultural dimension of power distance between Thai and British consumers.

H d: There is no difference between Thai and British consumers in term of Schwartz value.”(Power, Achievement, Hedonism and Self-direction)

H e. There is no relationship between Hofstede’s dimension and mobile phone choice criteria.

H f. There is no relationship between Schwartz value and mobile phone choice criteria.
### 5.2 Data Collection

A self-administered questionnaire was used to obtain the primary data and consisted of 11 questions divided into four sections. The objectives of the four sections were to: (i) collect demographic data and the mobile phone purchasing experience; (ii) identify whether Schwartz’s cultural values (Power, Achievement, Hedonism and Self-direction) had an impact on Thai and British consumers; (iii) examine to which extent values of “Individualism and Collectivism” and “Power Distance” influence Thai and British consumers; and (iii) identify the mobile phone buying decision factors and its relationship with cultural attributes.

A non-probability sampling method by using quota sampling was conducted in this research. The sampling units were both male and female students who had mobile phones. As students are the social economic groups and future business people, plus professional, therefore these students were representative for the population of British and Thai people. The sample frame for this research was the University of Surrey in UK. The questionnaires were measured on a five-point Likert Scale in accordance with cross cultural research conducted by Richardson and Smith (2007). In their research, the five-point Likert scales are adopted to examine the cultural constructs; high and low context, power distance and the media choice behaviour of two nationalities. Hence, the five point likert scale was suitable for this study. For the “Individualism/Collectivism” and “Power Distance” measurement, a score of less than 3 indicated that the respondent could be classified as “Individualist” and “High Power Distance”. Conversely, scores more than 3 illustrated that respondents could be classified as “Collectivism” and “Low Power Distance”.

### 5.3 Pilot Study

In order to determine the reliability of the research, a pilot study was undertaken to minimize the research error by testing the reliability of questionnaire. A pilot study was conducted for two weeks amongst 16 students at the University of Surrey. The variables used to develop the questionnaire are drawn from figure 1. The reliability of this research was measured using the Crobach alpha coefficient which indicates the level of inter-item consistency. The consistency alphas met the acceptable rate which was 0.723; thus, the scale used for this study is considered to be reliable. On the scale validity, the standard deviation (26.161) of the population was used to calculate the sample size by the formula: \( N = \frac{(z*SD)^2}{(E)^2} \) where \( N \) represents the minimum sample size, \( Z \) is the degree of confidence required, \( SD \) is the standard deviation of population and \( E \) represents the range of error around the sample estimate acceptable. Using the formula, the sample size was calculated using a 95% confidence interval which \( Z \) value equals 1.96, within plus or minus 5% of the population mean. Hence, the minimum sample size should be: \( N = (1.96 * 26.161)^2 / 86.6250*0.05)^2 = 140 \). Therefore, a minimum of 140 questionnaires should be issued in order to achieve reliability. The questionnaire was distributed to Thai students (70 sets) and for British (70 sets) students in the University of Surrey. After checking and collecting the final data, Independents- samples t test and Pearson Correlation Test were employed to further analyse the data.

### 6 ANALYSIS OF THE RESEARCH FINDINGS

The results of the first four hypotheses are shown in the following tables.

#### 6.1 Hypothesis Testing - Independent Sample-test

<table>
<thead>
<tr>
<th>Hypothesis Ha</th>
<th>British Mean</th>
<th>Thai Mean</th>
<th>Sig. 2-tailed</th>
<th>British Mean</th>
<th>Thai Mean</th>
<th>Sig. 2-tailed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buying Intention</td>
<td>3.56</td>
<td>3.96</td>
<td>0.003*</td>
<td>Price</td>
<td>3.79</td>
<td>3.84</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
Hypothesis Ha aimed to test whether differences in national culture impact on consumer buyer behaviour. The results indicated that “Buying Intention”, “Social Acceptance”, “Service”, “Brand Image”, “Promotion” and “Product Quality” show significant differences between Thai and British consumers (Sig.values <0.05). Thus, the null hypothesis is rejected; there is a significant difference between Thai and British consumers for these six variables. Table 3 shows that “Product Quality” and “Feature / Appearance Image” are the important elements for British consumers. However, the “Product Quality” “Promotion” and “Brand Image” are found as the important variables for the Thai consumers. Hypothesis Hb HC and Hd are aimed at testing whether any differences exist between Thai and British consumers in terms of Hofstede cultural dimensions and Schwartz’s cultural values.

Table 3. Independent Sample t –test of Hypothesis H_a

<table>
<thead>
<tr>
<th></th>
<th>Mean (British)</th>
<th>Mean (Thai)</th>
<th>Sig. t-test 2-tailed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Acceptance</td>
<td>2.53</td>
<td>3.62</td>
<td>0.000*</td>
</tr>
<tr>
<td>Service</td>
<td>3.57</td>
<td>3.86</td>
<td>0.053*</td>
</tr>
<tr>
<td>Brand Image</td>
<td>3.20</td>
<td>4.06</td>
<td>0.000*</td>
</tr>
<tr>
<td>Feature Image</td>
<td>4.01</td>
<td>4.04</td>
<td>0.843</td>
</tr>
<tr>
<td>Promotion</td>
<td>3.60</td>
<td>4.06</td>
<td>0.000*</td>
</tr>
<tr>
<td>Product Quality</td>
<td>4.61</td>
<td>4.09</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

Hypothesis Hb and Hc

<table>
<thead>
<tr>
<th>Hofstede Cultural Dimensions</th>
<th>Hypothesis Hb</th>
<th>Hypothesis Hc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power Distance (Hb)</td>
<td>Mean (British)</td>
<td>Mean (Thai)</td>
</tr>
<tr>
<td>Team Work</td>
<td>Prestige</td>
<td>2.91</td>
</tr>
<tr>
<td>Being Accepted</td>
<td>Impress Other people</td>
<td>2.97</td>
</tr>
<tr>
<td>Reference Group Influence</td>
<td>Successful</td>
<td>2.21</td>
</tr>
<tr>
<td>Group Opinion</td>
<td>Present Arguments</td>
<td>2.47</td>
</tr>
<tr>
<td>Family / Friend Discuss</td>
<td>Express Disagreement</td>
<td>2.76</td>
</tr>
<tr>
<td>Increase Interaction</td>
<td>Status</td>
<td>3.69</td>
</tr>
</tbody>
</table>

Table 4. Independent Sample t –test of Hypothesis Hb and Hc

The findings of Hypothesis Hb indicated that there is a difference between Thai and British consumers for five questions “Being accepted”, “Reference Group Influence”, “Family / Friend Discuss”, and “Increases Interaction” (Sig.values<0.05). Referring to the mean values (Table 4); there are some results that go against the assumption of Hofstede that Thais tend to be a collectivist culture. The calculated mean in the questions “Reference Group Influence” (2.9) and “Group Opinion” (2.7) indicates that Thai consumers responded as an “individualist society”. Conversely, Hofstede predicted that the British consumer is a “high individualism” one. The calculated means shows that the question of “Increase Interaction” has high scores (3.69). As the full statement of the question is “mobile phone has increased the frequency of interactions with family and friends. Hence, it is logic to explain that mobile phones tend to activate the collectivist attribute. For the Hypothesis Hc, the Sig. values for the four questions “Impress other people”, “Successful”, “Present Arguments” and “Status” are less than 0.05. Thus, the null hypothesis is rejected; there is a significant difference in the mean scores between Thai and British consumers.

Hypothesis Hd

<table>
<thead>
<tr>
<th>Schwartz Cultural Values</th>
<th>Mean (British)</th>
<th>Mean (Thai)</th>
<th>Sig. t-test (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power</td>
<td>3.27</td>
<td>3.76</td>
<td>0.002*</td>
</tr>
</tbody>
</table>
Table 5. Independent Sample t-test of Hypothesis of Hypothesis Hd

As shown in Table 5, in the three elements: “Power”, “Achievement” and “Self Direction”, the mean values of Thai respondents was greater than that of British ones. The Sig.-values of three variables: “Power”, “Achievement” and “Self –Direction” was less than 0.05. Therefore, the null hypothesis is rejected, implying a difference in Schwartz values in terms of Power, Achievement and Self-direction between Thai and British consumers. The results of hypotheses He and Hf are shown in the following tables.

6.2 Hypothesis Testing - Pearson Correlation Test

<table>
<thead>
<tr>
<th>Hypothesis He</th>
<th>Individualism</th>
<th>Power Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thai Factors</td>
<td>R</td>
<td>Thai Factors</td>
</tr>
<tr>
<td>Social Acceptance</td>
<td>0.733 Promotion</td>
<td>Service</td>
</tr>
<tr>
<td>Buying Intention</td>
<td>0.725 Service</td>
<td>0.313 Social Acceptance</td>
</tr>
<tr>
<td>Promotion</td>
<td>0.633 Social Acceptance</td>
<td>0.306 Feature Image</td>
</tr>
</tbody>
</table>

Table 6. Pearson Correlation Test of Mobile Phone Buying Decision factors that were highly positive correlated with cultural attributes of Individualism and Power Distance * (R= Pearson Correlation )

With reference to Table 6, all three variables: “Social Acceptance”, “Buying Intention” and “Promotion” are strongly correlated with the cultural attributes of “Individualism”. As the correlation coefficient R = 0.733, 0.725 and 0.633. Furthermore, there is a strong correlation with the cultural elements of “Power distance” which was observed through the “Service” and “Social Acceptance”, the correlation coefficient was r = 0.936 and 0.602, these outcomes demonstrate a strong relationship between two continuous variables. Thus, Thai consumers are extremely influenced by their culture in these four purchase decision factors. For British consumers, the results (Table 6) show that there is a strong relationship between “Promotions” and the cultural dimension of “Individualism” in which the correlation (r) was scored at 0.694. Moreover; there is a medium relationship between the “Social Acceptance”, “Brand Image”, “Product Quality” and “Power Distance” cultural dimension. Hence, it can be assumed that British consumers are more likely to be influenced by their cultural element in these four criteria.

<table>
<thead>
<tr>
<th>Hypothesis Hf</th>
<th>Thai Respondents</th>
<th>British Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Factors</td>
<td>R</td>
<td>Decision Factors</td>
</tr>
<tr>
<td>Feature/Appearance Image</td>
<td>0.535</td>
<td>Buying Intention</td>
</tr>
<tr>
<td>Social Acceptance</td>
<td>0.51</td>
<td>Price</td>
</tr>
</tbody>
</table>
The findings have indicated that Thai’s have strong correlation scores amongst “Feature/Appearance Image –Hedonism” and “Social Acceptance –Power”. Thus, Schwartz’s cultural values of “Hedonism” and “Power” are more influenced in these two buying decision variables. For British consumers, the strong correlation exists where the correlation is between “Buying Intention- Self-Direction” and “Price –Hedonism”. Hence British consumer’s cultural values in the Schwarz dimension are highly influenced by “Buying intention” and “Price”.

In this section, Consumer decision factors and cultural impact framework developed for the marketing of mobile phones in Thailand and the UK

Two frameworks were developed for marketing mobile phones in Thailand and the UK in relation to consumer decision factors and cultural impacts of consumers of these two cultures. Figure 2 is a diagrammatic representation of the results and findings that illustrate the key consumer decision factors and cultural attributes that influence Thai consumers when purchasing mobile phones. This framework (figure 2) can assist marketing managers dealing with Thailand’s mobile phone market. The findings also indicate that the “collectivism” cultural dimension is strongly correlated with “social acceptance”, “promotion” and “buying intention” variables (this is indicated by the + + + symbol on the diagram). Also, the strong correlation with “power distance” cultural dimension was observed through the “service” and “social acceptance” criterion. Furthermore, a strong correlation exists between “hedonism-feature/appearance image” and “power –social acceptance”. Thus, “collectivism”, “power distance”, “power” and “hedonism” represents important cultural attributes. The most important dimension has been found to be “collectivism” as this cultural element influences three variables (social acceptance, promotion and buying intention). Similarity, the results indicate that “social acceptance” factor is mainly influenced by “collectivism”, “power distance” and “hedonism” cultural elements of Thai consumers. Hence, managers have to be aware of “collectivism”, “power distance”, “power” and “hedonism” cultural dimension when dealing with Thai consumers purchasing mobile phones, especially “social acceptance”, “promotion”, “service”, “buying intention” and “feature/appearance image factors.

Figure 3 is a diagrammatic representation of the results and findings for UK consumers and indicates that “individualism”, “hedonism” and “self-direction” are the major important cultural attributes that influence mobile phone purchasing. The framework (Figure 3) shows that a strong correlation exists

---

| Brand Image | 0.484 | Product Quality | 0.433 |

*Table 7. Pearson Correlation Test of Mobile phone factors that were highly positive correlated with cultural attributes (Schwartz’s Cultural Values) * (R= Pearson Correlation)*

*Positive relationship (+), Negative relationship (-) 
Strong correlation $\rightarrow (+ + +, - - -)$, Moderate correlation $\rightarrow (+ +, - -)$, Weak correlation $\rightarrow (+, -)$

---
amongst “individualism -promotion”, “hedonism- price”, “self-direction – buying intention” and this is indicated by the + + + symbol on the diagram. Marketers should therefore consider the “individualism /collectivism”, “hedonism” and “self-direction” cultural dimensions carefully because they have a high impact on British consumers. Further, marketers need to take “promotion”, “price” and “buying intention” factors into consideration for British consumers alongside their cultural attributes.

7 CONCLUSIONS

The primary purpose of this research was to determine and evaluate whether differences in national culture impact on the buyer behaviour of Thai and British consumers when purchasing a mobile phone. By addressing the Hofstede’s cultural dimensions and Schwartz’s cultural values with respect to mobile phone buying decision factors, the findings indicate that there is a significant difference between Thai and British consumers. From the findings, “promotion” is the important element for both Thai and British nationalities as the results showed high correlation with the cultural attributes of “collectivism” and “individualism”. For Thai consumers, it is important to note that “social acceptance” acts as the main buying decision factor alongside their cultural attributes. The most important cultural dimension has been found to be “collectivism” as this cultural element influences three variables (See Figure 2). In contrast, a strong correlation occurred between “buying Intention- self-direction” and “price – hedonism” with British consumers. Furthermore, “individualism” is found as the one of the cultural values that has a high degree of influence of mobile phone’s price criterion (See figure 3). This knowledge may then be exploited for the development of a culturally informed customer interface design.

The implications of the findings are valuable for mobile phone marketing managers to understand better the cultural attributes of consumer behaviour when purchasing mobile phones. From the findings, it shown that Thai consumers purchase a mobile phone which enables them to have social connection and relates them to their peer groups. Indeed, Thai consumers are likely to be sensitive to the influence of their group orientation and reference groups such as families and friends. Triandis, (2004) stated that the concern of interpersonal relationships tends to be in collectivist societies rather than in individualist ones. Hence, it is advisable that promotion and communication strategies should send ‘collective’ messages which are group-oriented based and appeal to families. Conversely, British customers have a greater degree of influence and are a more individualist oriented culture as opposed to the Thai culture. According to Roth (1995), services that place emphasis on variety and hedonistic experiences are suitable in a high individualist culture. Thus, it is recommended that mobile phone companies offer a variety of mobile phones that can be personalized to the individual customer. Additionally, the research has shown that not all Thai’s behave as “collectivism” and not all British consumers are characterized as “individualists”. The model developed by Hofstede’s (1984), viewed culture as static, however the contribution that study provides is that culture is dynamic and can be interchangeable over time. An individual’s behaviour could be influenced and dominated by the dynamics of culture, technology, and in particular mobile phones. Therefore, it is recommended that marketing managers address these cultural dynamics as part of their ongoing re-segmentation, communication and promotion strategies.

An understanding of the different cultural dimensions on buyer behaviours will assist managers in the management of customer decision support systems. The findings of the research have shown how culture can inform the design and functions of the information system in order to create a more effective customer decision support system. Hence, managers should be able to tailor online communications and the design of the system to target the customer more effectively, which in turn will lead to improved customer relationship management (CRM). An appreciation of the cultural factors will provide managers with insights into how to develop marketing information systems, especially the promotional and communication strategies. In addition the results of the study suggests that collective messages which are group-based should be targeted towards the Thai culture, whereas
the individualist oriented messages such as personalized messages are more appropriate to the UK culture. The research instrument and analytical tools employed in this research will help marketing managers to track culture changes in their chosen markets. Moreover, managers can use the frameworks developed in Figure 2 and Figure 3 to guide what information features should be included and emphasized in order to create an effective decision support system that takes culture into account and accommodates cultural differences within an international setting.

References


Hofstede, G and Hofstede, G.J (2005), Cultures and Organizations: Software of the Mind, New York, McGraw-Hill USA.


Hofstede, G.(1997), Cultures and Organizations: Software of the Mind, New York ,McGraw-Hill,


Proceedings ECIS 2009


Morosini, P. (1998), Managing cultural differences, Pergamon


MODELLING USE CONTINUANCE IN VIRTUAL WORLDS: THE CASE OF SECOND LIFE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0002.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Virtual world, Continuance, Partial Least Squares, Structural Equation Modeling</td>
</tr>
</tbody>
</table>
MODELLING USE CONTINUANCE IN VIRTUAL WORLDS:
THE CASE OF SECOND LIFE

Abstract

There are now many virtual worlds in existence, a number of which have built considerable user bases. However, there is little empirical evidence to suggest what factors underpin their continued usage and success. This study set about determining why users continue to use virtual worlds, using the example of the popular virtual world of Second Life. The study adopts continuance theory and extensions to the basic model to examine the effects of enjoyment, habit and absorption. The results (n=339) suggest that continuance intention is driven by perceived usefulness, habit, absorption and enjoyment, which together provide a comprehensive explanation for virtual world behaviour ($R^2=0.565$). Interestingly, satisfaction did not appear to play a direct strong role in determining intentions. The paper rounds off with conclusions and implications for future research and practice in this very new area of inquiry.

Keywords: virtual world; continuance; enjoyment; absorption; habit; Second Life.
1 INTRODUCTION

The notion of a ‘virtual world’ is a relatively novel phenomenon that emerged from the juxtaposition of recent advancements in computer graphics, online gaming and social networking technologies. A ‘virtual world’ may be defined as “an electronic environment that visually mimics complex physical spaces, where people can interact with each other and with virtual objects, and where people are represented by animated characters” (Bainbridge 2007, p. 472). Such ideas were, until relatively recently, restricted to rather crude graphical representations and enabled limited interactivity. However, virtual worlds have become very sophisticated and increasingly realistic – often integrating advanced 3-D rendering, in-world currencies, avatar and object customization and development, property ownership and permissions, text and voice communication, and social networking tools – and the number of virtual worlds has increased dramatically in the last five years. Gartner (2007) predict that 80 per cent of active Internet users will be members of at least one virtual world by 2012.

If we restrict our discussion to virtual worlds that are not purely massively multiplayer online games (Barnes & Mattsson 2008), i.e. they have a freeform element, then there were around 70 virtual worlds in existence or under development as of the third quarter of 2008 (KZero 2008a). 23 worlds aimed at the under 10s (the largest being Poptropica with 20 million users and Barbie Girls with 15 million), 14 aimed at the pre-teens (the largest being Habbo with 100 million users, Neopets with 45 million and Club Penguin with 19 million users), 14 aimed at the groups from 13 to 20 (including Stardoll with 21 million users, IMVU with 20 million users and Gaia with 15 million users), and 21 aimed at the over 20s (the largest being Second Life with 15 million users). Some of these are centred around a specific brand (e.g. Neopets, Stardolls, Disney Fairies, Virtual MTV, Barbie World, Fusion Fall (Cartoon Network), Hello Kitty, Club PonyPals and Lego Universe) while others are more general (e.g., Second Life, Multiverse, Active Worlds, There, Qwaq and Kaneva). Further, they vary considerably in their content focus (KZero 2008b), including content creation (Second Life, Active Worlds, HiFiHi and Multiverse), Socializing/chat (Utherverse, There and IMVU), fashion/lifestyle (Frenzoo, Stardoll and GoSupermodel), education (Whyville, Medikidz and Jumpstart), music and media (vMTV, Onverse and vSide) and sports (Football Superstars, Empire of Sports and Sportsblox).

The combined population of users of the virtual worlds discussed above is well in excess of 300 million. The virtual goods sector is estimated at more than $1.5 billion per annum (Wu 2007). Second Life, the best known and broadest virtual world platform, has grown rapidly from 2 million residents in January 2006 to nearly 15 million residents in August 2008 (Second Life 2008a). In August 2008, there were nearly 20 million transactions and more than US$338 million were spent in the year to June 2008 (Second Life 2008b). Residents owned 1.714 million square meters of land and the million or so residents that logged in July 2008 spent 34.7 million hours in Second Life, nearly nine times that of July 2006 (Second Life 2008c). Evidence suggests that there is a high level of continued usage of virtual worlds such as Second Life.

Academic research into virtual worlds is still at a very early stage. With the massive recent growth in virtual worlds and their users and the very high sustained level of usage of these virtual worlds, we were very interested to find out why this might me the case. Thus, the research question for this study was: “Why do users continue their usage of virtual worlds?” As a basis for the research, we use continuance theory and three related theories that are likely to drive continuance: habit (Limayem et al., 2007), absorption or ‘persistent flow’ (Csikszentmihalyi 1990, Schaufeli et al. 2002) and enjoyment (van der Heijden 2003). As the context for the investigation, we focus on the largest and best-known of the freeform adult virtual worlds, Second Life.

The structure of this paper is as follows. In the next section we examine the theoretical development of a model of continuance in virtual worlds. The third section describes the methodology used in the study. Section four presents the results of the study. In the final section, the paper rounds off with a discussion, limitations, implications for research and practice, and conclusions.
2 THEORY AND RESEARCH MODEL

In this section we pull together the various strands of theoretical literature that we use to create our research model. In particular, we examine the salient literature on the concepts of use continuance, enjoyment, habit and absorption. Each of these is now examined in turn.

2.1 Continuance theory

The roots of continuance theory, as applied to information systems research, are from the marketing literature. Expectation-confirmation theory (ECT) emerged from the consumer behaviour and services marketing literature and has proven broadly robust in a number of service contexts (Dabholkar et al. 2000, Oliver 1993). The general thrust of ECT is the assessment of post-purchase intentions, as influenced by initial expectations about a product or service, subsequent adoption and use (consumption) and the formation of perceptions about performance as influenced by the confirmation or not of initial expectations, the latter determining the level of satisfaction with a purchase and subsequent repurchase or use discontinuance. Bhattacherjee (2001) was the first to fully formalize the theory into an ex-post framework that could be applied to the domain of information systems, adapting the theory to be applied post-acceptance and to encapsulate perceived usefulness (Davis et al. 1989, Davis 1993) as a replacement construct for expectations. Perceived usefulness has consistently proven to be an important construct in longitudinal adoption to post-adoption behaviour (Davis et al. 1989, Karahanna et al. 1999). Thus Bhattacherjee’s (2001) model relates the constructs of perceived usefulness and satisfaction to the extent of confirmation of a user’s expectations about an IS, whereby expectations that are fulfilled drive greater satisfaction and perceived usefulness. High levels of perceived usefulness are also posited to lead to greater satisfaction with a system. In turn, the outcome variable of continuance intention is determined by the level of satisfaction with an IS and the perceived usefulness of the system.

Bhattacherjee’s model has been successfully applied to individual user contexts involving the Web, such as online banking, and the Internet more broadly (Bhattacherjee 2001, Limayem et al. 2007). Virtual worlds are also distributed systems reliant on Internet technology, and the focus of this study is on individual users; Bhattacherjee’s theory (which appears to have broader application and generalization in any case) is adopted as a suitable core for a model examining continuance behaviour in this context.

2.2 Enjoyment

The applications of virtual worlds for organizations are potentially very broad, including product branding, promotional events, product development and testing, collaboration, process rehearsal and simulation, employee recruitment, education and training, commerce, and communication. However, at the individual user or consumer level, which is the focus for this study, virtual worlds can be construed as hedonistic systems where users experience pleasure or fun when using the system. Such uses may include, for example, socializing, romantic encounters, shopping for desirable personal items, customizing one’s personal appearance, playing games, fantasy, the creation of an alter-ego (Hemp 2006) and other experiences that may for example contribute to building self-esteem, social bonding and self-actualization. To have a pleasurable experience, individuals often seek sensations on multiple sensory channels (van der Heijden 2003), and clearly virtual worlds are equipped to provide such multimedia experiences by design.

Unfortunately, the IS continuance theory described above is built primarily on the utilitarian paradigm of providing instrumental value to the user. The consumer behaviour literature has determined strong support for repurchase and use intentions driven by both utilitarian and hedonic value in products or services (Bauer et al. 2005). Recently, the extension of this hedonic value conceptualization, hereinafter referred to as enjoyment, to the user acceptance of hedonic information systems has been...
formalized and empirically tested by van der Heijden (2003). Early research in virtual worlds has also confirmed the relationship between enjoyment and behavioural intention (Wu et al. 2008). This study adopts this theoretical development and the related theoretical relationship between enjoyment and intention to continue using a hedonic information system (as clearly implied, although not longitudinally tested in van der Heijden’s study). Further, in the spirit of the expectation-confirmation theory, as discussed above, the inclusion of a hedonic construct (enjoyment) as well as a utilitarian one (perceived usefulness) has clear implications for the fulfilment of expectations. Here, we theorize that hedonic value expectations are manifested or not in the same way as utilitarian ones. Thus, a user will hold expectations about the hedonic value of a system, in other words its perceived enjoyment, which will be confirmed or not by the adoption and use of the system, thereby triggering an evaluative response in the form of satisfaction.

2.3 Habit

Previous research has found that use continuance in information systems and other contexts can be predicted by the extent to which a behaviour has become automatic because of prior learning – typically conceptualized as habit (Limayem et al. 2007). Research has variously examined habit as a moderator between intention and actual behaviour, as a direct effect on actual behaviour, and as an indirect effect on behaviour that primarily determines intentions. Our focus in this research is on use intentions rather than actual behaviour, and so naturally we focus on that latter of these formulations. Our focus on intentions is in line with a core body of previous IS literature (Legris et al. 2003), whilst the focus on indirect habit effects is a view that is held in a number of previous studies that have examined the effects of habit (and the much used proxy construct of experience) on behavioural intentions (Legris et al. 2003).

In addition to the effect of habit on intention to continue using an IS, we also posit that habit is significantly influenced by satisfaction. Limayem et al. (2007) in their comprehensive definition, application and analysis of the habit construct in continuance theory find very strong support for the linkage between satisfaction and habit. Further, they find that frequency of prior usage has a positive association with habit, an assertion that we also make for the purposes of model development.

The creation of habit requires a stable context conducive to its formation through repetition or practice (Orbell et al. 2001); we would hold that such a context exists when focusing on individuals’ behaviour with respect to a single system such as a virtual world. This position is in line with that of Limayem et al. (2007) in their study of habitual use of the Internet.

2.4 Absorption

Csikszentmihalyi (1990) examines experiences that are ends in themselves and are not tied to external considerations: within such experiences, an individual is caught up in the “flow” of an activity and subsequently absorbed into it (Wang & Calder 2006). This work emphasizes the role of pleasure and enjoyment in the flow experience. Indeed, Hoffman and Novak (1996) describe flow as an intrinsically enjoyable state, which others have found to be strongly related to feelings of pleasure and arousal (Csikzentmihalyi 1990, Wang et al. 2007). Further, flow theory suggests that users absorbed in a flow experience may tend to lose behavioural control and get locked into patterns of greater usage and habitual behaviour (Csikzentmihalyi 1990). Turning to the use context in this study, previous work on flow in the context of Internet use and online gaming has found a significant relationship between flow and use intentions (Hsu & Lu 2004).

Flow is purported to be a multidimensional construct that has variously been defined in terms of concepts such as concentration, control, challenge, attention focus, curiosity, enjoyment and intrinsic interest (Webster et al. 1993, Wang et al. 2007). However, the concept of flow has some issues worthy of note. As Schaufeli et al. (2002, p. 75) suggest: “… [flow is a] complex concept that includes many aspects and refers to rather particular, short-term ‘peak’ experiences instead of a more pervasive and
persistent state of mind, as is the case with engagement.” Since our aim is to measure more persistent states of mind that influence intentions to continue use of an information system it would seem logical to focus on the concept of engagement rather than traditional flow, or more particularly the dimension of engagement that represents persistent flow states – that of absorption. As Schaufeli et al. (2002) state in their application of the absorption construct to the work context:

... absorption is characterized by being fully concentrated and deeply engrossed in one’s work, whereby time passes quickly and one has difficulties with detaching oneself from work. Being fully absorbed in one’s work comes close to what has been called ‘flow’, a state of optimal experience that is characterized by focused attention, clear mind, mind and body unison, effortless concentration, complete control, loss of self-consciousness, distortion of time, and intrinsic enjoyment … (p. 75)

From a theory perspective, we posit that an individual’s intention to continue using a virtual world will be driven by the level of absorption, or persistent flow, experienced in the media environment. In turn, absorption will be determined both by the level of enjoyment experienced in the virtual world and the level of habitual behaviour that has been established.

The theoretical developments above were combined with the basic IS continuance model for the basis of this study. Together these theories provide an extension to the typically utilitarian approach of continuance theory, which may be limited in its application, to capture broader elements related to hedonistic use (enjoyment), automatic behaviour (habit) and persistent ‘flow’ (absorption), all of which are likely to impact on the use of virtual worlds. The full research model tested in the research is provided in Figure 1.

![Research model](image)

**Figure 1.** Research model.

3 STUDY DESIGN AND METHOD

The sections below describe in detail the data collection process, the measurements used, and the type of data analysis executed.

3.1 Data Collection

We chose the context of Second Life since it is the largest virtual world for the adult age group and the most well-known of the freeform virtual world environments. Evidence also suggests that it is the virtual world platform with the broadest user base and range of applications (KZero 2008c).
Data collection involved the use of two traffic ‘bots’ in Second Life operating at busy traffic points. Each bot is essentially an avatar automated to deliver a survey advertisement and a URL in text form in-world. Each bot had an advertisement for the survey in its group name, above the avatar. Details of the survey were also provided in the profile of the avatar and respondents were requested to IM (instant message) the bot. Respondents initiate contact and are given details of the survey and the URL to begin the survey. The survey was provided using QuestionPro. To collect sufficient responses, each bot was placed at a high-traffic location selected from Second Life’s ‘popular locations’ list. The locations were chosen to be as generic as possible (to appeal to both genders, different ages and nationalities) and each focused on providing both free and paid-for digital content and on generating traffic through paid ‘camping’ activities (where individuals are paid small amounts of money for time spent ‘sitting’ at a particular location).

A monetary incentive (of L$250 or approx. $0.95) was provided to respondents for each completed survey. A non-conditional incentive was used, since there is evidence that this is likely to improve response rates in social science research over conditional incentives such as a prize draw (Church 1993). Further, evidence suggests that incentives do not necessarily bias sample composition or data quality (Brennan 1992) and are more likely to attract harder to reach groups, by providing motivation (Groves et al. 2000). There is no guidance on the recommended value of incentives for virtual world surveys, but we felt that L$250 was favourable compared to market research surveys delivered through Second Life by HippiePay and other companies. This is much smaller in monetary terms than the typical $10-$20 provided in other modes of social science research (Simmons & Wilmot 2004). The survey ran for approximately six months. After filtering for duplicates and invalid responses we had a total of 339 survey responses.

3.2 Measurement

All of the constructs and scale items used in the research were adopted from previously validated sources. The core of the model and measures for confirmation, perceived usefulness, satisfaction, continuance intention, habit and frequency of prior usage are adapted or adopted from Bhattacherjee (2001) and Limayem et al. (2007). The semantic differential measure for enjoyment employed in the instrument is that of van der Heijden (2003). Finally, the measure for absorption is taken from the engagement inventory of Schaufeli et al. (2002).

3.3 Data Analysis

Data analysis was performed using a variance maximization approach to structural equation modelling (SEM) and associated statistics for validity and reliability. More specifically, we used the partial least squares (PLS) technique with reflective indicators in Smart-PLS 2.0 (Ringle et al. 2005). The PLS technique has become increasingly popular in information systems research, marketing and in management research more generally in the last decade or so, influenced by its flexibility; indeed, PLS does not have the same distributional assumptions of normality for data, is able to handle small- to medium-sized samples (Chin 1998).

4 RESULTS

4.1 Respondent Characteristics

Of the 339 responses, approximately three-fifths were male (58.4%). The median age was 25-34 years, with only a third of the sample being 35 years or over. The average use experience was between one and three months, with around two-thirds using the virtual world for less than six months. However,
actual usage was high amongst the sample, with a median usage of between 4 and 10 hours per week, over five or six sessions per week.

4.2 Tests for Validity and Reliability of the Measures

Table 1 examines convergent validity – the extent to which theoretical scale items are empirically related. The loadings of the measures on their respective constructs in the model range from 0.756 to 0.946, with all being significant at the 0.1% level. Further, all of the constructs fulfill the recommended levels with reference to composite reliability (CR) and average variance extracted (AVE); as shown in Table 1, all items were higher than the cut-off of 0.50 recommended by Fornell and Larcker (1981). All the values of AVE are considered acceptable, ranging from 0.672 to 0.865. Similarly, the values of composite reliability are very good, ranging from 0.892 to 0.962, well above the reliability values of 0.70 and 0.80 that are typically advised for building strong measurement constructs (Nunnally 1978, Straub & Carlson 1989).

Table 1. Psychometric table of measurements.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Item</th>
<th>Loading</th>
<th>St. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absorption</td>
<td>When I am in Second Life I forget everything else around me.</td>
<td>0.807</td>
<td>0.026</td>
<td>31.327</td>
</tr>
<tr>
<td></td>
<td>It is difficult to detach myself from Second Life.</td>
<td>0.826</td>
<td>0.026</td>
<td>93.215</td>
</tr>
<tr>
<td></td>
<td>I feel happy when I am using Second Life Intensively.</td>
<td>0.825</td>
<td>0.019</td>
<td>34.654</td>
</tr>
<tr>
<td>Confirmation</td>
<td>My experience with using Second Life was better than I expected.</td>
<td>0.907</td>
<td>0.004</td>
<td>75.239</td>
</tr>
<tr>
<td></td>
<td>The benefit provided by Second Life was better than I expected.</td>
<td>0.909</td>
<td>0.018</td>
<td>77.417</td>
</tr>
<tr>
<td></td>
<td>Overall, most of my expectations from using Second Life were confirmed.</td>
<td>0.852</td>
<td>0.023</td>
<td>37.286</td>
</tr>
<tr>
<td>Continuance Intention</td>
<td>I intend to continue using Second Life rather than use any alternative technology.</td>
<td>0.907</td>
<td>0.003</td>
<td>74.001</td>
</tr>
<tr>
<td></td>
<td>My intentions to continue using Second Life rather than use any alternative technology.</td>
<td>0.912</td>
<td>0.004</td>
<td>90.249</td>
</tr>
<tr>
<td></td>
<td>If I could, I would continue my use of Second Life.</td>
<td>0.874</td>
<td>0.016</td>
<td>53.565</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>I would describe my overall experience of using Second Life as...</td>
<td>0.936</td>
<td>0.012</td>
<td>77.033</td>
</tr>
<tr>
<td></td>
<td>a. Enjoyable to disgusting</td>
<td>0.892</td>
<td>0.011</td>
<td>84.309</td>
</tr>
<tr>
<td></td>
<td>b. Exciting to dull</td>
<td>0.946</td>
<td>0.009</td>
<td>101.341</td>
</tr>
<tr>
<td></td>
<td>c. Pleasant to unpleasant</td>
<td>0.912</td>
<td>0.016</td>
<td>56.593</td>
</tr>
<tr>
<td></td>
<td>d. Interesting to boring</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habit</td>
<td>Using Second Life has become automatic to me.</td>
<td>0.863</td>
<td>0.017</td>
<td>75.438</td>
</tr>
<tr>
<td></td>
<td>Using Second Life is natural to me.</td>
<td>0.916</td>
<td>0.012</td>
<td>76.785</td>
</tr>
<tr>
<td></td>
<td>When faced with a particular task, using Second Life is an obvious choice for me.</td>
<td>0.879</td>
<td>0.015</td>
<td>50.485</td>
</tr>
<tr>
<td>Perceived Usefulness</td>
<td>Second Life is of benefit to me.</td>
<td>0.918</td>
<td>0.011</td>
<td>84.818</td>
</tr>
<tr>
<td></td>
<td>The advantages of Second Life outweigh the disadvantages.</td>
<td>0.808</td>
<td>0.016</td>
<td>54.748</td>
</tr>
<tr>
<td></td>
<td>Overall, using Second Life is advantageous.</td>
<td>0.912</td>
<td>0.011</td>
<td>84.165</td>
</tr>
<tr>
<td>Frequency of Prior Behaviour</td>
<td>In the last 7 days, how much time would you say you spent using Second Life?</td>
<td>0.909</td>
<td>0.015</td>
<td>62.741</td>
</tr>
<tr>
<td></td>
<td>In the last 7 days, how many times did you use Second Life?</td>
<td>0.852</td>
<td>0.021</td>
<td>53.159</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>How do you feel about your overall experience of Second Life use?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>a. Dissatisfied to satisfied</td>
<td>0.887</td>
<td>0.020</td>
<td>43.884</td>
</tr>
<tr>
<td></td>
<td>b. Dissatisfied to pleased</td>
<td>0.910</td>
<td>0.015</td>
<td>60.844</td>
</tr>
<tr>
<td></td>
<td>c. Frustrated to contented</td>
<td>0.855</td>
<td>0.018</td>
<td>48.083</td>
</tr>
<tr>
<td></td>
<td>d. Terrible to delighted</td>
<td>0.901</td>
<td>0.016</td>
<td>56.254</td>
</tr>
</tbody>
</table>

Table 2 examines discriminant validity – the extent to which question items measure the construct intended or other related constructs. A standard test for discriminant validity was used whereby the square root of average variance extracted for each construct is compared with the correlations between it and other constructs; discriminant validity is demonstrated if the square root is higher than the correlations. Table 2 clearly indicates that each construct shares greater variance with its own measurement items that with other constructs with different measurement items, with a good margin.
of difference, the closest margin being 0.122. To further confirm the discriminant validity of the scale items used in testing the research model we utilized the cross-loading method of Chin (1998); the method prescribes a requirement for measurement items to load higher on a construct than the scale items for other constructs and for no cross-loading to occur. Item loadings in the relevant construct columns were all higher than the loadings of items designed to measure other constructs; similarly, when glancing across the rows the item loadings are considerably higher for their corresponding constructs than for others.

<table>
<thead>
<tr>
<th></th>
<th>ABS</th>
<th>CONF</th>
<th>CONTINT</th>
<th>ENJ</th>
<th>HABIT</th>
<th>PU</th>
<th>PRIORBEH</th>
<th>SATIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absorption(ABS)</td>
<td>0.820</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confirmation(CONF)</td>
<td>0.583</td>
<td>0.890</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuance Intention(CONTINT)</td>
<td>0.584</td>
<td>0.753</td>
<td>0.905</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enjoyment(ENJ)</td>
<td>0.397</td>
<td>0.455</td>
<td>0.433</td>
<td>0.930</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habit(HABIT)</td>
<td>0.698</td>
<td>0.683</td>
<td>0.645</td>
<td>0.405</td>
<td>0.896</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Usefulness(PU)</td>
<td>0.546</td>
<td>0.693</td>
<td>0.691</td>
<td>0.453</td>
<td>0.666</td>
<td>0.909</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency of Prior Behaviour (PRIORBEH)</td>
<td>0.463</td>
<td>0.365</td>
<td>0.364</td>
<td>0.282</td>
<td>0.450</td>
<td>0.369</td>
<td>0.897</td>
<td></td>
</tr>
<tr>
<td>Satisfaction(SATIS)</td>
<td>0.447</td>
<td>0.481</td>
<td>0.403</td>
<td>0.636</td>
<td>0.485</td>
<td>0.462</td>
<td>0.320</td>
<td>0.896</td>
</tr>
</tbody>
</table>

Table 2. Correlations between constructs (diagonal elements are square roots of the average variance extracted).

Overall, these tests for validity and reliability provide us with a high degree of confidence in the scale items used in testing our research model.

4.3 Test of the Research Model

Figure 2 presents the results of PLS path modelling (Centroid Weighting Scheme) in Smart-PLS (Ringle et al. 2005). The shaded items are those that have been added to Bhattacharjee’s (2001) basic continuance model. A power analysis in G*Power 3.0 (Faul et al. 2007) shows that the sample (n=339) is sufficient for explaining even small population effects ($f^2 \geq 0.044; \alpha=0.05; \beta=0.2$) in our model, with a power of 1.000 for moderate population effects ($f^2=0.15$).

![Figure 2. Results of PLS analysis.](image-url)
The data supports all relationships in the model except two – the linkages between perceived usefulness and satisfaction and between satisfaction and continuance intention. Interestingly, these are core relationships in the basic continuance model. Not surprisingly, confirmation strongly influences and contributes to explaining perceived usefulness ($p<0.001$, $R^2=0.481$). Confirmation also strongly relates to enjoyment, with modest explanatory power ($p<0.001$, $R^2=0.207$), whilst confirmation and enjoyment together strongly influence satisfaction ($p<0.001$ and $p<0.01$ respectively) and explain a rather substantial 45.7 percent of variance.

Habit in our model appears to be significantly determined by satisfaction and the frequency of prior usage behaviour (both $p<0.001$ in our model); these two variables together explain a healthy 33.1 percent of variance in habit. Absorption has strong relationships with habit ($p<0.001$) and enjoyment ($p<0.01$), which together contribute to explaining a very considerable level of variance (50.4 percent).

Overall, the research model explains a very considerable level of variance in continuance intention – 56.5 percent. Four determinants contribute to explaining behavioural intention, specifically, in order of the strength of their relationships: perceived usefulness ($p<0.001$), habit ($p<0.001$), absorption ($p<0.01$) and enjoyment ($p<0.05$). In sum, users build their behavioural intentions to continue to use the virtual world of Second Life based on utilitarian assessment of value (perceived usefulness), hedonistic assessment of value (enjoyment), chronic disposition (habit) and via the persistent “flow” of the virtual world experience within which users are absorbed.

**DISCUSSION AND CONCLUSIONS**

This study has tested an extended continuance model in the context of virtual worlds, or more specifically, in the context of the virtual world of Second Life. The tested research model combines the ex post continuance intention model of Bhattacherjee (2001) with additional constructs for habit (Limayem et al. 2007), enjoyment (van der Heijden 2003) and absorption (Schaufeli et al. 2002), that are each posited to play an important role in continued virtual world usage. Using a sample of 339 respondents collected in Second Life we find strong support for the validity and reliability of all constructs. Furthermore, the model is tested using PLS and four of the determinants of continuance intention are clearly supported in the model (perceived usefulness, habit, absorption and enjoyment), along with all of the extended relationships proposed. However, although aspects of the expectation-confirmation theory are supported, including the extension to include hedonic value (enjoyment), the traditional link from utilitarian value (perceived usefulness) to satisfaction is not supported by the data. Similarly, the link from satisfaction to continuance intention is also not supported – only a linkage that is mediated by habit. This would appear to suggest that the standard continuance model is not supported in this context, and that a construct such as enjoyment (which bears some similarity to satisfaction) may be more suitable as a measure of the confirmed value derived from virtual world adoption and use than that of satisfaction per se. Users of the virtual world do not continue their use based directly on whether they are ‘satisfied’, but rather on an assessment of hedonic value and of utilitarian value, and based on established patterns of automatic behaviour and the chronic disposition to be absorbed in system use. Indeed satisfaction is mediated by habit in its path to continuance intention. Satisfaction appears to be a term more suited to non-hedonic product or service evaluation and is rather telling of the context from which it is derived – services marketing. Users of virtual worlds may be too bound up in the experience and the use of the system to be able to detach themselves and make an objective assessment of how satisfied they are or not with the virtual world within which they are habitually absorbed and use this as the basis for further usage decisions.

The nature of virtual worlds is clearly quite different to other product or service contexts within which ECT marketing theory is generally applied. The closest group of products/services would appear to be online gaming; although virtual worlds such as Second Life are not strictly games there is striking similarity to other hedonic systems such as online games and this is a domain that is likely to help in informing practice. Users of virtual worlds clearly place an emphasis on hedonic and utilitarian value derived, and these are areas within which developers can contribute to the success of virtual worlds.
Other aspects, such as habit and absorption, are dispositions of the individual and are not directly controllable; however, they can be manipulated indirectly through the design of the virtual world. Enjoyable experiences that leave users wanting more contribute to building habitual behaviour. Similarly, enjoyable experiences in realistic virtual worlds and compelling human-computer interaction can help to build persistent flow experiences and a chronic disposition where individuals find it hard to discontinue use of system – perhaps even an addictive behaviour, as has been shown for computer gaming environments.

Utilitarian value, as represented by perceived usefulness in our study, can be enhanced by creating additional instrumental value for users. Such value is created by providing additional functionality and content to virtual worlds that is clearly of benefit to residents and that will ultimately be used. This could include, for example: tools for communication, collaboration and social networking; tools for entrepreneurial or commercial endeavours; utilities for customization and creation of objects and other content; tools for the development of simulated environments; systems for establishing and managing a presence in a virtual world and in relation to the operator of the world; educational and training-oriented systems; other productivity tools; and content that appeals to the typical goals of individuals using the virtual world.

Hedonic value, as discussed by van der Heijden (2003, p. 696), can be enhanced by “the inclusion of hedonic content, animated images, a focus on colours, sounds, and aesthetically appealing visual layouts.” Within the context of Second Life this means creating ever more realistic and visually appealing 3-D rendered environments, with strong visual appeal and integration of quality audio; Second Life has continued to work in this space, including, for example, more realistic representations of sky and water – originally trailed through the WindLight test version but now built into the standard client. The key is providing emotional content including images, objects and sounds that stimulate the user, provoke a response and assist in creating an absorptive experience. The provision of emotional content should be dynamic and changing, rather than static, creating a compelling reason for users to return to the virtual world and contributing to the creation of a persistent absorptive effect on the user.

Our research is limited to the extent that we have focused solely on the virtual world of Second Life. Other virtual worlds differ according to aspects such as the payment model, gaming model, sectoral customer focus and technical design. Notwithstanding, the constructs used in our model would appear generic and robust enough to be applicable in a wide variety of combinations of these virtual worlds. It is possible that the expectation-confirmation model may reveal different results – in particular a link from perceived usefulness to satisfaction to continuance intention in different virtual worlds. Given the nature of Second Life, this is perhaps unlikely and we might expect Second Life to behave more as a barometer for other virtual worlds, many of which are only beginning to establish a presence and build a customer base. Future research should aim to test the model in further virtual world environments in order to confirm this assumption.

A further limitation of our research model is the absence of demographic and other factors, such as age, gender, geographical location, experience with Second Life, use intensity, personality and so on. Although demographic and use data were captured in the study, there was not appropriate scope in this paper to examine their direct or moderating effects in a research model. For example, we are likely to have a heterogeneous sample of a global nature; capturing and analyzing information on the sample may help to understand the behaviour of different groups of global consumers. New PLS techniques such as finite mixture analysis (FIMIX) may prove to be very useful in this regard, providing a means of respondent group segmentation and analysis (Ringle et al., in press). Further, since virtual worlds have elements of hedonic and utilitarian value and can be used for many purposes (including social and work purposes, for example) it would be useful to delve deeper into the specific purposes of usage and whether these have an impact on continuance behaviour in future studies.

This paper represents an early attempt to understand behaviour in virtual worlds from an information systems perspective. There is much to learn in this new environment and IS theory needs to develop further to assess virtual world behaviour. Virtual worlds different to standard domains of theoretical
application in technology acceptance of IS which, although they may exhibit intrinsic elements, tend to be largely extrinsic in use motivation, such as individual use of organizational systems, e-commerce systems and the Internet, mobile technology, and so on. Hedonic systems and related theory are underrepresented in the overall IS acceptance literature, which is unfortunate given the hedonistic nature of the system under investigation here. Virtual worlds appear too complex to be the focus of one discipline alone. This is an issue of concern for future research, which should look towards a multidisciplinary approach towards understanding hedonistic IS such as virtual worlds; in such an approach disciplines such as psychology, human-computer interaction and consumer behaviour are likely to play important roles.

Acknowledgements
The help of Mario Menti from GMI, Inc. is gratefully acknowledged in this project.

References


THE MODERATING ROLE OF UTILITARIAN/HEDONIC USER MOTIVATION ON USER BEHAVIOUR TOWARDS WEB 2.0 APPLICATIONS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0383.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Web 2.0, Behavior, Structural Equation Modeling, Technology Acceptance Model (TAM)</td>
</tr>
</tbody>
</table>
THE MODERATING ROLE OF UTILITARIAN/HEDONIC USER MOTIVATION ON USER BEHAVIOR TOWARDS WEB 2.0 APPLICATIONS

Chen-Ya Wang, Department of Information Management, National Taiwan University, Taiwan; Department of Information Management, Lunghwa University of Science and Technology, Taiwan, ntuimgrace@gmail.com

Seng-cho T. Chou, Department of Information Management, National Taiwan University, Taiwan, chou@ntu.edu.tw

Hsia-Ching Chang, Department of Informatics, University at Albany, State University of New York, USA, carriehe@gmail.com

Abstract

Web 2.0 is now an important internet application because of the integration of social interaction and web technologies. Previous information system studies usually specified their research context as a utilitarian system or hedonic system and the results were concluded within one specific system type. Web 2.0 application provides a flexible environment for different kinds of user motivations that can be used for utilitarian or hedonic purpose. This study extended the Technology Acceptance Model (TAM) by introducing a moderating factor into the model, in order to study users’ behavioral intentions in a Web 2.0 environment. We designed two task types of user motivation and conducted our experiment on two Web 2.0 websites. According to the PLS (Partial Least Squares) analysis, this study demonstrated that utilitarian and hedonic purposes had a moderating effect on the relationship between perceived belief and user attitude as well as the relationship between perceived information quality and perceived belief in the Web 2.0 application. The relationship between perceived usefulness and attitude was stronger in the utilitarian user motivation; whereas the relationship between perceived ease of use and attitude was stronger when the user had hedonic motivation to use the Web 2.0 application. We also found that perceived information quality had significant impact on the perceived usefulness and perceived ease of use.

Keywords: moderating effect, utilitarian/hedonic user motivation, perceived web quality, technology acceptance model (TAM), web 2.0 application
1 INTRODUCTION

With the rapid development of internet technology, numerous web applications have become available to internet users over the past decade. Web 2.0 is one innovative web application that provides an environment for human social interaction as well as information sharing. Based on the core spirit of Web 2.0: ‘participation’, many kinds of applications such as blogs (i.e. Weblogs), folksonomies, wikis, etc., have emerged to satisfy various user needs. The user studies related to Web 2.0 application have grown quickly during these years. However, most of them only choose one specific type of Web 2.0 applications and seldom discuss the user motivation influences while they use the Web 2.0 application (Hsu & Lin 2008). For this reason, this study chooses two types of Web 2.0 applications and designs two types of user motivations to examine the user behavior in Web 2.0 applications.

The Technology Acceptance Model (TAM) is widely recognized within the information system (IS) research field as a useful model for exploring user attitude and intention toward an information system. Many studies have included other factors into TAM (such as: subjective norm, output quality, intrinsic/extrinsic motivation, etc.) to provide different perspectives as well as to enhance the explanatory power of user behaviour (Davis et al. 1992, Jarvenpaa & Todd 1997, Venkatesh 2000, Venkatesh & Davis 2000, Shang et al. 2005). Other studies have explored the effect of the moderating factor on user acceptance of information system (e.g. experience, voluntariness, etc.) (Castañeda et al. 2007, Jarvenpaa & Todd 1997). The user motivations studies in the IS field usually viewed user motivations as determinants of user perceived belief, and divided user motivations into intrinsic and extrinsic motivation. In many studies, the TAM model was applied to the internet environment in order to examine the user acceptance of internet applications (Jarvenpaa & Todd 1997, Moon & Kim 2001, Shang et al. 2005). Related researches also considered other important factors in internet application, to enhance the explanatory power related to user behavior (Jarvenpaa & Todd 1997, Dishaw & Strong 1999, Venkatesh & Davis 2000, Shang et al. 2005, Hsu & Lin 2008). For example, the user’s perception of web quality is now recognized as a factor that influences the user’s perceived belief regarding the internet application (Lin & Lu 2000, Liu & Arnett 2000, Moon & Kim 2001, Cao et al. 2005, Éthier et al. 2006). In the internet environment, a user might use one web application for utilitarian or hedonic purpose and these goals might moderate the user’s acceptance of these applications (Childers et al. 2001). Web 2.0 applications have a special characteristic that can be used for utilitarian or hedonic motivation, providing the possible extension to explore user motivation from a different perspective. Our research model, based on TAM, will include the perceived web quality factor to explore the moderating influence of utilitarian/hedonic user motivations in Web 2.0 application.

In this paper, we propose a model that combines the perceived web quality factor with the TAM model to examine the moderating effect of utilitarian/hedonic user motivations on the user acceptance of Web 2.0 application. We designed four experimental tasks based on two user motivations (i.e., utilitarian and
hedonic motivation) and two web 2.0 applications (i.e., video share and blog websites) for participants to collect self-report questionnaire data. This study used SmartPLS 2.0 to analyze the experimental data and provided some implications for the development of the Web 2.0 application.

2 RESEARCH MODEL AND HYPOTHESES

The moderating role of utilitarian/hedonic user motivations in the context of the Web 2.0 application has never been studied thus far. Our model, based on TAM, assumes that utilitarian/hedonic user motivations will moderate the relationship between users’ perceived belief and attitude, as well as the relationship between users’ perceived web quality and perceived belief. Figure 1 shows our proposed model. The following parts will describe the hypotheses.

![Figure 1. Theoretical Model](image)

### 2.1 Technology Acceptance Model (TAM)

The Technology Acceptance Model (TAM) proposed by Davis (1989) attempted to explain and predict user acceptance of IS and posited that user acceptance can be determined by user behavioral intention. The intention is determined by user attitude and the user’s perceived usefulness regarding the IS. User attitude is determined by two salient beliefs: perceived usefulness (PU) and perceived ease of use (PEOU). Furthermore, perceived ease of use is a determinant of perceived usefulness. These relationships in TAM have been examined in many related studies (Davis 1989, Venkatesh & Davis 1996, Venkatesh & Davis 2000). TAM is now widely used to predict the acceptance of new technology applications, such as new information systems, new web applications, and mobile services, etc (Lin & Lu 2000, Cao et al. 2005, Ahn et al. 2007, Hsu & Lin 2008). Many scholars extended the TAM model by introducing more variables to enhance the explanatory power of TAM (Dishaw & Strong 1999, Venkatesh & Davis 2000, Shang et al. 2005, Ahn et al. 2007). The purpose of this study is to extend...
TAM model by including perceived web quality as the determinants of perceived belief, and examine
the moderating role of utilitarian/hedonic user motivations, and consequently, we only revalidated the
original relationships in TAM model in the Web 2.0 application. These hypotheses are:

H1: Perceived ease of use has a positive effect on perceived usefulness in the Web 2.0 applications.

H2: Perceived ease of use has a positive effect on user attitude in the Web 2.0 applications.

H3: Perceived usefulness has a positive effect on user attitude in the Web 2.0 applications.

H4: User attitude has a positive effect on user behavioral intention in the Web 2.0 applications.

H5: Perceived usefulness has a positive effect on user behavioral intention in the Web 2.0 applications.

2.2 Perceived web quality

Delone and Mclean (1992) proposed a model using information and system quality to evaluate the
success of an information system. Information quality captures the users’ perceived value of the output
produced by a system and can be measured by information accuracy, relevance, timeliness, and
completeness etc (Negash et al. 2003, Éthier et al. 2006). System quality is a measure of the
functionality of a system including: usability, availability, reliability, and response time, etc (Delone &
Mclean 1992, Negash et al. 2003, Cao et al. 2005). A service quality component was added to enhance
customer service settings in later studies (Kettinger & Lee 1994, Pitt & Watson & Kavan 1995).
Service quality measures how well the delivered service matched the customer expectations (Cao et al.
2005, Ahn et al. 2007). In recent years, many researches have explored the perceived web quality
impact on user acceptance in the context of web applications (Cao et al. 2005, Éthier et al. 2006, Ahn et
al. 2007). Most of them categorized perceived web quality into: information, system, and service
quality; they showed that perceived web quality had a positive impact on perceived belief (Lin & Lu
focuses on the moderating effect of utilitarian/hedonic user motivations on user behavior by conducting
four experiments. We will not examine the service quality component of perceived web quality. The
perceived web quality impact on user perceived belief will be examined with the following hypotheses:

H6a: Perceived information quality has a positive effect on perceived usefulness (PU) in regard to Web
2.0 applications.

H6b: Perceived information quality has a positive effect on perceived ease of use (PEOU) in regard to
Web 2.0 applications.

H7a: Perceived system quality has a positive effect on perceived usefulness (PU) in regard to Web 2.0
applications.

H7b: Perceived system quality has a positive effect on perceived ease of use (PEOU) in regard to Web
2.0 applications.
2.3 Utilitarian/Hedonic user motivations

Many traditional information system studies are under the research context of utilitarian system type. Some scholars have argued that the results of IS user behavior studies might differ when the system types are different. Van der Heijden (2004), for instance, validated the technology acceptance model from the hedonic system perspective. The distinction of hedonic/utilitarian systems is derived from the utilitarian and hedonic products concept of consumer behavior literature (Hirschman & Holbrook 1982, Holbrook & Hirschman 1982, Van der Heijden 2004). The consumer behavior studies showed that the utilitarian or hedonic nature of the product determined the intention to consume (Babin et al. 1994). Some studies in the marketing field also find that the utilitarian or hedonic shopping purposes influence consumer shopping behavior (Chitturi et al. 2008, Kim & Shim 2002).

In the IS field, much research has explored the impact of user motivations on user behavioral intention by defining user motivations as intrinsic or extrinsic (Davis et al. 1992, Venkatesh 2000). Beyond the traditional IS research, user motivations are also discussed in the context of internet application. Some studies defined two types (i.e., specific task and general task) of motivation to observe user search behavior (Kim & Allen 2002, Kim 2008). Other studies examined the user online shopping behavior and included the utilitarian/hedonic motivation factors of consumer consumption behavior that are widely used in the marketing field (Childers et al. 2001). Our study adopted the utilitarian/hedonic perspective that the Web 2.0 application could satisfy users with both utilitarian and hedonic purposes; therefore, this study designed the experimental tasks from just such a perspective. We assumed that the determining strength of perceived usefulness (PU) and perceived ease of use (PEOU) on user attitude would be different when the user had a different motivation. We also proposed that the strength of the relationship between perceived web quality and perceived belief will differ according to the utilitarian/hedonic user motivations. The moderating effect of user motivations on user attitude will be examined with the following hypotheses:

H8: The relationship between perceived ease of use and perceived usefulness is stronger in hedonic motivation than in utilitarian motivation.

H9: The relationship between perceived ease of use and user attitude is stronger in hedonic motivation than in utilitarian motivation.

H10: The relationship between perceived usefulness and user attitude is stronger in utilitarian motivation than in hedonic motivation.

H11a: The relationship between perceived information quality and perceived usefulness is stronger in utilitarian motivation than in hedonic motivation.

H11b: The relationship between perceived information quality and perceived ease of use is stronger in hedonic motivation than in utilitarian motivation.
H12a: The relationship between perceived system quality and perceived usefulness is stronger in utilitarian motivation than in hedonic motivation.

H12b: The relationship between perceived system quality and perceived ease of use is stronger in hedonic motivation than in utilitarian motivation.

3 RESEARCH METHOD

3.1 Research variables

There are three kinds of research variables in this study: (1) TAM related factors, (2) perceived web quality, and (3) utilitarian/hedonic user motivations. The TAM related measures (i.e., user behavioral intention, attitude, perceived usefulness, and perceived ease of use) were adapted from the original TAM and the updated TAM model, TAM2 (Davis 1989, Venkatesh 2000). The measure items of perceived web quality (i.e., information quality and system quality) were derived from prior studies that had been widely validated in related studies (Cao et al. 2005, Éthier et al. 2006, Ahn et al. 2007). The wordings of these measure items were slightly modified to fit the Web 2.0 application environment. A seven-point Likert type scale was adopted with anchors ranging from strongly disagree (1) to strongly agree (7). As for the utilitarian/hedonic user motivation, it was the control variable that we assigned with different types of experimental tasks to the participants of this research.

3.2 Experiment design

To examine the moderating role of user motivation, this study designed two types of experimental tasks: (1) utilitarian user motivation, and (2) hedonic user motivation. This study defines the utilitarian task as finding the information that the participants want to know related to their work tasks or school courses. On the other hand, the hedonic task is defined as finding the information in which the participants are interested in their own life, for example hobbies, leisure activities, etc. The participants were asked to define two subjects regarding the experimental task type that they are assigned and search information that they wanted to obtain in the Web 2.0 application. This study chose two Web 2.0 websites for the experimental environment setting. The first is a worldwide video share website (called Website A in the following parts), which allows users to easily upload and share video clips across the internet through websites, mobile devices, blogs, and email. The second is a famous blog website (called Website B in the following parts), which provides an easy way for general internet users to publish material of any topic that they choose to share or discuss. This blog website also provides functions to share photos, pictures, and videos etc. With the combination of the two user motivations and the two Web 2.0 websites, four types of experimental tasks emerge. The experiments were conducted in computer laboratories with twenty minutes allowed for each task. The participants were
randomly divided into four groups and were asked to complete the online questionnaire immediately after finishing their experimental tasks.

3.3 Data collection

Data were gathered from students registered in eight business courses in the business school at a university in Taiwan. Of the 408 questionnaires collected, 366 completed and valid surveys were received. According to the four types of experimental task, 102 valid responses were obtained from the utilitarian task experiment and 92 valid responses were collected from the hedonic task experiment in Website A. 88 valid responses were obtained from the utilitarian task experiment and 84 valid responses were collected from the hedonic task experiment in Website B. Among the usable respondents, 71.6% were part-time students who had weekday jobs and 49.7% were male. 48.1% of the participants were between 18 and 25 years of age. Participants in this research frequently used internet applications with which they connected to the internet on average more than three hours each day. 76.8% of the participants had more than six years of internet experience.

4 DATA ANALYSIS AND RESULTS

Partial least squares (PLS) was applied to test our research model using SmartPLS 2.0. PLS is appropriate for exploratory research because it is more prediction-oriented than other structural equation modelling tools (Éthier et al. 2006). Data analysis proceeded in two stages (Anderson & Gerbing 1988). First, the measurement model was evaluated using confirmatory factor analysis (CFA) to validate the reliability and validity of the constructs. Next, the structural model was estimated using hypotheses testing to test the significance of the path coefficients. Our research model was evaluated with three groups of sample data: (1) the sub-sample of the utilitarian experimental task, (2) the sub-sample of the hedonic experimental task, and (3) the full sample.

4.1 Measurement model

The adequacy of our measurement model was evaluated on the criteria of composite reliability, convergent validity, and discriminant validity. Table 1 shows the analysis results of our measurement model. The composite reliability values ranged from .909 to .970, well above the suggested criteria of .7 (Fornell & Larcker 1981), indicating a commonly acceptable level for confirmatory research. Convergent validity was assessed by the average variance extracted (AVE) for each construct, which should exceed the variance due to measurement error for that construct. The AVE of each construct ranged from .590 to .924, exceeding the acceptable value of 0.5 (Fornell & Larcker 1981). Discriminant validity examines the uniqueness degree of item measures in defining a construct (Gefen 2003). The satisfactory discriminant validity is achieved when the square root of AVE for each construct exceeds all correlations between that and other constructs (Fornell & Larcker 1981). As shown in Table 1, our
evaluation of discriminant validity was acceptable. The analysis result in this stage showed that the measurements in our study were reliable and valid.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Items</th>
<th>CR</th>
<th>AVE</th>
<th>R²</th>
<th>Correlation of constructs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Utilitarian model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIQ</td>
<td>7</td>
<td>0.962</td>
<td>0.785</td>
<td></td>
<td><strong>0.886</strong></td>
</tr>
<tr>
<td>PSQ</td>
<td>7</td>
<td>0.909</td>
<td>0.591</td>
<td></td>
<td><strong>0.769</strong></td>
</tr>
<tr>
<td>PU</td>
<td>4</td>
<td>0.970</td>
<td>0.892</td>
<td>0.627</td>
<td>0.748</td>
</tr>
<tr>
<td>PEOU</td>
<td>5</td>
<td>0.925</td>
<td>0.713</td>
<td>0.476</td>
<td>0.627</td>
</tr>
<tr>
<td>Attitude</td>
<td>2</td>
<td>0.948</td>
<td>0.902</td>
<td>0.640</td>
<td>0.716</td>
</tr>
<tr>
<td>Intention</td>
<td>4</td>
<td>0.959</td>
<td>0.854</td>
<td>0.627</td>
<td>0.632</td>
</tr>
<tr>
<td><strong>Hedonic model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIQ</td>
<td>7</td>
<td>0.954</td>
<td>0.750</td>
<td></td>
<td><strong>0.866</strong></td>
</tr>
<tr>
<td>PSQ</td>
<td>7</td>
<td>0.909</td>
<td>0.590</td>
<td></td>
<td><strong>0.768</strong></td>
</tr>
<tr>
<td>PU</td>
<td>4</td>
<td>0.964</td>
<td>0.871</td>
<td>0.661</td>
<td>0.653</td>
</tr>
<tr>
<td>PEOU</td>
<td>5</td>
<td>0.954</td>
<td>0.804</td>
<td>0.397</td>
<td>0.557</td>
</tr>
<tr>
<td>Attitude</td>
<td>2</td>
<td>0.961</td>
<td>0.924</td>
<td>0.549</td>
<td>0.622</td>
</tr>
<tr>
<td>Intention</td>
<td>4</td>
<td>0.946</td>
<td>0.815</td>
<td>0.662</td>
<td>0.542</td>
</tr>
<tr>
<td><strong>Full model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIQ</td>
<td>7</td>
<td>0.959</td>
<td>0.772</td>
<td></td>
<td><strong>0.879</strong></td>
</tr>
<tr>
<td>PSQ</td>
<td>7</td>
<td>0.909</td>
<td>0.591</td>
<td></td>
<td><strong>0.769</strong></td>
</tr>
<tr>
<td>PU</td>
<td>4</td>
<td>0.970</td>
<td>0.891</td>
<td>0.642</td>
<td>0.709</td>
</tr>
<tr>
<td>PEOU</td>
<td>5</td>
<td>0.941</td>
<td>0.762</td>
<td>0.444</td>
<td>0.601</td>
</tr>
<tr>
<td>Attitude</td>
<td>2</td>
<td>0.957</td>
<td>0.917</td>
<td>0.623</td>
<td>0.682</td>
</tr>
<tr>
<td>Intention</td>
<td>4</td>
<td>0.955</td>
<td>0.841</td>
<td>0.647</td>
<td>0.603</td>
</tr>
</tbody>
</table>

Table 1. Reliabilities, Convergent validity, and Discriminant Validity

Notes: a: PIQ is “perceived information quality”; PSQ is “perceived system quality”; PU is “perceived usefulness”; PEOU is “perceived ease of use”. b: CR is “composite reliability”. c: Diagonal elements in the “correlation of constructs” matrix are the square root of the average variance extracted.

4.2 Structural model

To test the hypotheses, the bootstrapping technique was used to produce the t-value and test the significance of the path coefficients. The results of hypotheses testing are presented in Table 2.

- TAM related hypotheses were all confirmed (H1, H2, H3, H4, and H5). With the exception of the positive relationship between perceived usefulness and user intention (H5) could not be found in utilitarian user motivation.
The positive relationship between perceived information quality and the user perceived belief were all supported (H6a and H6b).

H7a was not supported, which meant that the perceived system quality did not have a positive effect on perceived usefulness. On the other hand, H7b was supported, which meant that the perceived system quality had a positive effect on perceived ease of use.

H8, H9, and H10 proposed that the relationship strength between perceived belief and user attitude will different when users has different motivations. They were all supported. The relationships between perceived ease of use and perceived usefulness and user attitude were stronger in hedonic motivation (.538, .220) than in utilitarian motivation (.290, .164). However, the relationship between perceived usefulness and user attitude was stronger in utilitarian motivation (.682) than in hedonic motivation (.559).

H11a and H11b examined the relationship strength between perceived information quality and perceived belief. Only H11a was supported: the perceived information quality had a stronger relationship with perceived usefulness in utilitarian motivation (.509) than in hedonic motivation (.287). H11b was not supported: the relationship between perceived information quality and perceived ease of use had no obvious difference (.395 versus .332).

H12a and H12b examined the relationship strength between perceived system quality and perceived belief. They were all rejected. According to the original hypothesis, H7a was rejected: the perceived system quality had no significant effect on perceived ease of use. H12a was supposed not to be supported.

<table>
<thead>
<tr>
<th>Path</th>
<th>Utilitarian model</th>
<th>Hedonic model</th>
<th>Full model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SPC</td>
<td>T-value</td>
<td>SPC</td>
</tr>
<tr>
<td>H1:PEOU→PU</td>
<td>0.290</td>
<td>4.045***</td>
<td>0.538</td>
</tr>
<tr>
<td>H2:PEOU→A</td>
<td>0.164</td>
<td>2.611**</td>
<td>0.220</td>
</tr>
<tr>
<td>H3:PU→A</td>
<td>0.682</td>
<td>11.353***</td>
<td>0.559</td>
</tr>
<tr>
<td>H4:A→I</td>
<td>0.650</td>
<td>7.952***</td>
<td>0.590</td>
</tr>
<tr>
<td>H5:PU→I</td>
<td>0.171</td>
<td>1.854</td>
<td>0.277</td>
</tr>
<tr>
<td>H6a:PIQ→PU</td>
<td>0.509</td>
<td>7.430***</td>
<td>0.287</td>
</tr>
<tr>
<td>H6b:PIQ→PEOU</td>
<td>0.395</td>
<td>5.685***</td>
<td>0.332</td>
</tr>
<tr>
<td>H7a:PSQ→PU</td>
<td>0.091</td>
<td>1.458</td>
<td>0.109</td>
</tr>
<tr>
<td>H7b:PSQ→PEOU</td>
<td>0.370</td>
<td>4.917***</td>
<td>0.370</td>
</tr>
</tbody>
</table>

Table 2 Models summary

Note: * p<0.05, ** p<0.01, *** p<0.001; SPC is “Standardized Path Coefficient”.

Proceedings ECIS 2009
5 DISCUSSION AND CONCLUSION

The purpose of our study was to extend the TAM model by including the perceived web quality component and examine the moderating impact of user motivations on user behavior in the Web 2.0 application. We demonstrated that perceived web quality (i.e., information quality and system quality) affected the perceived belief (i.e., PEOU and PU) of users, with the exception that perceived system quality did not have significant impact on perceived usefulness, which meant that users care about information quality more than system quality. It provides a practical implication for service providers. When designing applications, service providers should think about how to well organize, store, and share user provided information to strengthen attraction to their provided services.

The results showed that utilitarian and hedonic motivations had a moderating impact on the relationship between perceived belief and user attitude as well as on the relationship between perceived information quality and perceived belief in the Web 2.0 application. An interesting finding was that the relationship between perceived usefulness and user attitude was stronger in the utilitarian user motivation. On the contrary, the relationships between perceived ease of use and user attitude and behavioral intention were stronger when the user had hedonic motivation. From the practical perspective, the developers and designers of Web 2.0 application should clearly define what type of values they want to provide to their users. For example, if service providers want to attract more utilitarian-oriented users or provide their users with more utilitarian-oriented values, they should devote greater efforts to determining which design of their services can strengthen the users’ perception of the usefulness of those services. Furthermore, the user motivation only affected the relationship between perceived information quality and perceived usefulness, which was stronger in utilitarian user motivation. Therefore, the service providers should endeavor to ensure the information quality of their services, especially when they aim to provide utilitarian value to their users.

TAM related hypotheses were all supported, with the exception of the H5, which meant that perceived usefulness did not have a direct relationship with user intention in utilitarian motivation. The result showed that the path coefficient between user attitude and user intention was greater in utilitarian motivation (.650 versus .590). It might imply that the relationship between perceived usefulness and user intention in utilitarian motivation is indirect and also mediated by user attitude. This phenomenon still needs further exploration in future research.

This study has some limitations that provide some opportunities for future research. First, Web 2.0 applications are not limited to these two types we chose for conducting experiment. Further research needs to include more kinds of Web 2.0 applications to reach a more general research conclusion. Second, though most participants had weekday jobs, still 28.4% of the sample was full-time university students. The data collection in future research has to recruit participants from various backgrounds to avoid sampling bias.
References


DISCUSSION OF FUNCTIONAL DESIGN OPTIONS FOR ONLINE RATING SYSTEMS: A STATE-OF-THE-ART ANALYSIS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0607.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-commerce (B2B / B2C / B2G / G2C), Reputation, Web 2.0, User participation</td>
</tr>
</tbody>
</table>
DISCUSSION OF FUNCTIONAL DESIGN OPTIONS FOR ONLINE RATING SYSTEMS:
A STATE-OF-THE-ART ANALYSIS

Winkelmann, Axel, University of Koblenz-Landau, Department of Computer Science, Institute for IS Research, Universitätsstraße 1, 56070 Koblenz, Germany, winkelmann@uni-koblenz.de

Herwig, Sebastian, University of Münster, European Research Center for Information Systems (ERCIS), Leonardo-Campus 3, 48149 Münster, Germany, sebastian.herwig@ercis.uni-muenster.de

Pöppelbuß, Jens, University of Münster, European Research Center for Information Systems (ERCIS), Leonardo-Campus 3, 48149 Münster, Germany, jens.poepelbuss@ercis.uni-muenster.de

Tiebe, Daniel, University of Münster, European Research Center for Information Systems (ERCIS), Leonardo-Campus 3, 48149 Münster, Germany, daniel.tiebe@ercis.uni-muenster.de

Becker, Jörg, University of Münster, European Research Center for Information Systems (ERCIS), Leonardo-Campus 3, 48149 Münster, Germany, joerg.becker@ercis.uni-muenster.de

Abstract

Ratings are important for building up trust among different parties. Since the arrival of the internet era in the 1990s, countless online rating systems have emerged. For example, Amazon.com established a rating system for books and other products in 1995. Today online rating can be found everywhere, be it e-commerce sites, social networks, and information or recommendation platforms. In most cases, users provide the input to these systems which is then aggregated and directed to appropriate recipients. The increasing relevance of these rating systems forms a new distinct research field with a growing need for research on the design, effects, and validity of rating systems. Hence, we contribute to the body of knowledge by conducting a thorough analysis of the state of the art of online rating systems. We especially focus on the functional design options by analyzing 102 different systems with the help of a criteria catalog of 237 criteria. In this paper, we discuss an excerpt of our findings and present a morphological box that categorizes functional design options for online rating systems.

Keywords: E-Commerce, Reputation, Web 2.0, User Participation.
1 INTRODUCTION

Ratings are important to people, especially for building up trust among business partners. The traditional offline rating systems have always been and still are word of mouth and gossip (Cheung & Luo & Sia & Chen 2007, Dellarocas 2001). Nevertheless, since the internet era begun in the 1990s, new online rating systems emerged. As some of the first, Armstrong and Hagel III (1995) dealt with these systems and they found out that customers fall back on advices and recommendations of other customers when executing online transactions. Moreover, they discovered that online rating systems can even be a means to improve customer retention (Armstrong & Hagel III 1995). Being one of the rather progressive in this field, Amazon.com has enabled its users to rate books and other products since 1995. Many other online stores have also included rating systems on their websites and even pure rating web-portals have emerged; e.g. Epinions.com was launched in 1999 which – according to the Alexa Rating – is the most popular web-portal for reviews on products and services today. All this contributes to make electronic shopping a social experience because of an intensive sharing of recommendations and reviews (Vossen & Hagemann 2007).

The advantage of online – in contrast to offline – ratings is scalability and formalization (Bolton & Katok & Ockenfels 2004, Dellarocas 2003). Scalability means that ratings can be gathered from and communicated to a multitude of parties, independent from time and place (Resnick & Kuwabara & Zeckhauser & Friedman 2000). Users can access a huge number of ratings provided by other users in an easy and cost-efficient manner (Cheung et al. 2007). According to Töpfer, Silbermann and William (2008) especially the fast diffusion of up-to-date rating information is a main advantage (Conte & Paolucci 2002, Töpfer et al. 2008). Through unification of gathering, aggregation and presentation ratings become more comprehensible and their acceptance increases (Resnick & Zeckhauser 2001).

In fact, the main field of application of online rating systems is e-commerce which has gained increasing importance over the last two decades (Kapell 2007, Sebralla 2008). Nevertheless, one key problem of online transactions is the lack of trust among business partners. Users of online shops often have problems in estimating the quality and reliability of offers and their counterparts (Ward & Lee 1999). To address this issue, rating and reputation systems have become an integral part of many e-commerce sites and are especially supposed to facilitate the identification of trustworthy business partners. For instance, each transaction at Ebay.com is expected to be rated by the involved users in order to build up their rating profiles which are accessible to any other registered user. Analyses have shown that there is an impact of the seller’s rating profile on the likeliness of a successful transaction as well as the achieved selling price (Armstrong & Hagel III 1995, Sebralla 2008).

Nevertheless, rating systems can also be found apart from e-commerce, e.g. in social networks that people use to manage their connections to other people. A typical example of a social network used for business purposes is LinkedIn, where users can write recommendations for colleagues, business partners or professional service providers. Similar to a traditional job reference, recommendations refer to their work and are listed in their profiles. Even more rating systems are found on non-business network sites like MySpace.com, Facebook.com or Skyrock.com where people are able to rate e.g. user profiles or videos.

The increasing relevance and spread of rating systems forms a new distinct research field. According to Peters and Reitzenstein (2008), there is an increasing need to do research on the forms, effects and validity of rating systems. For instance, relevant questions are how to aggregate rating input and represent valid rating results as well as the effects on (business) partner selection or trust building. So far, there is a lack of studies on the functional design of differing rating systems although it is the design of rating systems that determines the effect of their results to a large extent. In this paper, we will analyze the state of the art of existing rating systems and derive a consolidated set of design options which are discussed in respect to their influence on the entire rating system.
The remainder of this paper is structured as follows. The next section will provide a brief theoretical background on online rating systems. Thereafter, we will present the research method we applied for our state-of-the-art analysis. Then, we will describe excerpts from our results which we will subsequently discuss in order to provide a consolidated morphological box of design options. We will conclude with explaining the limitations of our study and providing an outlook on further research.

2 THEORETICAL BACKGROUND

The basic idea of online rating systems is to let users rate entities by means of web applications and hence to collect, aggregate and distribute reviews on entities (Resnick et al. 2000). The aggregated ratings about a given party or entity can be used to derive a score, e.g. a trust or reputation score, which is communicated to other parties. These scores can assist these parties in deciding whether or not to transact with certain other parties in future (Josang & Roslan & Boyd 2007). Typically, rating systems have a central authority that collects ratings and disseminates the rating results/scores.

The terms rating, review and reputation are closely linked. “Reputation is what is generally said or believed about a person’s or thing’s character or standing.” (Josang et al. 2007) It can be considered as a collective measure that is based on the reviews or ratings from members in a community. Web applications that gather reviews or ratings in order to compute and communicate reputation scores are called rating systems or reputation systems which we regard as synonyms throughout this paper. However, recommender systems must not be confused with reputation or rating systems. Recommender systems generate individualized recommendations to guide users in a personalized way by applying collaborative filtering or similar (Burke 2002); e.g. Amazon.com’s “Customers Who Bought This Item Also Bought” application. Rating systems can rather be described as collaborative sanctioning systems because they are based on the assumption that all users in a community should judge the performance or quality of an entity (e.g. transaction partner, product, information or multimedia contents). The aim is to sanction poor entities and to give an incentive for improvement (Josang et al. 2007).

A rating itself can be understood as a certain believe to which degree the rated entity is useful for a given objective (Miceli & Castelfranchi 2000). The criteria for this rating are usually determined by the objective whereas the entity determines the values that are assigned to each criteria (Schuler 2004). The appraisal usually happens subjectively and depends on the skills and experiences of the party that rates the entity. Obviously, a rating is only a single opinion which also refers to a specific situation and time (Sabater-Mir & Paolucci 2007). Furthermore, as an empirical analysis showed, users tend to rate positively and avoid negative statements (Resnick & Zeckhauser 2001).

The use of online rating systems requires an adequate design of the underlying mechanisms. According to Chen, Hogg and Wozny (2004), one of the main decisions refers to the gathering of information. Operators of such systems must determine which users are allowed to rate which entities. Especially, deliberate manipulations by single users must be avoided (Dellarocas 2003). To respect user privacy, users should have to opt-in before being subject of ratings (Ziegler 2008).

3 RESEARCH METHOD

Our main research objective was the analysis of the diverse functional designs of online rating systems leading to a consolidated set of design options. The visual design was not in scope of our analysis. Our research process included the following six steps:

- Selection of websites to be analyzed
- Identification of rating systems that are part of the online service
- Iterative design of the criteria catalog
- Appraisal of the online services according to the criteria
- Analysis of the appraisal data
- Consolidation of criteria into a morphological box of design options.
Selection of websites to be analyzed

Our initial literature review did not reveal any reference study that has analyzed rating systems in respect to their functional design. Consequently, we needed to decide on an appropriate sample of websites to be analyzed. The selection of the sample was supposed to be transparent and objective with enough rating mechanisms needed for a thorough analysis.

To meet these requirements, we selected the Top50 websites of the Alexa Traffic Ranking (April 2008; see Table 1). This popular ranking is basically derived from users that have added the Alexa Toolbar to their Browser. Although the traffic ranking is only an approximation, it provides a suitable means for the identification of the most popular websites. Due to adult contents and contents limited to languages apart from German or English, we excluded thirteen websites from the initial sample.

Table 1. Overview of Alexa Top50 (April 2008)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Website</th>
<th>Type of Service</th>
<th>Reason for Exclusion</th>
<th>Rank</th>
<th>Website</th>
<th>Type of Service</th>
<th>Reason for Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yahoo.com</td>
<td>Web portal</td>
<td></td>
<td>26</td>
<td>Photobucket.com</td>
<td>Picture portal</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Google.com</td>
<td>Search engine</td>
<td></td>
<td>27</td>
<td>Google.com.br</td>
<td>Search engine</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Youtube.com</td>
<td>Video portal</td>
<td></td>
<td>28</td>
<td>Amazon.com</td>
<td>E-Commerce portal</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Live.com</td>
<td>Search engine</td>
<td></td>
<td>29</td>
<td>Imdb.com</td>
<td>Film rating portal</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Mix.com</td>
<td>Web portal</td>
<td></td>
<td>30</td>
<td>Visitable.ru</td>
<td>Social network</td>
<td>No English or German language</td>
</tr>
<tr>
<td>6</td>
<td>Myspace.com</td>
<td>Social network</td>
<td></td>
<td>31</td>
<td>Google.it</td>
<td>Search engine</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Wikipedia.org</td>
<td>Online encyclopedia</td>
<td></td>
<td>32</td>
<td>Google.es</td>
<td>Search engine</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Facebook.com</td>
<td>Social network</td>
<td></td>
<td>33</td>
<td>Google.cn</td>
<td>Search engine</td>
<td>No English or German language</td>
</tr>
<tr>
<td>9</td>
<td>Blogger.com</td>
<td>Blog community</td>
<td></td>
<td>34</td>
<td>Imageshack.us</td>
<td>Picture portal</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Yahoo.co.jp</td>
<td>Web portal</td>
<td>No English or German language</td>
<td>35</td>
<td>Youporn.com</td>
<td>Video portal</td>
<td>Pornographic content</td>
</tr>
<tr>
<td>11</td>
<td>Orkut.com</td>
<td>Social network</td>
<td></td>
<td>36</td>
<td>Wordpress.com</td>
<td>Blog portal</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Rapidshare.com</td>
<td>Online storage service</td>
<td></td>
<td>37</td>
<td>Google.co.jp</td>
<td>Search engine</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Baidu.com</td>
<td>Search engine</td>
<td>No English or German language</td>
<td>38</td>
<td>Yandex.ru</td>
<td>Web portal</td>
<td>No English or German language</td>
</tr>
<tr>
<td>14</td>
<td>Microsoft.com</td>
<td>Enterprise homepage</td>
<td></td>
<td>39</td>
<td>Flickr.com</td>
<td>Picture portal</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Google.co.in</td>
<td>Search engine</td>
<td></td>
<td>40</td>
<td>Friendster.com</td>
<td>Social network</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Google.de</td>
<td>Search engine</td>
<td></td>
<td>41</td>
<td>Skyrock.com</td>
<td>Social network</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Qq.com</td>
<td>Web portal</td>
<td>No English or German language</td>
<td>42</td>
<td>Adultsearchfinder.com</td>
<td>Social network</td>
<td>Pornographic content</td>
</tr>
<tr>
<td>18</td>
<td>Ebay.com</td>
<td>Auction portal</td>
<td></td>
<td>43</td>
<td>Go.com</td>
<td>Web portal</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Hill.com</td>
<td>Social network</td>
<td></td>
<td>44</td>
<td>Odnoklassniki.ru</td>
<td>Social network</td>
<td>No English or German language</td>
</tr>
<tr>
<td>20</td>
<td>Google.fr</td>
<td>Search engine</td>
<td></td>
<td>45</td>
<td>Google.com.mx</td>
<td>Search engine</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Aol.com</td>
<td>Web portal</td>
<td></td>
<td>46</td>
<td>Bbc.co.uk</td>
<td>News portal</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Mail.ru</td>
<td>Web portal</td>
<td>No English or German language</td>
<td>47</td>
<td>Craigslist.org</td>
<td>Advertising portal</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Google.co.uk</td>
<td>Search engine</td>
<td></td>
<td>48</td>
<td>Dailymotion.com</td>
<td>Video portal</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Sina.com.cn</td>
<td>Web portal</td>
<td>No English or German language</td>
<td>49</td>
<td>Redtube.com</td>
<td>Video portal</td>
<td>Pornographic content</td>
</tr>
<tr>
<td>25</td>
<td>Fz2.com</td>
<td>Online storage service</td>
<td>No English or German language</td>
<td>50</td>
<td>Cnn.com</td>
<td>News portal</td>
<td></td>
</tr>
</tbody>
</table>

Identification of rating systems that are part of the online service

Having selected the sample, the next step was to identify the rating systems on these websites. We had to consider that the Alexa Top50 includes portals like e.g. Yahoo.com which offer a huge range of separate online services. To avoid accessing every single page of such portals, we restricted our search to services that were linked on the home pages. Some services provided further sub-services which we also only included into our search when linked at the service home page. Where possible, we did at least one rating at each system and analyzed several ranked entities. Altogether, we identified 102 rating systems by searching the online services.

Iterative design of the criteria catalog

In order to gather the characteristics of each rating system systematically and to allow for comparisons, we iteratively elaborated a list of 237 criteria. The criteria are mainly related to functional aspects of rating systems but also cover the environment in which they are embedded. Lacking an existing list of criteria for this particular purpose, we started to derive relevant criteria by a literature review. We continuously extended the criteria catalog in the course of analyzing the broad range of existing online rating systems. We designed each criterion in a binary manner, i.e. a system either possesses the characteristic or it does not.
Appraisal of the online presences according to the criteria

We applied the list of criteria to each rating system and assigned the matching values. In some cases it was not possible to determine if a system possesses a certain characteristic. Therefore, we also assigned ‘not applicable’, and ‘not testable’. A criteria is ‘not applicable’ if its prerequisites are not met, i.e. it is not possible to edit a review comment when the system just allows to rate an item. ‘Not testable’ was assigned when we could not determine the value without doubt, e.g. because of missing documentation.

Analysis of the appraisal data and consolidation of the criteria

Due to page restrictions we can only present exemplary excerpts of our appraisal results in the following section. We already ordered these excerpts according to the key dimensions of the morphological box we will present in section 5. The morphological box (Zwicky 1969) was developed through a critical discussion of the long list of criteria and cross consistency assessments by all authors in order to aggregate and structure the appraisal results (Bailey 1994).

4 ANALYSIS OF THE STATE-OF-THE-ART

The following presentation of our analysis results is structured along six key dimensions of the functional design of rating systems. The first dimension is the service provider who operates the rating system. Secondly, we identified which kinds of entities are rated. Furthermore, we had a closer look on how the rating is entered and aggregated. Finally, we analyzed the representation of the ratings as well as the incentives provided for reviewers in order to evaluate entities.

4.1 Service Providers

In our sample, the range of service providers that use rating systems is very broad. For example, we analyzed producers, retailers, review sites and broadcasters. The service provider diversity goes along with their diverging objectives. Consequently, rating systems can fulfill different functions. For our sample we were able to identify six different functions of rating systems as visualized in Table 2.

<table>
<thead>
<tr>
<th>Functions of Rating Mechanisms</th>
<th>No. of Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product evaluation at price/product portals</td>
<td>5</td>
</tr>
<tr>
<td>Sales promotion</td>
<td>8</td>
</tr>
<tr>
<td>Partner selection</td>
<td>9</td>
</tr>
<tr>
<td>Rating of information</td>
<td>16</td>
</tr>
<tr>
<td>Neutral rating mechanisms</td>
<td>27</td>
</tr>
<tr>
<td>Community rating</td>
<td>37</td>
</tr>
</tbody>
</table>

Table 2. Functions of Rating Systems

77 of our 102 investigated rating systems focused on the central entity of the service, i.e. it is possible to rate the entity that the website was actually intended for. Moreover, 12 of these 77 systems strongly depend on the rating systems to fulfill the intended website service. For example, Pixnay.bebo.com is a social community that is limited to people rating each other’s pictures.

4.2 Entity

The entity can be seen as the central element of each rating system as it determines most of the design parameters. Concerning entity types we distinguished between subjects, abstract subjects and objects (see Figure 1). Subjects are active actors, in general other users. Objects are e.g. goods or pieces of information. The concept of an abstract subject means that an object or event is rated but the aggregation is assigned to a subject. For instance, at Ebay.com buyers rate a seller through transactions made and ratings are linked to the user profile of the buyer although actually an event (transaction process) was rated. The information behind this kind of ratings is limited to the specific situation and therefore not directly transferable to other contexts and situations.
Some rating systems allow the rating of different entity types. For example, some city guide portals allow the rating of subjects (companies) and objects (points of interest). Most systems (78) offer the assessment of objects only. Ratings about products or services prevail. The object usually does not belong to the website owner (this is only the case for 11 systems) but is provided by users (29) or external third parties (48).

### 4.3 Gathering of Ratings

The gathering of ratings should be designed in a way that it allows users to enter ratings as easily as possible. However, 84 of the analyzed systems required an enrolment for the online service. About two third of these systems (62) asked for a valid email-address. At 11 systems, users had to additionally qualify for ratings, e.g. through membership (2) or through a specific member status (3; e.g. reaching a certain score at Answers.yahoo.com). In other cases (6) the accomplishment of a certain transaction is a pre-requisite for rating, e.g. at Ebay.com.

In 16 cases, the user (7 cases) or some other stakeholder is able to restrict the rating of entities. For example, at Pixnay.bebo.com a registered user is able to decide whether other users are allowed to rate his/her picture or not.

The rating entry forms of the service providers vary. Some are very detailed and therefore require a lot of input from the user. Furthermore, some allow a meta-rating such “X out of Y users found the rating very helpful”. However, most turn out to be very simple with just one or few rating options.

### 4.4 Aggregation of Ratings

93 out of 102 systems consolidate single ratings to one holistic score (see Table 3). This can be problematic as a consolidation reduces the rating information (e.g. it blurs the differences between positive and negative ratings). Some services such as ‘Karma’ at Orkut.com allow the entering of single values such as trusty, cool and sexy without consolidating the assessments. Furthermore, 8 systems display the number of ratings for each value (e.g. 17 positive, 3 neutral, and 23 negative ratings).
Table 3. Aggregation of Entities

78 rating systems compute the aggregated score as an average rating value (see Table 3). The other quarter of rating systems uses a cumulative approach such as counting of votes (12), cumulating based on scale values (5) or cumulating with subtraction of negative ratings (4) as well as relative values (1). Some websites such as Propeller.com or Epinions.com use further information for the calculation of the scores. At Propeller.com, the score is also based on comments and the number of users that have read the article. The exact calculation is not explicitly described to avoid manipulations. At Epinions.com, a score is additionally based on the number and the age of ratings.

4.5 Representation of Ratings

The representation determines the explanatory power of ratings (Sabater-Mir & Paolucci 2007) and relies on the data input (Dellarocas & Wood 2008). 89 systems work with ratings that are visible for all users (see Table 4). Of those 13 remaining, 4 are only visible for registered users and 9 rating systems require a qualification prior to accessing the ratings. In most cases (85), the total number of ratings is quoted on the profile page of the rated entity (e.g. in the user profile, product details, etc.). However, 28 rating systems use ratings for enriching search results on individual entities. 28 systems display the number of ratings for each scale values such as 17 users rated “very good”, 24 users rated “good”, etc.
The providers of 32 systems state that they approve each rating (see Table 5). E.g. Expedia.com claims to take up to 10 days for checking each evaluation. Our tests indicated that 12 out of these 32 systems did not approve the ratings but allowed an immediate visibility of our judgments. 69 systems allowed an immediate visibility of ratings without any editorial approval.

Some services provide a sorting or selection mechanisms. A good example is the filter mechanism of IMDb.org that provides extensive options for selecting movie ratings of specific rater groups (see Figure 2).

![Figure 2](image)

**Figure 2. Selection of ratings at IMDb.org**

### 4.6 Incentive Schemes

Incentive schemes can be used for encouraging users to rate specific entities. The reason behind this is the fact that a rating provides value for the consumer rather than for the evaluator. The visibility of login names or real names as an incentive to excel evaluators is only used by 51 systems (see Table 5). 26 systems allow a linkage to the individual profile page of the evaluator. 22 systems enable an aggregation of ratings within the own profile. Furthermore, a rating of evaluators and specific labels are used by a fraction of services. Only two online services, Epinions.com and Orbitz.com, offer monetary compensation for rating efforts. Refunding at Epinions.com is based on the number of users that read the rating. Orbitz.com offers coupons to each reviewer.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>No. of Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visibility of evaluator per rating</td>
<td></td>
</tr>
<tr>
<td>Visibility of login name</td>
<td>Yes 26</td>
</tr>
<tr>
<td>Visibility of real name</td>
<td>No 23</td>
</tr>
<tr>
<td>Visibility of login name</td>
<td>Not applicable 24</td>
</tr>
<tr>
<td>Visibility of real name</td>
<td>Not testable 3</td>
</tr>
<tr>
<td>Visibility of given ratings within own profile</td>
<td>22 4 76 0 102</td>
</tr>
<tr>
<td>Ranking of evaluators</td>
<td>Yes 5</td>
</tr>
<tr>
<td>Usage of labels</td>
<td>No 17</td>
</tr>
<tr>
<td>Expert labels</td>
<td>Yes 9</td>
</tr>
<tr>
<td>Character labels</td>
<td>No 21</td>
</tr>
<tr>
<td>Monetary compensation of rating effort</td>
<td>Yes 2</td>
</tr>
<tr>
<td>Character labels</td>
<td>No 100</td>
</tr>
<tr>
<td>Visibility of given ratings within own profile</td>
<td>6 3 93 0 102</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criteria</th>
<th>No. of Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visibility of evaluator per rating</td>
<td></td>
</tr>
<tr>
<td>Visibility of login name</td>
<td>Yes 26</td>
</tr>
<tr>
<td>Visibility of real name</td>
<td>No 23</td>
</tr>
<tr>
<td>Visibility of login name</td>
<td>Not applicable 24</td>
</tr>
<tr>
<td>Visibility of real name</td>
<td>Not testable 3</td>
</tr>
<tr>
<td>Visibility of given ratings within own profile</td>
<td>22 4 76 0 102</td>
</tr>
<tr>
<td>Ranking of evaluators</td>
<td>Yes 5</td>
</tr>
<tr>
<td>Usage of labels</td>
<td>No 17</td>
</tr>
<tr>
<td>Expert labels</td>
<td>Yes 9</td>
</tr>
<tr>
<td>Character labels</td>
<td>No 21</td>
</tr>
<tr>
<td>Monetary compensation of rating effort</td>
<td>Yes 2</td>
</tr>
<tr>
<td>Character labels</td>
<td>No 100</td>
</tr>
</tbody>
</table>

**Table 5. Incentive Schemes**

### 5 DISCUSSION OF RESEARCH RESULTS

Our analysis – based on 237 criteria applied to 102 different rating systems – shows a wide variety of characteristics among rating systems. In our sample, a quarter of all systems are used as a means to gather and provide user information for specific entities e.g. a certain type of car. The support of sales activities and product comparison are common areas of use. There are some elements that are similar in nearly all rating systems. However, it is not possible to identify a best practice approach that is suitable for every situation. Every type of entity and business model has its own requirements on rating systems.

Hence, we will discuss the key parameters and parameter values we derived from our findings. We included these key parameters into a morphological box (Zwicky 1969). They were determined
through a critical discussion of all identified criteria as well as cross consistency assessments by all authors (Bailey 1994). The goal was to aggregate and structure the results of our analysis.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provider</td>
<td>Commercial</td>
</tr>
<tr>
<td>Entity</td>
<td>Non-commercial</td>
</tr>
<tr>
<td>Type</td>
<td>Subject</td>
</tr>
<tr>
<td>Reference to service</td>
<td>Abstract subject</td>
</tr>
<tr>
<td>Integrity (Encapsulation)</td>
<td>Object</td>
</tr>
<tr>
<td>Specific</td>
<td>Service area only</td>
</tr>
<tr>
<td>Unspecific</td>
<td>Domain</td>
</tr>
<tr>
<td>Provider spanning</td>
<td>Domain-spanning</td>
</tr>
</tbody>
</table>

**Figure 3. Conceptualization of Design Options**

The searched online services are of either commercial or non-commercial interest. It turned out that 100% of all analyzed rating systems have a commercial background. A further analysis of non-commercial services may add to our framework, although we believe that their rating systems will deal with the same entities and objectives.

The rated entities turned out to be either subjects such as individual people, abstract subjects or objects. In most cases (77%) the rated entities had been objects such as hotels, products or services. Rating systems for objects are usually not restricted to certain brands. Only at YahooShopping.com we found that it is possible to rate some shoe brand, while it is not possible to rate other brands.

Furthermore, providers use rating systems not only for their specific service but also in other service areas or even across various providers. However, there is no standardized mechanism for exchanging rating information between platforms. With the emergence of Open Social (OpenSocial.org) at the end of 2007, which is a standardized API for social applications across multiple websites, we expect the interaction among different platforms will increase within the next years.

Linking the permission for rating to certain prerequisites can improve the rating quality (Dellarocas 2000). In most cases, the user group is restricted (82%). Hence, it is possible to automatically validate whether a user is authorized to conduct a rating or not. However, there are also some rating systems which do not require a user registration. The decision whether to provide an user-restricted or unrestricted rating system very much depends on the acceptance of people, the general business model behind the website and the importance of the rating. If, for instance, the user has to be registered in order to use services anyway, a restricted rating system is preferable. But if the rating system is the main part of the service (e.g. Pixnay.com’s picture rating), a unrestricted rating system might limit the acceptance of such a platform. In some contexts it may even be sensible to set up further prerequisites before a rating is admitted. For example, at Ebay.com people have to conduct a purchase first before they are able to rate the seller of the product. Qualification can also be useful in areas of high interest but low general knowledge. Hence, a qualification mechanism is sensible for websites that need specific expert knowledge or for social networks where only friends are allowed to rate other friends.
We found that most ratings are direct ratings that express an opinion directly (‘I like Mandy’, ‘This hotel receives 5/5 stars’, etc.). For the provider, this sort of rating is very easy to aggregate and calculate. However, there are academic discussions and studies about problems with fraud (Gregg & Scott 2006). In addition, a bilateral direct rating may lead to revenge ratings. Even at Ebay.com, the old bilateral direct rating system, which was restricted to transactions, turned out to be very problematic and has been changed to a more or less unilateral rating system recently. In few cases we also found some relative rating systems on the internet where ratings are made based on the comparison with other entities (Dellarocas 2003). Unfortunately, such an approach is only possible in social applications with a high degree of cross linking (Botsch & Luckner 2008). In general, relative rating systems can help overcome some (but not all) fraud issues since a rating is only possible in comparison to other subjects or objects. Nevertheless, this kind of rating systems is not applicable to all services. For instance, relative ratings may not be helpful in situations where the entity cannot be compared to other objects (e.g. other services) or the evaluator only knows a limited number of comparable entities (e.g. comparing a hotel to other hotels in this region).

Providers of rating systems tend to aggregate individual ratings in order to receive an overall score although they risk losing relevant individual information. For example, an average mark does not give any information about the number of positive and negative ratings. Hence, it is sensible to display ratings for each score value (e.g. for positive, neutral, and negative ratings, see Google.com/products as an example). Since many ratings systems calculate scores out of single ratings, the method for this calculation is one of the key design options.

Most aggregation approaches for ratings have proven to be very simple. Usually, all ratings counted equally. Simple Systems compute an average of all ratings or a total score as the sum of positive scores minus the sum of negative scores. Advanced methods compute a weighted average of the ratings, where the rating weight can be determined by different factors, e.g. trustworthiness of the rating participant, age of the rating or distance between rating and current score. Furthermore, even more sophisticated statistical methods can be applied. In almost any case (98%), the aggregated score is only based on individual user ratings but there are also other factors that might influence ratings (e.g. at Epinions.com).

Retrieving ratings or scores is usually open to everybody that accesses the according online service. The high number of generally visible ratings indicates the importance of ratings for individual business models. E.g. e-commerce websites heavily rely on user ratings in order to offer additional value to potential buyers.

The need for incentives to attract ratings largely depends on the general activity and involvement of users on the platform and hence the business model itself. For example, pure rating platforms like Pixnay.com do not need any form of incentive for rating at all. In contrast, platforms that sell services that have to be used before rating them (e.g. hotel stays, craftsmen services, etc.) will need incentives in order to persuade users to return again to their websites for rating.

The social aspect of profile building and reputation is a basic incentive that is very popular among online services. These services either use the real name of a person or a pseudonym. In case of using the real name there will be a connection to the real world. Depending on the service, this can be an advantage (e.g. the user can prove that he is an expert in his profession) or a disadvantage (users fear a loss of privacy). Advanced incentives can be of monetary or other beneficial nature. For example, Epinions.com offers money for reviews. Some services offer indirect incentives such as the opportunity to participate in a raffle as a reward for entity rating. Hence, they provide additional value to the user with relatively low effort.
6 LIMITATIONS AND FUTURE RESEARCH

Our analysis is based on a large sample of websites and their rating systems. However, due to language barriers we were only able to analyze German and English websites. Furthermore, we only considered the most popular ones. There may be a possibility that there will be other rating systems in other cultural regions (e.g. Asia) that would additionally contribute to our body of knowledge. At the end, our criteria catalog was made out of 237 criteria and we believe it to be fairly comprehensive. Nevertheless, the analysis of additional rating systems might add even further criteria.

Unfortunately, we were only able to test the rating algorithms from the user perspective and could not look behind the curtains. However, we are convinced that our observations are quite accurate. Future research should also include less popular websites in order to understand differences between rating systems on high traffic websites and low traffic websites.

Today, most rating systems are on objects rather than subjects. We believe that the emergence of social networks and social applications will change the focus of rating systems towards users. Social networks may offer new rating systems within the next years and will raise additional questions regarding data privacy, self-determination, etc. Furthermore, new APIs such as Open Social may change the way of interaction among websites and hence the design of cross-platform rating systems. We believe our recent “snapshot” of the state-of-the-art of rating systems to be a valuable contribution to these and further emerging questions that have to be addressed by future research.

References


CONTROLLING COMPUTER-BASED MULTITASKING THROUGH PROVISIONING SYSTEMS IN CO-LOCATED LEARNING SETTINGS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0138.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Behavior, Behavioural theories, Usability, IT/IS management</td>
</tr>
</tbody>
</table>
CONTROLLING COMPUTER-BASED MULTITASKING THROUGH PROVISIONING SYSTEMS IN CO-LOCATED LEARNING SETTINGS

Truman, Gregory, Information Systems Area, IE Business School, Maria de Molina 12, 4 Derecha, 28004 Madrid, Spain, gregory.truman@ie.edu

Abstract

Computer-based multitasking behavior has become commonplace, however its efficacy in co-located settings is disputed. An important question therefore is how computer-based multitasking can be controlled when traditional organizational control mechanisms are infeasible or unavailable. We conceptualize and operationalize computer-based multitasking in terms of on-task and off-task uses. Our research objective is to examine a provisioning system’s effectiveness with respect to restraining off-task uses while leaving on-task uses unaffected in a co-located learning setting. We conclude that provisioning systems alone are not sufficient to effectively control computer-based multitasking—to restrain off-task use, so it may be advisable to augment provisioning systems with behavioural modification and reform efforts.

Keywords: Multitasking, provisioning systems, co-located learning, technology-mediated learning, e-learning, agency theory, longitudinal, field study
1 INTRODUCTION

Technological advances related to information and communication technologies are transforming learning processes in various contexts (Hiltz and Turoff 2005, Alavi and Gallupe 2003, Alavi 1994, Young 2005, Arbaugh 2000). Of assorted nature, these transformations have generated interest among information systems, management education, education and education psychology researchers (Balkovich, Lerman, and Parmelee 1985), and have created a field of inquiry called technology-mediated learning (TML). In this paper, we follow Leidner and Jarvenpaas’ (1995) characterization of TML as learning experiences that are modified in some way through use of information and communication technologies.

TML learning is facilitated by applications such as email, chat, blogs, virtual whiteboard, video-conferencing, document management, assessment facilities, and the Internet in general among others (Tu 2000, Navarro and Shoemaker 2000, Ross and Schulz 1999). These applications may be combined into one system or “package” (Arbaugh 2000, Hiltz and Wellman 1997), which we refer to here as technology-mediated learning applications (TMLAPP). TMLAPP running on contemporary computing platforms are expressly designed to facilitate computer-based multitasking. As a result, computer-based multitasking behaviour has attained a commonplace nature in both virtual and co-located learning settings.

In this paper, we are specifically interested in examining computer-based multitasking in co-located learning settings. We argue that the co-located setting is critical to study, because it is this setting that appears to heighten contention over both the appropriateness and efficacy of computer-based multitasking behaviour (Gomolski 2006, Wallis et al. 2006, Chudoba et al. 2005, Wasson 2004, Shellenbarger 2003, Hwang 2002). While the reasons for this are likely manifold and not entirely understood, one can reasonably claim that co-located settings produce behavioural norms with respect to participants’ attention and effort. In co-located learning settings, where people organize with the intention of performing some learning task, behavioural norms produce expectations that each participant will give undivided attention and effort. Since computer-based multitasking in co-located learning settings is palpable, and may suggest—although is not definitive of—diversion of attention and effort, contention over multitasking appropriateness and efficacy heightens. For instance, approaching the issue from an anthropological perspective, Wasson (2004) concludes that multitasking behaviour may require or effect change in standards of politeness to the extent that it becomes culturally embedded.

2 LITERATURE REVIEW

Currently, there is no consensus regarding the efficacy of engaging in computer-based multitasking in co-located settings. Some suggest that a potential of bringing TMLAPP into co-located settings, especially ones of large numbers of participants, is to establish back-channel conversations, where participants’ attention and effort may be stimulated through enhanced content sent over back-channels. On the other hand, some question whether it is possible to concurrently multitask and pay attention to others in effective ways (Gomolski, 2006). Due to its very nature, multitasking behaviour may create distraction effects for other participants. In fact, in some co-located contexts computing devices or Internet access are altogether banned (Guernsey, 2003).

Consequently, an important question is—how can computer-based multitasking be controlled such that the attending advantages are maximized and disadvantages minimized? In this paper, we conceptualize and operationalize use behaviour in terms of on-task and off-task uses vis-à-vis the use context. Our research objective is to examine the effects and effectiveness of employing a
provisioning system (PS) for the intended purpose of restraining off-task uses. We utilize agency theory to frame our hypotheses and we use longitudinal data from a quasi-experimental field study to operationalize objective measures on actual use in a co-located setting.

2.1 Controlling use through provisioning systems

In this study and for purpose of subsequent discourse, we assume that at least one co-located participant assumes the role of facilitator, who is responsible for organizing and coordinating participants’ activity directed to task completion. Additionally, we define participants on-task (off-task) use behaviours as those that are (are not) related to the task at hand. Ideally participants will remain on-task however anecdotes and commonplace experience suggest that they will engage in off-task uses as well. While computing devices may be banned altogether, this is often not desirable or feasible for a variety of reasons. As an alternative and in an effort to restrain off-task uses, a facilitator may employ a dynamic provisioning system (PS) to limit participants’ access to certain applications and devices. For example a facilitator may allow access to shared file servers and printers within the local network, while disallowing network-based applications such as email and chat. By selectively allowing some applications and devices while eliminating others, the practical value of computing devices in co-located contexts is presumably enhanced by balancing on-task uses against off-task ones.

In practice, the options of prohibiting computing device use altogether or allowing unimpeded use represent end-points on a continuum of various provisioning levels, each allowing access to some applications and devices to varying degrees. The assorted provisioning levels that are made available are determined by the PS’s design, which may be customized for the user entity. Generally speaking a PS enables the facilitator to specify infrastructure or application service provisioning. The PS that we examine here offered provisioning options for email, shared (local) file servers, local printing, an Intranet application and the Internet. Each service could be denied or allowed, although the five provisioning options were not entirely independent of one another. For purpose of subsequent discourse, we use the term ‘access setting’ to refer to the applications and devices that are provisioned during co-located sessions. We qualify an unrestricted (restricted) access setting as less (more) impeded—allowing participants access to relatively more (fewer) applications and devices.

3 THEORETICAL FRAMEWORK

We argue that a generalized co-located context involving at least one facilitator and participant may create conditions that satisfy agency theory assumptions (Eisenhardt 1989, Eisenhardt 1985, Ouchi 1979). We predicate this argument on others’ general assertions that agency theory is applicable to a broad range of organizational design issues with respect to control mechanisms (Eisenhardt 1985, Ouchi 1979). More specifically Eisenhardt (1989) states that "The agency structure is applicable to a variety of settings, ranging from macro-level issues […] to micro-level dyad phenomena such as blame, impression management, lying, and other expressions of self-interest. (italics ours, p.58)", and concludes that researchers may "...use agency theory in their study of [a] broad range of principal-agent issues facing firms. (p.72)"

Broadly characterized, agency theory may be applied to any dyad relationship that involves cooperative behaviour between its members. Commonly members assume roles that exist in superior-subordinate, supervisor-worker or leader-apprentice arrangement with respect to one another. In addition the relationship typically involves the principal in delegation of a task to an agent, and in subsequent control over agent’s task performance and/or outcomes. A frequent assumption about the principal-agent relationship is that the principal and agents’ objectives or interests with respect to performance or outcomes will differ, diverge or depart. Consequently the principal’s primary problem lies in extracting the agent’s cooperative behaviour so that behavior serves the principal’s objectives.
In reviewing agency theory's contributions to organization theory, Eisenhardt (1989) identifies similarities and differences between two research lines—positivist agency theory and principal-agent research. Both lines share assumptions about people, organizations and information. For instance, agency theory assumes that human behaviour and decision-making is driven by self-interest. Second, there exists at least partial goal conflict between principals and agents. Third, there exists information asymmetry between the principal and agent such that the principal cannot fully monitor the agent's behaviour and actions. Intentionally absent any mathematical rigor, our application of agency theory falls into the realm of positivist agency theory, where researchers are generally focused on identifying situations where the principal and agent have conflicting goals, and on describing control or governance mechanisms that restrict the agent's self-serving or opportunistic behavior. Positivist research broadly identifies two control or governance mechanisms for curbing opportunistic behaviour — outcome-based contracts and information systems (Eisenhardt 1989).

We argue that a co-located learning setting may create a situation where the goals of the facilitator (principal) and of the participants (agents) begin to conflict or diverge. For example, we assume that the facilitator has interest in engaging, soliciting input from, or imparting information to participants. Said differently, the facilitator wants the participants' attention and effort focused on the learning task in order to facilitate the learning process and to yield successful learning outcomes. While participants may fully comply with the facilitator's wishes, their behaviour may succumb to self-serving interest as they direct attention and effort to other work or personal tasks or leisure activity. To the extent that participants' engage in self-serving behaviour, co-located facilitator's and participants' goals begin to diverge or conflict.

The facilitator is able to monitor some of the participants' behaviour. However, from a practical perspective, the facilitator's monitoring capacity is constrained by assorted factors such as the number of participants, the nature of participants self-serving behavior, the facilitator's cognitive abilities, and the participants' physical arrangement among others. In short, there is in general information asymmetry between facilitators and participants with respect to participants' attention and effort and, consequently, to facilitator's and participants' goal alignment. Therefore when participants can engage in self-serving behaviour with minimal concern or regard for detection, they are likely to do so.

With the introduction of computing devices into co-located learning settings, participants' ability to pursue self-serving behaviours is significantly expanded because much work and personal tasks and leisure activity is computer-mediated. For instance, a participant may check the value of personal stock holdings, compose a letter to a relative, or check the latest standings of local professional sports teams. To the extent that participants' self-serving behavior increases, agency theory predicts that goal conflict between facilitator and participants increases.

The facilitator's ability to monitor participants' behaviour is further compromised by introduction of computing devices. Assume a simulated learning task where participants take on the role of institutional investors at a financial brokerage firm and the task of revising company-wide investment guidelines. A participant may use the computing device to look up on-line information about investment instrument prices or holdings in order to provide input to the task. This use behaviour is on-task. In contrast that same participant may look up current information about his own investments. This use behaviour is off-task. From the facilitator's perspective, the participant's computing device use is virtually indiscernible with respect to on-task or off-task usage, unless the facilitator has view of the participant's computing device screen which is seldom the case. In short, because the difference between a participant's on-task and off-task uses is effectually indiscernible for the facilitator, information asymmetry effectively increases thereby exacerbating the monitoring problem.

In summary, from an organization design perspective, control or governance mechanisms ameliorate information asymmetry between the co-located facilitator (principal) and participants (agents) and thereby bring about goal alignment between them. Governance mechanisms take on at least two forms according to the positivist's research realm—outcome-based contracting and information systems (Eisenhardt 1989). The specific choice between these is generally guided by efficiency.
considerations—i.e., which control mechanism is the least costly to implement. In co-located learning settings, neither outcome-based contracting nor information systems is a feasible alternative from a practical perspective.

Where these control or governance mechanisms are impractical, an alternative way to bring about goal alignment is through socialization (or clan) structures (Ouchi 1979). While socialization structures are many and varied in practice, on a conceptual level we envision socialization processes that correspond to Ouchi's (1979) description of "…a highly formalized and lengthy period of socialization during which [participants] are subjected not only to skill training but also to value training or indoctrination. (italics ours, p.837)."

Absent any outcome-based contracting, information systems or socialization structures for attaining goal alignment between principal and agent, a co-located learning setting may offer the option of utilizing provisioning systems (PS) that can effect control through "brute-force." For instance, under the assumption that the facilitator wants to discourage participants’ off-task behaviour, the facilitator can dynamically provision devices and applications so as to restrict access to select ones. By restricting use of some devices and applications for some duration of the learning task, the facilitator will effectively reduce the potential for participants' off-task behaviour. In short, the facilitator can provision a restricted access setting, which should lead to reduced off-task use levels among participants. We submit the following hypothesis.

**Hypothesis 1:** Off-task use behaviors are lower under a restricted access setting when compared to an unrestricted access setting.

By provisioning a restricted access setting, the facilitator is restricting specific devices and applications from participants’ use. While this should have the intended effect of reducing off-task use, the devices and applications are no longer available for on-task use either. Since devices and applications can be used for both purposes, one may argue that on-task use behaviour will also be reduced. However, dynamic provisioning should allow the facilitator to provision an unrestricted access setting on opportune occasion, thereby providing access to devices and applications when their use is efficacious to the learning task's objectives and activities at hand. Therefore as another goal of facilitators is presumably to leave participants’ on-task behaviour unimpeded or unaffected, we argue that the impact on participants' on-task uses should be negligible assuming that the facilitator dynamically provisions an unrestricted access setting on opportune occasions. We submit the following hypothesis.

**Hypothesis 2:** On-task use behaviors are not significantly different between an unrestricted access setting and a restricted access setting.

### 4 RESEARCH METHODOLOGY

#### 4.1 Research setting

The research design is a longitudinal quasi-experimental field study that involves subjects who enrolled in a required freshmen course at a business college in the Northeastern United States. The setting is a two-semester integrated course that presents a multidisciplinary perspective by interweaving topics and concepts from different academic disciplines (or functional areas) into one course, which we refer to here as Business Introduction. The pedagogical approach of this course is heavily applied, and much of the curricula was designed and delivered under the assumption that each participant would have a dedicated laptop computer during co-located sessions. Business Introduction is offered in seven sections, and each section is capped at 60 participants. Two faculty members teach each section. In January 2006, the course received an Excellence in Education Award from the U.S. Association for Small Business and Entrepreneurship (USASBE) for its innovative and entrepreneurial pedagogical approach.
4.2 Research design

For this study, subjects came from four Business Introduction sections, which were instructed by two lecture teams. Each team instructed one section in unrestricted access and another in restricted access to avoid confounding effects between lecture team and access restrictions. Unrestricted access allowed use of all applications including course management application, shared file servers and printers, the email system and the Internet. Restricted access prohibited network applications such as Internet browsing, email and chat, except on occasions where the facilitators deemed them to be pedagogically relevant. This exception normally involved allowing Internet access for some learning sessions. In addition, lecture teams alternated between sections to control for order effects and eliminate the second section advantage of receiving the lecture after it had been delivered to the first section.

The lecture teams shared teaching materials. Their efforts to ensure uniform course requirements across sections resulted in a relatively standardized curriculum. However, curriculum delivery was not scripted and was subject to idiosyncratic lecturing styles and the emergent nature of discussion periods. Facilitators were generally non-directive in controlling participants' computer use however. The subjects were free to use their laptop without limitations under unrestricted access setting and with email, chat and Internet access limitations in the restricted access setting.

4.3 Sampling

Taught by two facilitators, each of seven sections capped enrolment at 60 participants. We solicited four sections for participation, and 122 subjects volunteered to participate. This represents a 51% (=122/240) participation rate. From these 122 subjects, we obtained 95 monitoring log files and 89 surveys of which 72 overlapped. Four subjects switched sections after term onset, so we have primary data on 68 subjects.

Although the subjects were not informed about the specific research questions, they were made aware that participation involved unobtrusive monitoring of their computer use during Business Introduction sessions (primary data), filling out a survey (primary data), and releasing course performance data (secondary data). Monitoring occurred only during the first term. While subjects were not randomly assigned to sections, the access setting was designated for each section well after the participant registration period, so no self-selection bias with respect to "getting into" an unrestricted section could have occurred.

4.4 External Validity—Monitoring Awareness

To the extent that monitoring affected subjects’ behaviour, external validity is undermined thereby yielding no generalizable conclusions. Thus, we took several measures to reduce subjects’ monitoring awareness. First, recruitment during a May-June timeframe provided a three to four month lag between consent and onset of monitoring to mitigate awareness effect over use behaviour. The research project was never raised by the facilitators, and subjects were not reminded of their participation until asked to fill out the survey at monitoring conclusion. Finally, the monitoring program was installed and uninstalled through an unobtrusive SMS (server-side) procedure, and it was configured to run in stealth mode—the monitoring program icon never appeared on the desktop, task bar or task list. Log files were copied from the learners’ computers’ hard drives by an undetectable server-side custom program called the “harvest” program. Thus, no subject was ever alerted to the monitoring program.

Responses from a relevant survey item strongly suggest that most subjects did not change their use behaviour in response to being monitored. Eighty-four percent of subjects disagreed with the statement “I changed how I used my computer because I know that I was being monitored.” Fifty-seven percent
strongly disagreed with the statement. In addition, a consensus opinion emerged from the four facilitators, three of whom had previously taught Business Introduction, that subjects' did not alter their use behaviour due to monitoring awareness. Altogether then—the lag between monitoring consent and onset, the monitoring and harvest programs being undetectable, and other efforts to keep the research project at low profile helped to mitigate subjects’ awareness of being monitored.

4.5 Variables and Measures

Subjects’ use was monitored only during co-located sessions, and use activity is logged to the local drive so all use is recorded independent of any network connection. Periodically, the log file was copied from each subject's local hard drive to a file server by a custom-designed server-side program called the “harvest” program. The “harvest” program required that the subject establish a network connection, but that connection could occur at any time. The precise timing of log file “harvesting” was jointly determined through the random occurrence of subject network login and time of last harvest, and generally occurred about every seven to ten days. The harvest program operated inconspicuously.

Sample use logs are shown in Tables 1 and 2. Fields include application name, window title and keystrokes among others. Each time a subject set the window focus, a log record was created. The window focus start and end times are recorded as AppGotFocus and AppLostFocus, respectively. Divided into active and inactive parts, elapsed time is the difference between focus times. Active time generally reflects the amount of time that the subject operated any input device, which in this context includes the keyboard and mouse. Inactive time generally reflects idle time with respect to input device manipulation.

Each sample highlights use that is either predominantly on-task or off-task. Table 1 illustrates mostly on-task use where the subject is taking notes on a problem-solving exercise involving breakeven analysis. Table 2 shows mostly off-task use where the same subject at a different time is in an instant message dialog. (See Keystrokes column.) If the record represents on-task (off-task) use, then the on-task column is coded as ‘1’ (‘0’) and the off-task column is coded as ‘0’ (‘1’). Some records are coded ‘0’ in both columns, because they represent neither on-task nor off-task use, or they are indeterminate with respect to use behaviour. In no case was a record coded as ‘1’ in both columns, because no use can be simultaneously characterized as on-task and off-task. The coding process was partly subjective and not entirely error free, so the data were reviewed by two independent coders. Agreement level was 82% (89,260 out of 109,057), and the records in non-agreement were removed for subsequent analyses.

Four measures each for on-task and off-task uses are computed from the data—focus change, keystroke, active time and inactive time. Focus change is a count on the number of records, and represents the number of times that a subject sets focus to an application (window). Keystroke is the sum on the key counter column, and represents the number of keystrokes made by the subject. The key counter values include nonprintable character key depressions, such as shift, control, alt and function keys among others, despite these being absent in the keystrokes columns. Active (inactive) time is the sum of the active (inactive) time column, and represents the amount of active (inactive) time in seconds. All use measures are ratio scaled.

Aggregate results are shown by on-task, off-task and neither use activities in Table 3, which presents nominal and proportional figures. These results indicate that about two-thirds (one-third) of all use activity is off-task (on-task). Records coded as neither account for a small percentage, and these are removed from subsequent analyses. We also eliminate inactive time due to a critical interpretability problem.

4.6 Data Analysis

Our hypotheses require that we test for differences between groups on a set of ratio-scaled measures.
<table>
<thead>
<tr>
<th>Name</th>
<th>MatAll</th>
<th>Date</th>
<th>AppGotFocus</th>
<th>AppLostFocus</th>
<th>ApplicationName</th>
<th>Keystrokes</th>
<th>WindowFile</th>
<th>AppKeyCounter</th>
<th>ElapsedTime</th>
<th>ActiveTime</th>
<th>InactiveTime</th>
<th>On-task</th>
<th>Off-task</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNXX01</td>
<td>363</td>
<td>30-Sep</td>
<td>9:56:17 AM</td>
<td>9:57:45 AM</td>
<td>WINWORD.EXE</td>
<td>sales-Variable BCost-Fixed Cost= Profit</td>
<td>Quantitative P:1CORR - Microsoft Word</td>
<td>69</td>
<td>00:01:28</td>
<td>00:00:43</td>
<td>00:00:45</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>364</td>
<td>30-Sep</td>
<td>9:57:45 AM</td>
<td>10:04:48 AM</td>
<td>WINWORD.EXE</td>
<td>Purchases (increase Inv.) Cost G= S= (decreases Inv.)  $ - 25(322)</td>
<td>Quantitative P:1CORR - Microsoft Word</td>
<td>258</td>
<td>00:07:03</td>
<td>00:03:02</td>
<td>00:04:01</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>365</td>
<td>10:04:48 AM</td>
<td>10:05:02 AM</td>
<td>30-Sep</td>
<td>Explorer.EXE</td>
<td>Program Manager</td>
<td>0</td>
<td>0:00:00:14</td>
<td>00:00:07</td>
<td>00:00:07</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>366</td>
<td>30-Sep</td>
<td>10:05:02 AM</td>
<td>10:05:05 AM</td>
<td>WINWORD.EXE</td>
<td>Quantitative P:1CORR - Microsoft Word</td>
<td>0</td>
<td>0:00:00:03</td>
<td>00:00:03</td>
<td>00:00:00</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>367</td>
<td>10:05:05 AM</td>
<td>10:06:42 AM</td>
<td>30-Sep</td>
<td>AcrobatReader.exe</td>
<td>Acrobat Reader - [ProblemAssign.pdf]</td>
<td>0</td>
<td>0:01:37</td>
<td>00:00:04</td>
<td>00:01:33</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>368</td>
<td>10:06:42 AM</td>
<td>10:07:53 AM</td>
<td>30-Sep</td>
<td>WINWORD.EXE</td>
<td>Quantitative P:1CORR - Microsoft Word</td>
<td>46</td>
<td>0:01:11</td>
<td>00:00:54</td>
<td>00:00:17</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>369</td>
<td>30-Sep</td>
<td>10:07:53 AM</td>
<td>10:24:55 AM</td>
<td>WINWORD.EXE</td>
<td>No Units Unit Cost $ NOx P 100 $5 500</td>
<td>Quantitative P:1CORR - Microsoft Word</td>
<td>762</td>
<td>00:17:02</td>
<td>00:06:22</td>
<td>00:10:40</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>370</td>
<td>30-Sep</td>
<td>10:24:55 AM</td>
<td>10:25:00 AM</td>
<td>Explorer.EXE</td>
<td>Program Manager</td>
<td>0</td>
<td>0:00:00:05</td>
<td>00:00:00</td>
<td>00:00:05</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>371</td>
<td>30-Sep</td>
<td>10:25:00 AM</td>
<td>10:25:02 AM</td>
<td>aim.exe</td>
<td>AOL(R) Instant Messenger(TM)</td>
<td>0</td>
<td>0:00:02</td>
<td>00:00:02</td>
<td>00:00:00</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>372</td>
<td>30-Sep</td>
<td>10:25:02 AM</td>
<td>10:25:19 AM</td>
<td>aim.exe</td>
<td>XX/02's Buddy List Window</td>
<td>0</td>
<td>0:00:17</td>
<td>00:00:01</td>
<td>00:00:16</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>373</td>
<td>10:25:19 AM</td>
<td>10:32:08 AM</td>
<td>30-Sep</td>
<td>WINWORD.EXE</td>
<td>FCI Units ocontcontributionion= (33120/18 units) Sales - Fixed xed Costs = variable costs= profits DaSales= 33120 -(Sales *.40)=0 sa=Sales=55,200 Sales= 1,840/price</td>
<td>Quantitative P:1CORR - Microsoft Word</td>
<td>314</td>
<td>00:06:49</td>
<td>00:03:21</td>
<td>00:03:28</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>374</td>
<td>30-Sep</td>
<td>10:32:08 AM</td>
<td>10:33:08 AM</td>
<td>WINWORD.EXE</td>
<td>Quantitative POSitive Problems NOTE Satall notes9.30 Quantitative</td>
<td>Document1 - Microsoft Word</td>
<td>92</td>
<td>00:01:00</td>
<td>00:00:47</td>
<td>00:00:13</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CNXX01</td>
<td>375</td>
<td>30-Sep</td>
<td>10:33:08 AM</td>
<td>10:38:20 AM</td>
<td>WINWORD.EXE</td>
<td>Ex: Car A is going 60 MPH, Car B is traveling 60 mph. - They are 6 mi apart. A fly is traveling at 12.0 MPH. When are they going to collide? 3 minutes, which was the fly will be 6 miles away from the point it left, hilltops original point</td>
<td>notes9.30 Quantitative - Microsoft Word</td>
<td>326</td>
<td>00:05:12</td>
<td>00:01:40</td>
<td>00:03:32</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

| Table 1: Mostly on-task use |

Proceedings ECIS 2009
<table>
<thead>
<tr>
<th>Name</th>
<th>Match</th>
<th>Date</th>
<th>App/TaskGroup</th>
<th>ApplicationName</th>
<th>Keystrokes</th>
<th>WinTitle</th>
<th>AppKeyCounter</th>
<th>ElapsedTime</th>
<th>ActiveTime</th>
<th>InactiveTime</th>
<th>On-task</th>
<th>Off-task</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:00 AM</td>
<td>aim.exe</td>
<td></td>
<td>XXX01 - Instant Message</td>
<td>0</td>
<td>0:00:03</td>
<td>0:00:03</td>
<td>0:00:00</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:03 AM</td>
<td>aim.exe</td>
<td></td>
<td>XXX02's Buddy List Window</td>
<td>0</td>
<td>0:00:02</td>
<td>0:00:02</td>
<td>0:00:00</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:05 AM</td>
<td>aim.exe</td>
<td>hey</td>
<td>XXX01 - Instant Message</td>
<td>4</td>
<td>0:00:03</td>
<td>0:00:03</td>
<td>0:00:00</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:08 AM</td>
<td>EXPLORER.EXE</td>
<td>Business College - Information Technology - Microsoft Internet Explorer</td>
<td>0</td>
<td>0:00:01</td>
<td>0:00:01</td>
<td>0:00:00</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:10 AM</td>
<td>EXPLORER.EXE</td>
<td>Program Manager</td>
<td>0</td>
<td>0:00:01</td>
<td>0:00:01</td>
<td>0:00:00</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:12 AM</td>
<td>aim.exe</td>
<td>MSN Messenger</td>
<td>0</td>
<td>0:00:02</td>
<td>0:00:02</td>
<td>0:00:00</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:14 AM</td>
<td>aim.exe</td>
<td>XXX03 - Conversation</td>
<td>0</td>
<td>0:00:02</td>
<td>0:00:02</td>
<td>0:00:00</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:14 AM</td>
<td>aim.exe</td>
<td>MSN Messenger</td>
<td>0</td>
<td>0:00:00</td>
<td>0:00:00</td>
<td>0:00:00</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:19 AM</td>
<td>aim.exe</td>
<td>hey! what time is ur flight?</td>
<td>31</td>
<td>0:00:05</td>
<td>0:00:05</td>
<td>0:00:00</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:20 AM</td>
<td>EXPLORER.EXE</td>
<td>Business College - Information Technology - Microsoft Internet Explorer</td>
<td>0</td>
<td>0:00:01</td>
<td>0:00:01</td>
<td>0:00:00</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:27 AM</td>
<td>aim.exe</td>
<td>XXX01 - Instant Message</td>
<td>7</td>
<td>0:00:07</td>
<td>0:00:07</td>
<td>0:00:00</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:29 AM</td>
<td>EXPLORER.EXE</td>
<td>Business College - Information Technology - Microsoft Internet Explorer</td>
<td>0</td>
<td>0:00:02</td>
<td>0:00:02</td>
<td>0:00:00</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:48 AM</td>
<td>EXPLORER.EXE</td>
<td>XXX04 - Instant Message</td>
<td>16</td>
<td>0:00:19</td>
<td>0:00:19</td>
<td>0:00:06</td>
<td>0:00:13</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:47:52 AM</td>
<td>EXPLORER.EXE</td>
<td>XXX04 - Instant Message</td>
<td>6</td>
<td>0:00:12</td>
<td>0:00:12</td>
<td>0:00:05</td>
<td>0:00:07</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:48:04 AM</td>
<td>EXPLORER.EXE</td>
<td>XXX04 - Instant Message</td>
<td>0</td>
<td>0:00:05</td>
<td>0:00:05</td>
<td>0:00:00</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:48:09 AM</td>
<td>EXPLORER.EXE</td>
<td>XXX04 - Instant Message</td>
<td>0</td>
<td>0:00:05</td>
<td>0:00:05</td>
<td>0:00:00</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:48:09 AM</td>
<td>EXPLORER.EXE</td>
<td>XXX04 - Instant Message</td>
<td>17</td>
<td>0:00:17</td>
<td>0:00:17</td>
<td>0:00:03</td>
<td>0:00:14</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:48:26 AM</td>
<td>EXPLORER.EXE</td>
<td>Business College - Information Technology - Microsoft Internet Explorer</td>
<td>0</td>
<td>0:00:03</td>
<td>0:00:03</td>
<td>0:00:00</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:48:29 AM</td>
<td>EXPLORER.EXE</td>
<td>XXX04 - Instant Message</td>
<td>35</td>
<td>0:00:12</td>
<td>0:00:12</td>
<td>0:00:07</td>
<td>0:00:05</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:48:41 AM</td>
<td>EXPLORER.EXE</td>
<td>XXX01 - Instant Message</td>
<td>15</td>
<td>0:00:18</td>
<td>0:00:18</td>
<td>0:00:04</td>
<td>0:00:14</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:48:59 AM</td>
<td>EXPLORER.EXE</td>
<td>XXX04 - Instant Message</td>
<td>27</td>
<td>0:00:10</td>
<td>0:00:10</td>
<td>0:00:10</td>
<td>0:00:00</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:49:09 AM</td>
<td>EXPLORER.EXE</td>
<td>Business College - Information Technology - Microsoft Internet Explorer</td>
<td>0</td>
<td>0:00:09</td>
<td>0:00:09</td>
<td>0:00:09</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CNXX0</td>
<td>1</td>
<td>27-Sep</td>
<td>9:49:18 AM</td>
<td>EXPLORER.EXE</td>
<td>XXX05 - Conversation</td>
<td>48</td>
<td>0:00:29</td>
<td>0:00:29</td>
<td>0:00:15</td>
<td>0:00:14</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2: Mostly off-task use
More specifically, the independent variable has two levels—restricted and unrestricted. For each hypothesis, the dependent variable has three measures—focus change, keystrokes and active time. When one must test for significant differences between two groups on two or more dependent variables, one may use Hotelling's T-square statistic.

<table>
<thead>
<tr>
<th>Focus Change</th>
<th>Nominal Figures</th>
<th>Active Time</th>
<th>Proportional Figures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off-task</td>
<td>69,027</td>
<td>1,265,658</td>
<td>680,022</td>
</tr>
<tr>
<td>On-task</td>
<td>15,362</td>
<td>652,371</td>
<td>315,403</td>
</tr>
<tr>
<td>Neither</td>
<td>2,706</td>
<td>6,238</td>
<td>11,352</td>
</tr>
<tr>
<td>Total</td>
<td>87,095</td>
<td>1,924,267</td>
<td>1,006,777</td>
</tr>
</tbody>
</table>

Table 3: Use Measures

5 RESULTS

The results in Table 3 show that about two-thirds of all use is off-task and about one-third is on-task. Thus, a majority of use behaviour is off-task although a sizeable amount is on-task. When given tools to be more effective in a co-located learning setting, subjects appear to gravitate toward off-task use behaviours. This is unfortunate from a facilitator’s perspective.

<table>
<thead>
<tr>
<th>Use Behaviour</th>
<th>Access Setting</th>
<th>Hotelling's $T^2$</th>
<th>Focus Change</th>
<th>Keystrokes</th>
<th>Active Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off-task</td>
<td>Restricted</td>
<td>p&lt;.05</td>
<td>681</td>
<td>12,859</td>
<td>7,588</td>
</tr>
<tr>
<td></td>
<td>Unrestricted</td>
<td>p&lt;.01</td>
<td>1,279</td>
<td>23,155</td>
<td>11,905</td>
</tr>
<tr>
<td>On-task</td>
<td>Restricted</td>
<td>n.s.</td>
<td>246</td>
<td>13,277</td>
<td>5,560</td>
</tr>
<tr>
<td></td>
<td>Unrestricted</td>
<td>p&lt;.05</td>
<td>222</td>
<td>7,057</td>
<td>4,128</td>
</tr>
</tbody>
</table>

Table 4: Use Measures by Use Behaviour, Access Setting

The Hotelling's T square test results are shown in Table 4. The results indicate that off-task use was significantly different between restricted and unrestricted access settings on the three use measures at p<.05. Off-task levels were higher in the unrestricted access setting, which is consistent with expectations. Levels of focus change (p<.01) and keystrokes (p<.05) in the unrestricted group were almost twice that of the restricted group. The unrestricted group had a greater level of off-task active time (p<.05) compared to the restricted group as well. Hypothesis 1 is supported. On-task use was not significantly different between unrestricted and restricted groups on the three use measures. This is also consistent with expectations, assuming that the facilitator dynamically provisions services as needed. Hypothesis 2 is also supported.

6 DISCUSSION, CONCLUSION AND LIMITATIONS

The collective findings are largely consistent with expectations, but it is important to remain mindful that overall use is characterized by a preponderance of off-task use. Viewing the data differently (see Table 5), we see that in restricted access off-task use is significantly higher than on-task use at p<.01 across the three measures. This highly significant finding stems from the large difference on focus change (p<.01), where the off-task level is almost three times the on-task level. Thus, despite the PS's infrastructure and application restrictions, participants continue to find ways to engage in off-task uses. Additionally, for participants who were provisioned all infrastructure and application services
(unrestricted), their use was overwhelmingly off-task as indicated by $p<.01$ significance levels for the overall Hotelling’s T square test and for each use measure.

<table>
<thead>
<tr>
<th>Access Setting</th>
<th>Use Behaviour</th>
<th>Hotelling’s $T^2$</th>
<th>Focus Change</th>
<th>Keystrokes</th>
<th>Active Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restricted</td>
<td>Off-task</td>
<td>681</td>
<td>12,859</td>
<td>7,588</td>
<td></td>
</tr>
<tr>
<td></td>
<td>On-task</td>
<td>246</td>
<td>13,277</td>
<td>5,560</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Significance</td>
<td>$p&lt;.01$</td>
<td>$p&lt;.01$</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Unrestricted</td>
<td>Off-task</td>
<td>1,279</td>
<td>23,155</td>
<td>11,905</td>
<td></td>
</tr>
<tr>
<td></td>
<td>On-task</td>
<td>222</td>
<td>7,057</td>
<td>4,128</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Significance</td>
<td>$p&lt;.01$</td>
<td>$p&lt;.01$</td>
<td>$p&lt;.01$</td>
<td>$p&lt;.01$</td>
</tr>
</tbody>
</table>

Table 5: Use Measures by Access Setting, Use Behaviour

Our findings support the view that PS are effective at reducing off-task use while leaving on-task use unimpeded. Stated differently, PS may offer one means to balance the advantages and disadvantages of computer-based multitasking. Nevertheless, a preponderance of computer-based multitasking was off-task despite restriction of certain devices and applications to reduce off-task use. A main explanation for this breach lies in participants’ use of local applications, which lay outside the PS’s “reach.” Another explanation lies with some users devising clever ways to circumvent the PS control, although this was relatively uncommon. In either case, the findings suggest that PS are only partially effective as an organizational control mechanism for bringing about goal alignment between facilitator as principal and participants as agents in a co-located learning setting. Consequently, additional ways to control computer-based multitasking are needed.

Considering established theories, we contend that socialization structures as described in Ouchi’s (1979) theoretical work may be necessary to restrain off-task use. Extending from Ouchi’s (1979) description of them as "...a highly formalized and lengthy period of socialization during which [participants] are subjected not only to skill training but also to value training or indoctrination. (p.837),” we imagine value training and indoctrination could manifest in formal procedures to bring about behavioural modification and reform with the objective of restraining off-task uses. We envision formal procedures that would inform about appropriate computer-based multitasking uses, where notions of appropriateness encourage participants to stay on-task and resist off-task uses. These formal procedures would be executed at the onset of the learning process in order to establish behavioural norms at the beginning, and be delivered on a repeated and frequent basis for continual reminder until conclusion of the learning sessions. The formal procedures would include punitive measures, which would be clearly communicated and reliably enforced in case of violation.

An additional option may be the use of a “public” display that is large and viewable by all participants, where the “public” display could show any participant’s personal screen. While simultaneously showing all participants’ personal screens on the public display is not practical, skilful design and implementation of contemporary infrastructure and applications can allow the facilitator to selectively put one participant’s personal screen on “public” display. A viewing of a personal screen by all participants would likely be effective at discouraging off-task use, assuming that procedures were fully disclosed to participants beforehand. On a theoretical level, this is akin to using an information system such that the principal (facilitator) is informed about the agent’s (participant’s) actions (computer-based multitasking behaviour) in an effort to bring about goal alignment (reduce off-task use).

We conclude that participants’ computer-based multitasking is characterized by greater levels of off-task use when compared to on-task use by a significant margin. We also conclude that provisioning systems alone are not sufficient to effectively control computer-based multitasking—i.e., to restrain off-task use, so it may be advisable to augment provisioning systems with behavioural modification and reform efforts.

One obvious limitation relates to generalizability of these results. The data were collected from co-located settings of a college-level classroom setting. The roles of instructors as principals and students...
as agents, and the attendant role dynamics, may not adequately overlap or correspond to those of co-located facilitators and participants in ordinary business co-located settings. Thus, the findings may not generalize to this situation.

Another limitation relates to an implicit assumption in our hypotheses development that off-task and on-task use behaviours are independent of one another. It may be that some dependency exists in that as off-task use increases then on-task uses may decrease or vice versa. This may occur under circumstances where a user has no slack cognitive resource, so that any increase in off-task use will result in a decrease in on-task use. While this may be true in some instances, we argue that it is more likely that slack cognitive resources are available in co-located contexts, thereby allowing off-task and on-task uses to increase or decrease independent of the other.

7 REFERENCES


MANAGEMENT COLLUSION: Keeping the lid on the 'can of worms'

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0484.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Management control, Management practices, Deception tactics, Human computer interaction (HCI)</td>
</tr>
</tbody>
</table>
MANAGEMENT COLLUSION: KEEPING THE LID ON THE CAN OF WORMS

This paper suggests that issues surrounding non-compliance should not be dismissed as resistance but instead should be further studied by managers and developers, leading to accommodation of differing views. The technological frames of reference strand of social shaping of technology theory is used to overlay the issues arising from a case study looking at non-compliance with information systems. This procedure highlights underlying antecedent organizational conditions which are likely to underpin non-compliant behaviour. These antecedent conditions include acceptance, control and proceduralisation. Examination shows that change in these areas is within the remit of managers and yet noncompliance continues and is accepted by managers. This raises the question not of why or how these behaviours exist, but why they continue to exist in the light of their unmasking and exposure. This paper suggests that non-compliance should not be dismissed as resistance but should addressed by managers and developers leading to a skeleton or framework for understanding problems and developing organisationally aligned solutions.

Keywords: Social Shaping, Technological Frames, Resistance, Workaround.

1 INTRODUCTION

The purpose of this paper is to examine complicity and collusion by managers who fail to address issues underlying non-compliance to information systems. Several terms in this paper are deliberately selected to differ from the usual terms used and the term non-compliance is used instead of the more usual term resistance, with all associated negative connotations. In addition terms such as complicity and collusion are used to describe managerial knowledge of non-compliance along with their inaction in addressing such behaviours, where these are identified as managerially controllable. Collusion implies secret understanding with negative intent and complicity implies tacit conspiracy in underhand matters. This paper builds on work identifying categories of non-compliance with information systems (Ferneley and Sobreperez, 2006; Ferneley, Sobreperez and Wilson, 2005). The author now ties in the technological frames of reference aspect of social shaping of technology (Bijker, 1995; Orlikowski & Gash 1994) and applies these to the case study data to identify relevant social groups. This paper applies, across case study sites, the notion of technological frame incongruence across identified relevant social groups, and uses indicators for non-compliance to construct contrasting technological frames. In this way, this paper applies the lens of technological frames theory to the findings of the original case study (Ferneley & Sobreperez 2006) and uses the identified groups of conditions underpinning non-compliant behaviour to create the structure of the technological frames. This action outlines clearly that relevant groups of employees who interact with information have widely disparate views of the nature and purpose of the information they use, and that this can stifle congruence or commensuration of attitudes to information systems.

In addition, the application of technological frames helps to identify areas where management, supervisors and developers take different standpoints and dismiss or minimise non-compliant behaviours and their subsequent negative impact on data integrity, accuracy of reporting, and organisational culture. An organisational ideal is the seeking of technological frame congruence (Orlikowski and Gash, 1994) and the comparison of stances across different dimensions helps to identify the differing attitudes, beliefs and viewpoints of the relevant social groups. In this way areas in which these differing standpoints have contributed to conditions which are precursors for non-compliant behaviour are highlighted. This action may assist in the categorising, across other contexts, those behaviours which are harmless, hindering or, vitally, essential to the completion of tasks, processes or roles within an organisation (Davis and Hufnagel, 2007; Ferneley and Sobreperez, 2006). Further, this action of categorising could become important in the identification of behaviours to
ignore and behaviours requiring action in individual organisations. Non-compliant behaviours capable of mitigation by management action are continuous in their manifestation and the action of classification could assist managers in the seeking of curative or alleviating systems

2 THEORETICAL AND RESEARCH BACKGROUND

Where users show opposition to information systems, this has traditionally been seen as recalcitrant, negative behaviour, labelled as resistance, to be minimised and overcome as much as possible throughout an organisation (Franz and Robey, 1984; Lyytinen and Hirschheim, 1987a). More recently, other researchers have pointed out that resistance can be seen as a positive force and may highlight the shortcomings of systems or of organisational arrangements (Hirschheim and Newman, 1988; Levine, 1997; Markus, 1983; Mumford, Land and Hawgood, 1978). This seems to be especially true where an organisation has distinct hierarchy and operators are distanced from those who specify, develop and implement systems (Lapointe and Rivard, 2005; Marakas and Hornik, 1996; Martinko, Henry and Zmud, 1996; Prasad and Prasad, 2000). Resistance studies are generally seen as control versus resistance, management versus workers, repressive power and subsequent response, overlooking the ambiguity and complexity often found (Thomas and Davies, 2005) and assuming polarity of management and workforce.

Some attempts have been made to create environments in which resistance is minimised and this has resulted in the development of Technology Acceptance Models (TAM) (Davis, 1993; Ventakesh and Davis, 2000) which predict the impact of external variables on the internal attitudes and behaviour of people in new technology situations. However, widespread use of these has failed to significantly affect this area, or account for social influence in the adoption and utilization of new information systems (Malhotra and Galletta, 1999); developers and managers continue to view any non-compliance as resistance, with associated negative connotations, to be overcome for business reasons.

The literature has largely been managerialist in outlook focussing on the behaviour of workforce and theories of resistance from the exclusive point of view of managers (Jermier, Knights and Nord, 1994; Lyytinen, 1987b; Markus, 1983). This has involved taking a view from the outside and above, looking at the workforce as different and separate from those who manage them, who make decisions governing their use of information systems, and who determine implementation and control of new information systems (Boje, 1995).

Managerialist ideology became popular with the rise of corporate capitalism over family-owned businesses and assumptions are that managers are different kinds of people from ordinary working people, that they have skills, knowledge and attitudes that differentiate them in a superior way (Chandler, 1977). It is argued that managers are infused with a paternalistic and ‘know-best’ attitude through virtue of their position, they have a ‘responsibility’ to the organisation greater than that of lower workforce members and are agents and functionalists of the survival and growth in profitability of the organisation (Alvesson, Mats and Willmott, 1996; Boje, 1995; Hoskin and Maeve, 1993a; Hoskin and Maeve, 1993b).

The flattening of hierarchies occurring with harmonisation and removal of layers of middle managers purports to drive down decision-making and to ‘empower’ lower levels but rarely are these other than operational decisions (Appelbaum, Hebert and Leroux, 1999; Harley, 1999). Thus the unitarist approach favoured by the rise of the Human Relations school of management thought is contrasted with the pluralist ‘us and them’ idea of managers versus workforce implying tensions, conflict, and resistance to systematic and continuous workforce subjugation. Shoshana Zuboff (1988) considered the use of computer technology as a radical discontinuity in industrial history. She observed that more data is now available on the minutiae of work based tasks than ever before, and one of her findings was that power lies with access to that data. She construed that this could constitute a fundamental organisational change and would have far reaching effects for employment, social relations and power relationships providing managers did not block the development of intellectual skill of workers to
protect their own authority. She concluded that technology could be either restricting or unleashing depending on the choices made within the organisation and that these choices might threaten the fundamental distinction between managers and the managed. Thus true empowerment of the workforce at operational level becomes a threat to the power and authority of managers.

The issue of differing groups of people holding differing views of technology is introduced by the Technological Frames of Reference (Orlikowski and Gash, 1994) strand of Social Shaping of Technology Theory. The essence of this theory is that groups of people interact with technology in different ways and are identified by shared tacit knowledge, assumptions and understanding. Groups are identified as having similar purposes and objectives of technology and similar understanding of data use, data quality and ultimate information utility.

Orlikowski and Gash (1994) suggest an ideal of frame congruence, where differing groups of people are aligned on the structure and content of their frames and thus share expectations and assumptions around key technological aspects. The existence of incongruent technological frames implies differences in the view of division of labour, autonomy of employees, and status and position of individuals are consciously or implicitly built into information systems by systems planners and designers (Hirschheim and Klein, 1989).

Should incongruence be identified, the implication is that in order to move towards congruence, one or several relevant social groups will need to amend or change their frames of reference in order to resolve incongruence. Clearly these amendments and attitude changes will be expected of the user or operator and not by the designer, consultant, manager, executive champion or financer of the project Davidson (2006). The very idea of communicating and cooperating with users and seriously considering their views and outlooks undermines the status, power and privilege of managers and developers. The groups involved in IT often have different priorities and goals (Dunlop and Kling, 1991), and added to this are social and hierarchical differentials over issues such as salaries, employment benefits and social status, and this creates the potential for conflict and controversy.

In the light of the case study, there are clear issues of resistance to or non-compliance with Information Systems and underlying antecedent conditions for these acts of non-compliance are identified as issues surrounding proceduralisation, acceptance and control (Sobreperez, 2008). This paper highlights the fact that the adjustment of these determinants to minimise non compliance is within the remit of managers and explores reasons for continuation of situations where these are overlooked or ignored.

There is much literature on the tendency of managers to support the status quo and supporting analysis of the reasons why they do this (Buchholz, 1978; Hambrick, Geletkanycz and Fredrickson, 1993; Hambrick and Mason, 1984). Top executives are not necessarily open minded about change and it is often in their interests to maintain an existing state of affairs where their position, reputation and importance are understood, accepted and believed (Jost, Banaji and Nosek, 2004). Propositions expounded in Hambrick, Geletkanycz and Fredrickson’s (1993) paper are that: the longer an executive has worked in the organisation or in the industry; the greater the organisations current performance; and the more they are allowed discretion in their decision making, the greater will be their commitment to the status quo.

The upper echelons share psychological factors such as beliefs, knowledge assumptions and values (Hambrick and Mason, 1984) and these can be seen to coincide with factors that underpin relevant social groups (Orlikowski and Gash, 1984). The upper echelons theory concludes that decisions made by managers are the outcomes of these behavioural factors rather than any attempt to optimise organisational efficiency or effectiveness.
3 RESEARCH APPROACH

The GarmentCo Case study was selected as an example of a production-based conveyor system where operators follow what appear to be rigid step-by-step instructions to complete orders within a pre-arranged workflow construct. This provided the opportunity for detailed study of actual practices taking place, variations from procedures and the situations in which these occurred. In addition there was the opportunity to study the call-centre approach of the office where orders were taken.

In contrast, the fire service case study site was selected due to the unique nature of the organisation and the extraordinary context of incidents. In this case study, each incident was different in a way that cannot be true for office or production based systems. The fundamental stage of the firefighter’s work is dynamic risk assessment due to the diversity in site, progress, and danger of each incident.

Although there are two case studies, it may be argued that there are multiple incidents of issues at two domains. The GarmentCo study includes several stages in the composition of outfits and also covers the ordering and despatch sections. The fire service includes reference to many divergent incidents where few factors are repeated.

GarmentCo hires out men’s formal clothes for weddings and other formal functions through high street retailers. Outfits are distributed by company vehicles or other couriers to the retail branch through which they were ordered and after the event, they are returned through the retail branch and company vehicle to one of the main processing warehouses. There they are checked, brushed, dry cleaned or laundered as appropriate, and returned to stock to be available for the next order. The workspace shop floor at the Lancashire depot is arranged into six zones. Each zone is used to process and store a particular type of clothing and each garment is bar-coded either with an iron-on or card label attached to the garment. An automatic conveyancing system known as Gartner is used to transport an order around the shopfloor where the appropriate garments and other items are added. Each zone has at least one operative selecting (picking) garments, scanning the barcode and attaching them to the appropriate order. On collecting a ticket the user types in their user number to one of the ‘Picking’ terminals and scans the barcode on the ticket. This assigns the order to the operator and the ticket contains the garments that need to be ‘picked’ from that area. The garment is attached to the hanger and the order is placed onto the conveyor system. This ticket follows the conveyor system around the factory, in order, through all of the areas from ‘Picking’ to ‘Dispatch’. The Gartner Transfer File communicates with the system tells the order where to drop off, and comes into action at every scan point on the conveyor system. An employee will assign themselves to the order by keying in their user number at the station and scanning the order ticket. This will update the system to show that the user has ‘picked’ the required garment and when.

The regional Fire Service case study is concerned with the mobilisation of fire engines to incidents and the reporting of said incidents. As incidents are reported to the Fire Service a centralised control office records initial incident details including incident location, who has reported the incident, Fire Service personnel and fire engines that are dispatched immediately and subsequently, the route or routes taken by fire engines, dispatch and arrival times and a log of all communications with the deployed Fire Service teams. After the incident a detailed electronic report is completed categorizing and reporting on the incident, the report is semi-structured and any level of Officer can be assigned the responsibility of completing the report. Structured attributes of the information system include cause of fire, location within the address, degree and speed of fire spread, number of casualties, other emergency services involved, specific equipment used and arrival and departure times. Free format responses include incident handling strategies and lessons learnt. These reports are collated and summarized by a centralised office that then presents the abstracted results to management who allocate human and physical resources from this data. In addition the summarized data is reported to central government who allocate funding and make policy decisions based on the data.

Data collection methods at GarmentCo included non-participant observation of work practices in the warehouse and telephone office, supplemented by semi structured interviews with warehouse staff and...
managers of finance, stock and personnel functions. At the Fire Service, data was collected through focus group interviews with groups or watches of firemen supplemented by interviews with senior fire officers and administrative staff at local and national statistics offices.

4 DATA ANALYSIS

Data was analysed using inductive coding and grounded theory principles of categorising and clustering data (Strauss and Corbin, 1994; Wolcott, 1982). The aim was to allow a conceptual framework to emerge during the course of study as the data was gathered, facilitating an open minded approach to analysis, although it is ‘impossible to embark upon research without some idea of what one is looking for and foolish not to make that quest explicit’ (Lincoln and Guba, 1985). Interview and focus group transcripts were read several times by two researchers and a coding framework was developed through discussion and verification of terms.

Following the development of the initial framework the analysis was conducted by firstly identifying key terms and concepts within the case studies, these terms were then restated to ensure that they remained as descriptive and literal as possible. Terms were cross validated to ensure the meaning was maintained and appropriate para-phrasing was used. An iterative clustering process was then performed with different clustering permutations being trialled. Clusters were then named (or coded) and combined to derived ‘meta-clusters’. Comparisons were made at cluster boundaries to test the cluster coherence and mini-theories (or ‘memos’) were generated for each cluster. At the highest, most abstract, level the core category is a summary of the concept of workaround rationale. The central theme of the conceptual framework is therefore ‘workarounds are manifestations of employee deviation from the intended use of technology’. This fits with the definition ‘intentionally using computing in ways for which it was not designed, or avoiding its use and relying on alternative means’ (Gasser, 1986, P216) from the earliest paper that specifically mentions workarounds.

The initial first level categorisation scheme was applied both at the time of data collection and by retrospective analysis on the completed transcripts and supporting field notes. Because qualitative data analysis is an open and iterative process applying the initial coding scheme to the case studies resulted in the emergence or induction of a richer coding set as the initial codes were ‘extended’, ‘filled in’, ‘bridged’ and ‘surfaced’ (Eisenhardt, 1989).

The case study data was analysed in two stages, within-case analysis and cross-case analysis (Miles and Huberman, 1994). Within-case analysis was performed to allow unique patterns to emerge and to gain a richer understanding of each case study. Individual acts of resistance or workaround were analysed identifying the conditions stimulating deviation from the prescribed or intended process, the objects affected by the behaviour and the associated threats. Cross-case analysis was undertaken using analytical induction to identify both common themes and unique patterns or behaviours. Abstracting from the individual cases focussed on the causes, pre-existing conditions or antecedents that generated the resultant behaviours. Segments of interview text were coded allowing an analysis of interview segments on a particular theme. The specific interest was in finding evidence of rationale for the various manifestations of workaround activity. Applying clustering techniques across the case studies identified sets of common sub-clusters.

The following set of tables display quotes from members of relevant social groups and are set against the three identified antecedents.

<table>
<thead>
<tr>
<th>Antecedent Condition</th>
<th>Supervisor</th>
<th>Shop floor Operators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enforced proceduralisation</td>
<td>The system is highly routine and cannot be overridden</td>
<td>We break down orders to find matches we can pick together</td>
</tr>
<tr>
<td>Unwillingness to Control</td>
<td>Targets are set too low and workers slack off when they reach them</td>
<td>Some targets are way too high, impossible, we don’t even try</td>
</tr>
<tr>
<td></td>
<td>We don’t care as long as the job gets done</td>
<td>Don’t know what they use it for, don’t care</td>
</tr>
<tr>
<td>Lack of Acceptance</td>
<td>Performance monitoring is not reflected in</td>
<td>It’s not worth the time to log out and</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
pay back in again it takes ages

Table 1 Supervisors and shop floor workers (GarmentCo)

<table>
<thead>
<tr>
<th>Antecedent Condition</th>
<th>Supervisors</th>
<th>Telephone Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enforced proceduralisation</td>
<td>The system ensures staff are hitting targets and working efficiently</td>
<td>The system does not allow for missing garments, damaged garments, mislabelling or amendments</td>
</tr>
<tr>
<td>Unwillingness to Control</td>
<td>All telesales operators have to fill in a sheet monitoring time spent away from the telephone. Telesales Staff are monitored in everything they do</td>
<td>There’s loads of things we can do to get a skive, she can’t know everything, sorting post, chasing parcels, dealing with input errors are all unmonitored</td>
</tr>
<tr>
<td>Lack of Acceptance</td>
<td>My job is to supervise the performance, productivity and conduct of staff and to report this management</td>
<td>We are timed with a stopwatch and there has to be a balance between accuracy and speed.</td>
</tr>
</tbody>
</table>

Table 2 Supervisors and Telephone Operators (GarmentCo)

<table>
<thead>
<tr>
<th>Antecedent Condition</th>
<th>Senior Fire Officer</th>
<th>Firefighters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enforced proceduralisation</td>
<td>Information is used in court for insurance, to refuse housing and occasionally to remove children from families. Fire Officers are very careful how they enter data</td>
<td>The report I saw was not the fire I went to</td>
</tr>
<tr>
<td>Unwillingness to Control</td>
<td>They ask for equipment they do not need, we need to rationalise resources and there is a macho culture associated with how much equipment a station has, therefore they need to justify its use.</td>
<td>We sometimes request equipment we ‘may’ need and then justify it later by describing the incident as though we needed the kit</td>
</tr>
<tr>
<td>Lack of Acceptance</td>
<td>We are encouraging greater use of information systems to support record keeping and to instigate knowledge sharing and knowledge management across the service. Firefighters are trained on the use of the FDR1 system for 2 day each at least</td>
<td>I have never had any training on the use of the information system. There is active discouragement of IT use among the lads</td>
</tr>
</tbody>
</table>

Table 3 Senior Fire Officers and Firefighters (Fire and Rescue Service)

<table>
<thead>
<tr>
<th>Antecedent Condition</th>
<th>National Statisticians</th>
<th>Local Statisticians</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enforced proceduralisation</td>
<td>The system has been in use since 1994 and works well to capture the data required by government</td>
<td>Standardisation removes ambiguities and forces firefighters to use ‘unknown’ and ‘other’ options.</td>
</tr>
<tr>
<td>Unwillingness to Control</td>
<td>County Fire Brigades are accredited after 18 months of external measurements and are then allowed to perform their own internal quality checks. Not everything collected is used, political considerations have changed at every level since FDR1 was designed</td>
<td>At the end of the training we gave up and went back to manual systems. Sending information to the ODPM is the most important work we do</td>
</tr>
<tr>
<td>Lack of Acceptance</td>
<td>The intention is that Brigades provide routine FDR1 information as Electronic</td>
<td>Accredited County Brigades send their information directly.</td>
</tr>
</tbody>
</table>
DISCUSSION

Resistant Behaviours The empirical case studies conclude that at the top level there are three pre-
conditions or antecedents for non-compliant behaviour with regard to information systems (Ferneley
and Sobreperez, 2006; Ferneley et al., 2005). These are identified for the purposes of this study as
enforced proceduralisation, lack of acceptance and unwillingness to control. Under enforced
proceduralisation, systems are overly procedural and employees must perform jobs in a certain
sequence so that the system can record them. This sequence is numerical and therefore chronological
despite situations such as at GarmentCo when greater efficiencies can be gained by batch processing
(selecting several of the same garment at once) or by pre-emptive processing (taking jobs out of
sequence) because it is known that similar non-standard items are required. At the Fire and Rescue
Service, the enforced proceduralisation was in the sequence of the FDR1 form and the lists of possible
responses available for selection in answer to questions. Under lack of acceptance, user failure to
identify to the system was widespread across both case studies. There seemed to be two reasons for
this: firstly the time taken to log in and out again using low specification desk top machinery and a
less than contemporary network technology was seen to be wasted and secondly any attribution was
viewed with suspicion and seen as surveillance. Employees were happy not to be personally
accountable for the recording of events as this was a way to sidestep any possible retribution for
errors, omissions or incorrect usage. The third antecedent is termed unwillingness to control, this
nomenclature is intended to convey that managers do not attempt to eliminate or minimise the non-
compliant behaviours outlined under proceduralisation and acceptance. The issue of collusion is
highlighted by the contradictory and opposing views of information systems usage and utility.

Management Collusion. Across both case studies, some interviews with managers revealed
awareness of resistant behaviours but no indication of actions intended to address them. Other
interviews did not even acknowledge that such behaviours occurred and managers were much more
likely to state the official organisational view despite evidence to the contrary. It is clearly within the
control of managers to write user guides and policies, to train users to use technology in a particular
way, to ensure access to support, to establish particular controls and to ensure that those who enter
data have understanding of where their input is used, by whom and for what purpose. It was clear that
users of information were happy to continue to use information they suspected or knew to be incorrect
which leads us to the question why this might be the case. Burawoy (1979) uses the metaphor of the
fig leaf to explain how the exploitative nature of the capitalist labour process uses consent to mask the
true relationship between those who sell their labour and those who profit from it. He describes a
mask that sometimes slips but generally does a good job of obscuring exploitation. In this way
collusion becomes the tacit agreement of both (or all) parties that they must work together and reach
agreement on aspects of their working life.

Managers appear to collude in several ways. Firstly, Managers deny or refute any occurrences of non-
compliant behaviours, this is illustrated by the manager’s comment

_The system is highly routine and cannot be overridden_

And the opposing view of operators

_We sometimes break down orders to find matches we can pick together_

Secondly Managers contradict or oppose evidence of non-compliant behaviours, note the comment
below from telesales office supervisor

_All telesales operators have to fill in a sheet monitoring time spent away from the telephone.
Telesales staff are monitored in everything they do_
And the opposing view from a telesales operator

*There’s loads of things we can do to get a skive ....... Sorting post, chasing parcels, dealing with input errors are all unmonitored.*

There are many examples of direct contradictions in the data collected, operator level staff giving one account of what happens and supervisory or management staff giving other, opposing views. The reasons for this may be complex but several seem to be apparent. Firstly, commitment to the status quo as mentioned earlier may well be a strong influence; secondly many senior managers throughout organisations are of an age where information systems were not taught in schools. They will tend to have learned anything they know about information systems on the job and in an applied way and may be unaware of steps that need to be taken to re-configure systems. Similarly in many organisations there are no direct lines of communication between departments and no systems for requesting modifications. Managers are likely to be wary of displaying their lack of technological knowhow by pointing out modifications that are very simple or too complex to implement.

We do not need to believe that managers consciously deceive workers or that all workers are exploited to see that some collusion is necessary simply to perform daily activities. Surveillance and monitoring are coercive and division of labour between managers who conceive work tasks and workers who execute them form constituents of the class division. However this view does not take into account the liberal view that if we believe that managers are custodians of the interests of all then we must see control by the few in the name of the many (Sewell, 2008). In the functionalist view (Durkheim, 1911) any kind of workplace ‘deviancy’ is a dysfunctional response that deviates from the normal and rational and consideration of this issue and the roles of managers directs the questioner to ask not how it arises but why it remains (Hacking, 1983). Information is only a small part of organisational decision processes, data are not merely an intellectual commodity but a political resource, whose redistribution through new information systems affects the interests of particular groups (Keen, 1981). There are few studies of resistance by managerial and professional employees, and polarity is often assumed in that the ‘workforce’ includes those ‘other’ than the management, and that managers work for and workforce against, the organisation. Thomas and Davies (2005) however, argue that different focuses of power require different focuses of resistance and that resistance by managers must take a different form and emphasis over that of the workforce. It may also be pointed out that management takes many forms from supervisor to managing director and that rather than viewing a polarity between managers and workforces, there is a continuum which includes many different levels of management, with all but the top level subordinate and answerable to the level above.

**Culture and Control.** The trompe l’oeil of information presented to managers and senior staff is underpinned by collusion in deviations and workarounds by managers and workforce (Ferneley et al., 2005). In the Fire Service in particular, a working class, masculine culture prevailed. Firefighters saw themselves as brave men and heroes, they were unconcerned with the ‘paperwork’ of recording systems and emphasized their ‘proper’ work of facing danger to save lives and property. They were strongly unionised, protective of their masculine role as ‘breadwinner’ and resisted the introduction of female firefighters. A culture of suspicion and criticism of management motives and practices from people distanced from the coalface of firefighting arose. This stark separation between workforce and management paradoxically created and reinforced a culture of distance which increased their exposure to management prerogative and their vulnerability to managerial discretion (Collinson, 2003).

Studies in conformity show that people are preoccupied with themselves as valued objects in the eyes of those in authority, and subordinate their own subjectivity in the process (Collinson, 2003). The conforming aspect of following a personal career path in an organisation has been identified (Grey, 1994) and aspirant individuals may perceive organisational and social relations as instrumental to career progress. In this way, the workplace becomes competitive as hierarchical positions are chased by many and allotted to few. An aggressive and competitive culture emerges and individuals may not wish to ‘rock the boat’ by highlighting deviation, alienating subordinate colleagues and tackling a difficult and possibly insurmountable issue.
Another interesting point is that the relevant social groups in this study are identified by their job title and status in the hierarchy. Clearly these do not necessarily map across to the different elements of the same technological frame; there may be many senior managers in any organisation who feel that systems are imposed from above, that they have no ownership of the data, and that there are significant differences between what actually happened and the recorded version. In this way then, members of a particular relevant social group, who share attitudes, opinions and points of view over their interaction with technologies, may cut across management hierarchies, job titles and salary scales. Resistance can then be seen as multi-directional and in this way, can be seen as gradually erosion of the micro-politics of power, by the use of low levels but continuous and changing methods and manifestations which weaken, destabilize and undermine established power structures enabling the construction of subtly different power relations across an organisation (Thomas & Davies 2005).

6 CONCLUSION

Information from different groups across two single case study has been examined and reveals differences in attitudes perceptions and stances around information systems. The technological frames construct is used to contrast thoughts, opinions and mindsets of different groups towards non-compliance with information systems. Three indicators are employed to provide structure for the frames of reference and these were identified within the case study as precursors for non-compliant behaviour. The overlaying of technological frames theory and application of case study findings points out very clearly the differing attitudes, perceptions and expectations of the diverse groups, both to systems usage and to any non-compliant behaviour. The differing frames views of relevant groups indicate issues that are within the control of managers and developers and that could be addressed in future planning and implementation. In this way non-operational groups take responsibility for cultural attitudes, control of data entry, proceduralisation and systems acceptance.

The author suggests that issues surrounding non-compliance should not be dismissed as resistance but should addressed by managers and developers and become a skeleton or framework for understanding problems and developing organisationally aligned solutions. The contribution of this study then is to point out that case study findings can interlock and intersect across theoretical areas and the use of the technological frames of reference provides a useful way to clearly demonstrate the existence of relevant social groups and the congruence or otherwise of their technological frames, therefore clearly highlighting areas of concern.

The framework used by Orlikowski and Gash (1994) is augmented by three further subdivisions of proceduralisation, acceptance, and culture and control, which serve to focus more clearly on issues of non-compliance and seek understanding and accommodation of viewpoints rather than dismissal of issues important to users and to successful implementation of information systems across and organisation. Analysis of the data raised the issue of management collusion where managers are endorsing behaviours they know to be non-compliant and possible reasons for this are examined. The view that managers are somehow different from the rest of the workforce and are immune from the forces that underpin resistant behaviour is questioned. The conclusion is that most managers are are subject to the same antecedents as more subordinate workers in addition to those attributable only to their status and position. They are therefore under pressure to meet deadlines, match performance benchmarks or reach targets and are answerable to upper echelons where these are not met. They are at a disadvantage if they lay blame on the workforce they are managing, as this clearly reflects on themselves as managers and their effectiveness, achievements and success. As managers often need to describe their career path in terms of success in order to reach the next level of management in the same or a different organisation, it seems clear that individual managers would not wish to reveal difficult, and possibly insurmountable problems with resistant workers, it is in their own interests to keep the lid on the ‘can of worms’.
7 REFERENCES


Durkheim, E. (1911) *De la Division du travail social*, Alcan, Paris.


Government and self-government in the information society

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0651.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Information society, Critical theory, Discourse analysis, IT policy</td>
</tr>
</tbody>
</table>
GOVERNMENT AND SELF-GOVERNMENT IN THE INFORMATION SOCIETY

Abstract

Research on the information society and the policies and strategies for its creation has tended to discuss them rationally as the national, and occasionally international or regional, responses to changes in the competitive environment. The predominant notion of the information society in various levels of governance has only rarely been critically examined. The paper provides a Foucauldian analysis of the constitution of the information society as a political and policy imperative at the level of the European Union and the multiple effects it had for its member states. Drawing on ideas on governmentality and regimes of truth, I argue that the European Commission continually shaped the rationality and identity of the information society it heralded, by managing to set itself as the legitimate locus of policy for the information society. In revealing the dominant discursive truths about the European information society, the research discusses how the truth claims about the construction of a particular version of the information society and the legitimate loci of its government shaped the degrees of freedom of the Greek policy makers through a range of disciplining and self-disciplining practices.

Keywords: information society, Foucault, governmentality, policy, European Union, Greece
1 INTRODUCTION

The pervasiveness of information and communication technology (ICT) in all fabrics of the society and economy has dramatically expanded the domains in which it is being discussed. Notions of the knowledge economy and the information society have made their way into political discourse and government action. A wide range of policies concerning ICT have made their way into political discourse and government action. A wide range of policies concerning ICT have been the subject of academic attention, but the implicit assumption has often been that such institutional interventions arise rationally through the careful consideration of national needs and aspirations. Although this may be true for the technological innovators, there are reasons to question it for the countries which are technological laggards.

The paper discusses the normative and mimetic pressures that result in national initiatives for ICT adoption and innovation. I am arguing that information society policies originating at a regional or supra-national level play a significant role in national decisions to pursue particular types of actions with regards to ICT through processes of discipline and self-discipline. The paper discusses the way in which one such supra-national organisation, the European Union, through its European Commission, has shaped the form of legitimate national government action of one of its members, namely Greece.

The work is conceptually based on the theoretical problematizations of Foucault’s later studies on governmentality as the government and self-government of conduct. In particular, the paper will appropriate the concepts of regime of truth and practices and techniques of government to question the construction of the information society at the EU level and the implications for Greek policy makers.

In investigating the thought that is embedded in the discourse and the practices of government I attempt to reveal how Greek policy makers willingly assume the interpretation of the information society construed by their European counterparts and constitute themselves as subjects through discipline and self-discipline.

The purpose and contribution of the paper is to present a theoretically informed discussion of the construction of the information society through European policies, and its interplay with national problematizations on how to appropriate ICT. The paper is adding to the discussion of the information society and ICT policy, extending it through its parallel consideration of two levels of government simultaneously.

The paper is structured as follows. In the next section, I will be reviewing the literature on the information society and ICT policy, to highlight the lack of critical accounts as well as the limited efforts to provide a theoretical conceptualisation of the phenomenon. I will then introduce the concept of governmentality, which forms the overarching conceptual bridge of the paper, and explain the dimensions of truth, rationality, techniques and identity formation which will directly inform my analysis. In the following section a narrative of the case as it evolved over the past twenty years will be presented. In the Analysis, I will critically examine the information society policies as they evolved in the period 1994-2006, investigate the role of the European Union in making information society policies, and discuss the types of governable practices that rendered the information society an object to be governed and managed in a way that transgresses national borders.

2 LITERATURE REVIEW

The discussion about the role of IT in increased interconnectedness of states, polities and economies through the processes of globalization has primarily focused on the ability of multinationals to transcend national borders to create their own conditions of operation (Dunning 2002). However, in a seminal paper, King, Gurbaxani et al. (1994) point to the importance of government intervention to facilitate innovation in the production and use of IT.
The way the states have been implicated in intensifying connections through, and in relation to, ICT have received much less attention (Walsham 2000). Early research on the information strategies of advanced nations include Kahin (1997) on the US and Moore (1998) on the UK. The turn of the century has seen the emergence of academic literature on various national strategies for the information society (Cronberg 1997; Silva and Figueroa 2002; Iosifidis and Leandros 2003; Hall and Loefgren 2004; Chen, Gao et al. 2005; Miller 2005; Sadagopan and Weckert 2005). These studies have produced a wealth of information about the different approaches of information society policies followed by states worldwide, but have, been mostly descriptive, and have made little progress to theorise on the ways in which such policies came to be considered as political priorities at specific points in time.

Kubicek and Dutton (1997) do attempt to turn our attention to the political nature of the creation of such information society strategies and policies. They analyse the social construction of the national information infrastructure rhetoric, demonstrating the conscious and fortuitous choices made in its creation, and the impact it had on policies thereafter. Moreover, the role of donor agencies and multinational consultancies in determining national ICT strategies has been highlighted (Madon 2000; Soeftestad and Sein 2002; Ciborra and Navarra 2005). In a more critical vein, Godin (2005) critiqued the role of OECD for turning the knowledge society into an easily digested buzzword taken up and consumed by its member-states.

Despite these efforts to explain how information society policies have been shaped, there is a need to challenge the means-ends rationalistic assumption of the creation of information society policies, and to further explore the role of institutional actors in shaping them. I suggest that a closer inspection can reveal a layered and multifaceted phenomenon which challenges the reasons why national strategies and the resulting institutional interventions have risen in importance globally and more specifically in the European area.

3 THEORETICAL FRAMEWORK

The research is grounded on Foucault’s concept of governmentality, which he developed in his later lectures (2007), when he attempted to veer away from power and discourse as autonomous forces and bring agency back into the discussion. Governmentality reflects the concern with how government takes place, but addresses government in a broader way than the study of the institutions of government or even governance. It instead concentrates on the ways in which individual and collective behaviour becomes the object of government. The “conduct of conduct” (Foucault 1982), i.e. the way agents are constituted as subjects of a government rationale becomes the overarching theme.

Government is closely intertwined with thought and the production of truth. To govern means to govern based on a rationale, or a process of thinking, which in certain points in time assumes the status of truth. Thought congealed in practices of government becomes taken for granted and is rarely challenged (Rose 1999). The production and reproduction of truth determines the options that appear as feasible for individual action, and renders other options unthinkable. What becomes important, then, is to understand the thought, or rationality, of government in order to reveal the way agency is constructed.

Understanding the rationality of government requires the investigation of the material arrangements on which government is embedded. Rather than assuming the existence of invisible power relations enabling the government of agents, a Foucauldian analysis of government focuses on the practices and techniques which provide a material assemblage of heterogeneous elements, which allow the government of behaviour over a distance (Dean 1996; Dean 1999). Techniques of government, such as tools, procedures, vocabularies and technologies, make the government material, and help produce an analysis not merely of ideology, but also of the structures in which it is embedded (Kumar 2005)

Drawing from the above discussion, I will be appropriating the concepts of regimes of practice and regimes of government, techniques of governments and the idea of a field of visibility, i.e. a domain
which is rendered visible through particular practices and techniques of government, which might otherwise remain in obscurity.

The conceptual lens of governamentality allows the researcher to examine with an inherently critical eye the construction of the information society as a political and policy object in need of government attention in the European space, as well as the impacts this has had for the development of national information society policies that shape in various degrees the appropriation of ICT through the society and economy.

4 METHODOLOGY

The research is a historical case study (Mason, McKenney et al. 1997). This method of research allows the phenomenon to be examined in context over a period of time, so that the historical interconnections and patterns can be investigated. Studying the historical evolution of a phenomenon is in line with a Foucauldian tradition of research, as it allows the researcher to examine how social life is construed and thus challenge its current form.

The research draws on two sources with regards to empirical data. Firstly, a wide range of documents of policy was collected and analysed discursively. An analytics of government was pursued, in line with Foucault’s later works, according to which questions are asked of who governs, and under which truth claims and techniques (Dean 1999; Rose 1999). The aim was to understand the ways in which the information society was construed as a political and governable object, rhetoric, as well as an imperative for national action.

Secondly, semi structured in-depth interviews were conducted with a wide range of informants from the ICT policy domain in European Commission, as well as Greece. The informants’ accounts were critically examined to understand the value assigned to the information society policies and initiatives, as well as to challenge the rationality and techniques through which the information society was constituted as an object to be governed.

Analysing the data consisted of determining patterns, trying out causal chain scenarios, as well as establishing empathy with the protagonists of the story (Mason, McKenney et al. 1997). The researcher’s impressions were compiled in an analytical narrative, which highlighted important and recurrent themes, as well as temporal linkages.

The analytical narrative recounted the policies designed and implemented on ICT for almost two decades, from 1985 to 2006, tracking in parallel the history of Greek and European policies in ICT, focusing on the efforts to promote the wider use of ICT, rather than on the research and regulation aspects. For the purposes of the paper, the period from 1994 to 2006 is being considered. A narrative for this period is provided below, to form the background of the analysis.

In the Analysis section below, I address two questions. To begin with, I attempt to understand under which rationality the information society became an object to be governed, through the successive attempts to make it the object of policy at the European and national levels. In doing that, I explore how its identity was constructed and shifted through time. Secondly, I seek to explore how the Greek responses to the information society were shaped under specific regimes of truth and practices.

5 GOVERNING THE INFORMATION SOCIETY – AN EVOLVING REGIME OF TRUTH

Governing the technological artefact has always been within the aspirations of the EU. Even if not always in a visible political position, governing the deployment ICT, usually through research programmes, has been within the scope of the EU. In an insightful analysis of the role of technology in furthering the purposes of European integration, Barry (2001) has argued that the EU has attempted to
enforce homogeneity and inter-nation coordination through the creation of technical standards, while the construction of digital networks has advanced closer integration among member-states and the EU itself. Around the middle of the 1990s, the governing of the technological artifact emerged as a new object, the information society, which was to become the centre of public attention and policy. The information society was to be governed, measured and managed. Rendering the information society an object of government begets further questioning. By which rationality would this new object be governed? Who was to govern it and how?

In 1993, the information society as a rhetorical object was brought into the highlight in the Bangermann report, one of the most influential European policy documents in the domain of ICT, released under the title *Europe and the global information society*. It highlighted the role of ICTs in improving the competitiveness of the European economy, which faced the increasingly technology-based competition from the United States and Japan. The information society was constructed as primarily an economic object, and as such it needed to be managed according to the dominant neoliberal paradigm. As an economic object, the private sector would be “entrusted” with its creation (European Commission 1994, p.10). Its motor would be private innovation, produced to be marketed, and safeguarded by well-established patent systems. As an extension of the market, it needed to abide to the rules of competition; the national and regional governments needed to make sure of it. The advent of the information society was also understood to increase the urgency for further liberalisation and privatisation; market barriers needed to be lifted to allow an unmanaged European information society to face up to the challenge of global competition. Moreover, individuals were called to take ownership of the changes and take action and risks to respond to them. A distinct regime of truth emerged around the ‘European’ information society: the business in the information society was to be an innovative one and the individual an entrepreneurial one.

The European Commission’s action plan, in 1994 (European Commission 1994), detailed how the information society was to be governed. The private sector was to have “prime responsibility” for the financing and deployment of the information infrastructures, whilst member-states were urged to “promote the information society” to their SMEs. Moreover, timelines were set for the creation of European guidelines and regulatory frameworks on tariffs, intellectual property rights, competition, electronic protection, privacy and standardisation. This and only this was to be the legitimate area of action for the state in the information society.

The continuing policy-making activity, however, points to a desire, or perceived need, to make the information society the object of public policy. The action plan at a first level set a number of instructions for achieving a goal. More than that, however, it presupposed a level of knowledge about the information society and in this way attempted to render it programmable. The European Union’s action plan minimally attempted two things: it made the information society a distinct social arrangement whose nature and aspects are familiar and known, and which can thus be legitimately governed through specific instructions. Furthermore, by making the information society into a global social arrangement surpassing the capabilities of distinct states, it set the corner stone for rendering the European Commission itself as the appropriate locus of policy decisions.

In Greece, the *Greek strategy for the information society: Tool for employment, growth and quality of life* made its appearance in 1995. The product of advisors in the Ministry of Industry, Energy and Technology, it saw the information society as comprising an innovative private sector fuelling the economy with new digital products and services delivered over infrastructures created by the enterprises themselves. The SMEs would be managerial and innovative in order to survive in the particular vision of the economic and market conditions in the information society. The state was to take a light-touch approach, responsible for educating the population and ensuring fair distribution of the benefits.

Despite the heavy state-centric character of the country, and its limited technological base (Thomadakis 1995), the Greek strategy of 1995, which inaugurated the ‘information society’ as a rhetorical object for the first time in Greek policy-making, subscribed to a particular version of truth
about the information society, originating in the EU. It reproduced the discourse on the innovative business and the entrepreneurial individual as the pillars of the information society, even though it was succinctly at odds with the capabilities of the Greek public and private sectors.

The strategy was also drawn at a particularly high level. It set decade-long targets, but specified no actions or timelines to implement them in the short term. No programme of actions was specified, as if a level of knowledge about the discussed social phenomenon could not be achieved to lead to its being rendered programmable. Seeking an explanation for the reason of existence of this strategy, and of more to come, I return to this point when I look at the practices of self-government and discipline later on.

Back at the EU level, in 1999, the European Commission proposed the eEurope initiative as the new EU strategy for the information society. It was created by the newly founded Information Society Directorate General and was endorsed as an integral part of the Lisbon Strategy in 2000. Literature has already discussed the materially different character of this strategy from the previous one, particularly in the way it upheld the social and cultural dimensions, instead of the market and economic rationalities of the information society (Goodwin and Spittle 2002; Chadwick and May 2003; Berleur and Galand 2005). What is more pertinent to this analysis is the shift in what was held true for the information society and how it should be governed.

More specifically, the information society no longer revolved around the entrepreneurial individual and the innovative business; it was a different object to be governed. The information society was now fraught with the dangers of digital exclusion of the unconnected, of unemployment for the under-skilled, of consumer exploitation, and of obsolete government processes unresponsive to new demands. For these dangers to be addressed, private entrepreneurship and innovation, although necessary, were no longer considered enough. The information society was now in need of more governing; it was in need of overt government intervention (Berleur and Galand 2005; Liikanen 2005).

Two action plans, in 2000 and 2002, attempted to render the new vision of the information society programmable and governable. The action plans prescribed a list1 of e-priorities, with e-government, e-health and e-inclusion featuring prominently among others. A further list of twenty government processes to be modernised and offered over the Internet by all member-states was established. Interventions like these, most of them of a non-compulsory nature, needed to take place at the national level in set timelines. Data was also to be collected and collated by the European Commission documenting the progress at the national level. A new unit within the Commission was established to undertake this exercise of benchmarking across sectors and countries. Best practices were to be presented and shared in the Commission’s numerous fora attended by national policy makers.

The further attempts to govern the information society mark two important changes in the regime of truth. Not only was the information society now to be governed by the state, but the European Commission was to govern the governing of the information society. On the one hand, the state intervention for the creation of the information society became not only the legitimate, but also the recommended way of doing it. Taking a hands-off approach became a hard policy to justify nationally and internationally, as national backwardness was now visible to, and thought to affect, the whole Community.

On the other hand, the constitution of the Commission as a legitimate locus of governing the information society was further strengthened, but now something qualitatively different defined its role. Not only did the Commission legitimately speak of, and act upon, the information society, but it

---

1 The list included twenty government transactions, of which twelve targeted businesses and eight targeted citizen, and which were considered primary targets for reaping the benefits of the ‘information society’. Four stages of reform were specified: provision of information online, possibility to download forms, possibility to initiate the transaction online, but need to complete the transaction offline, and finally possibility to complete the transaction online. The stage of computerisation was understood to be an important measure of development of the information society.
could now govern the governing of the information society by national authorities too. Through discreet, yet effective, disciplining mechanisms, constituted through the action plans and strategies, the Commission attempted to assume the authority to determine the ‘correct’ national conduct for a phenomenon no longer presented as one of national sovereignty, but re-presented as one of global/European validity. How effective these disciplining mechanisms were and the extent to which they indeed managed to govern national conduct is taken up on in the following section.

In the national context, the White Bible for Greece’s Entry into the Information Society was created and publicised in 1999. The strategy, which effectively constituted nothing more than a non-binding white paper, documented its authors’ vision for the role of ICT in the progress country. The choice of name is quite emotive. It was neither a strategy, nor a policy, not an action plan. It was a bible, a document of conviction, of the dangers of not following a specific path and hopes of a better future in the opposite case. It was about rights and obligations: the rights of citizens in the information age, and the obligations of the state to ensure the future of its citizens and businesses in the new socio-economic arrangement.

As a bible, the value of the White Bible could not be discussed in terms of its merits, neither could it be challenged by counter-arguments but through the positioning of the discussant in a opposing paradigm, or, in a different regime of truth. But the White Bible reiterated the dominant at the time regime of truth about a state-led information society and the role of the state in fostering the transformation. It discussed e-services and e-rights for the new type of citizens turned consumers of public and private electronic services, and the state’s role as educator, protector and procurer. Challenging its dogma would have meant openly opposing a version of truth about how to pursue the information society that was made legitimate by important institutional players. It did not happen but much later in the day, when the Bible’s ‘doctrine’ was juxtaposed to the reality of implementation.

In 2003, the drafting of a new information society strategy began in Greece by two academics-turned-policy implementers for the information society, and keen to leave their distinct mark. The draft, which was never finalised, maintained the truth of the information society as shaped by social, instead of market values, but advocated the need for a more horizontal mobilisation of the civil society to compliment the vertical role of the state. An ‘improved’ vision of the information society is upheld, simultaneously drawing from the dominant regime of truth, and emergent regimes about the power of networks (mainly the Internet) to connect and empower discreet individuals into powerful collectives.

Back in the EU, a new information society initiative, i2010, produced in 2005, stressed the provision of favourable business environments through interventions and policy-making. It employed different discourses of media policy and digital convergence to suggest “proactive” intervention to enable market growth, innovation, and “quality of life”, understood as social and digital inclusion. Different discourses, same rationality of government. The emphasis on benchmarking as a technology of government had given its place to the need for harmonization of technical standards and regulatory frameworks to facilitate the emergence of a European information society. The Commission’s role of bringing about these changes was considered nearly commonsensical.

In Greece, a new strategy was initiated yet again in 2006, called Digital Strategy. The product of labour of very few policy advisors, it reframed the question of the information society in terms of quality of life and economic productivity. The discourses appropriated were very similar to i2010, but for the first time in the Greek policies on the information society, the Digital Strategy was more than a vague picture of a vision. It broke down the information society as an object into manageable pieces and time-boxed them. The document itself, which was printed in an impressive illustrated publication and circulated widely in the political, policy and business cycles, through its timetabled actions, rendered the information society into a object that was new, yet known and well understood. In a governable fashion, the state knew what the correct response of its functions, businesses and citizens was. The information society was being rendered governable at the national level.
6 THE INFORMATION SOCIETY WITHIN A REGIME OF DISCIPLINING AND SELF-DISCIPLINING PRACTICES

The previous section discussed the constitution of an evolving regime of truth about the information society through the production of a series of European policies and the way it was reproduced in Greek policy-making. I also argued that the Commission attempted to govern not only the information society through its constitution of an object of policy, but also the governing of the information society by the member states. In the following section, I go on to discuss distinct ways in which the Commission was partly successful in its attempts to govern the behaviour of at least one member-country, Greece, through mechanisms which had as a direct impact to govern the behaviour of the national policy-makers of the information society.

6.1 Shaping the degrees of freedom

The role of EU-produced documents as objects that rendered the information society a known, familiar and thus governable social arrangement, through the production of concrete action plans and timelines, was discussed above. Unlike these, however, the Greek information society documents, at least the ones up to 2006, made little effort to re-present the information as a well-known object to be governed, and thus could not effectively function as means of governing other actors’ conduct with regards to the information society and its national implementation. Their purpose then needs to be sought elsewhere.

Interviews with Greek policy makers revealed deep contradictions as to the reason of creating these documents. An example of this can be investigated in the case of the creation of the Greek White Bible produced in 1999. A key policy-maker behind the drafting of the White Bible found it “disgraceful” that, before the creation of the White Bible Greece was the only European country not to have an information society strategy, as if it highlighted a shortcoming of policy attention or shrewdness on the part of Greek policy makers in appreciating the magnitude of the impeding change. Pointing out that the White Bible was not the first Greek information society strategy, as the 1995 one preceded it by a few years, he dismissed it as inappropriate.

This can be interpreted as a question of legitimacy with regards to dominant truth on what the information society and the role of the state in fostering it should be at this specific point in time. In this respect, the 1995 strategy was not legitimate in 1999 on two counts. Firstly, because it was incompatible with the shifted dominant regime of truth about the particular version of the information society pursued by the Commission at the time. It was out of tune with the prevalent congealed thought about what kind of object the information society was and how it was meant to be governed. Secondly, because not creating a new strategy was incongruent with the expected national policy response to the object of the information society. For the Greek policy makers creating the White Bible in 1999 (and there is evidence that the same hold true for the two subsequent strategies) was thought to demonstrate a rational and purposeful approach towards planning and investing in ICT, which they believed to be congruent with what was expected of them by the Commission. The content of the strategy was less significant than its perceived symbolic value as the outcome and instrument of rational deliberation on managing and constructing an information society.

So, the whereas the purpose of the European documents was to render the information society an object that was known well enough to be governed through action lines and deadlines, the purpose of the Greek documents was less to act as a set of instructions, but rather to symbolically demonstrate to be compatible with the expected pattern of behaviour, and the accepted truths about the information society at specific, important in terms of timing, points in time.

The Commission’s attempts to govern the information society established a regime of practice where the creation of an information society strategy was the appropriate thing to do. The regime of practice
effectively defined the degrees of freedom of Greek policy-makers, who, working within the context of a country with limited capacity in technology policy, saw their options to be already defined.

What this further means is that the assumed position of the Commission as the legitimate governor of the governing of the information society was not challenged by the Greek policy-makers. They responded to the EU-originating regimes of truth about the information society by upholding the authoritative position of the Commission in speaking of, and initiating action on, the information society.

6.2 Mechanisms of discipline

The Commission employed benchmarking as one of the core methods of achieving compliance with the targets of its version of the information society, by “track[ing] progress towards the agreed targets” on a web space (European Commission 2000) and ensuring that well-performers were congratulated, and laggards made visible to all. Interviews with Greek policy-makers revealed the role that cross-comparison and benchmarking tables played in their decisions to act. As a mechanism of government, the benchmarking exercise has the capacity to re-define an array of disparate countries as a homogenous area of comparison. Re-presenting all European countries in the same space has the potential to create a European technological zone to be governed (Barry 2001).

The effects of the benchmarking exercise in governing the Greek policy makers’ actions have been however mixed. All of the higher-level policy makers interviewed in Greece were acutely aware of the “dire image of the country abroad” owning to its consistent positioning at the bottom of the benchmarking tables. The comparison with other Southern European countries, as well as the new accession countries was thought to be particularly damaging. Their efforts were “geared towards closing the gap”, which would then be rendered visible in the benchmarks. Interviewees in Brussels also extensively utilised the benchmarking tables in our meetings to point out that Greece consistently held the last position and that even their persistent efforts had not managed to reverse the situation.

Despite the prominent position of the benchmarking exercise in the rhetoric of Greek and European high-level officials, and the related discourses on catching up, the visible effects of its governing of practice on all but the higher-level policy-makers were more ambiguous. The practice of benchmarking was often deconstructed and challenged both discursively and in practice. For example, numerous interviewees in Greece brought up the list of twenty proposed government services to be offered over the Internet across Europe and utilised it as a yardstick for measuring the country’s progress towards the “information society”. The same informants were however often critical of the local relevance or success of these projects as well as of the way they were “imposed” on them and measured. The constitution of a unifying and comparable European technological space was in practice defied in favour of the uniqueness of national socio-economic conditions. The practise of benchmarking was relegated to a secondary role, which was more symbolic than governing.

7 DISCUSSION

The analysis above provided evidence in support of the argument that the information society was constituted as a political object to be governed by the Commission, which continuously shaped its identity and the claims of truth by which it would be governed. Through practices of discipline and self-discipline the degrees of freedom of Greek policy-makers in ICT were shaped through their interaction with the Commission’s information society policies.

The creation of information society strategies has been usually discussed as the purposeful response to the intensification of information production, processing and dissemination in the whole fabric of the society and economy. Their sudden proliferation at around the turn of the century has rarely been questioned. In an insightful critique, Mosco (2004) addresses this question by analysing the mythical aspects of the discourse on ICT, in order to explain how, repackaged into the information society or
the digital era, the ICTs captured the imagination of bricoleurs of governments around the world. His study reveals the consumption side of the discourses that imaginative bricoleurs of governments create, furthermore providing evidence for the disciplining and self-disciplining effects of these myths, a concept similar to this discussion of regimes of truth.

The research falls within the critical stream of thinking, in attempting to challenge the dominant perceptions of social reality through “providing alternative readings (Alverson and Deetz 2000, p.17). Providing alternative reading a Foucauldian lens requires more than analysing language to reveal the hidden notions of government. Language is constitutive of government, as government can only take place under a certain description. A regime of intelligibility is what allows government to govern specific parts of social life. As such, language does not only describe acts of government; it also makes them possible. Discursive practices are both part of, and help constitute a regime of practice, i.e. a specific way of making sense of the world. Regimes of practices define what holds true in different points in time and determine what legitimate social action is. The analysis has attempted to show the distinct ways in which language actively determined the government and self-government of people and artefacts. In doing that, it extends existing studies of the European information society (De Miranda and Kristiansen 2000; Goodwin and Spittle 2002; Chadwick and May 2003) which critique the ideology behind the European information society discourse, but make no reference to the impact this discourse for social action.

What the analysis has further revealed is the important and often neglected role of regional, supranational organizations in setting the agenda by filtering global trends and technological concerns for their members. Working alongside trends of globalization, such organizations effectively scan the technological landscape for solutions that match their aims and promote them, thus influencing a zone around them where ambiguity and available options are reduced. The EU helps isolate, in certain ways, member-states from the direct impact of globalization.

Finally, it is important to note that the arguments of this study are derived from an in-depth study of one European country. As such, they are not necessarily generalizable to the rest of the European countries which face a different mix of socio-economic conditions. For example it would be expected that not all countries subscribed to the dominant regime of truth claims and practices produced by the EU. Countries that find it hard to articulate a technological vision of their own may more susceptible to such interventions. It is expected that the influence of regional organizations will be reduced in countries with stronger technological traditions. However, further research would need to examine whether similar findings arising in other cases of financial, technological or policy interplay at different levels of governance, such as international donors and developing countries.

8 CONCLUSION

This paper sought to critically examine the constitution of the information society as an object of political and policy discourse, and challenge the rationalistic assumption that information society strategies respond to clearly defined needs, by revealing the dominant role of supra-national institutions in determining the available options for national governments. The paper critically discussed the constitution of the information society by the European Commission as a discursive political and policy object and examined the role it played in the emergence of national information society strategies.

The discourse analysis of the information society strategies has revealed how the construed identity of the information society and the rationality by which it would be managed shifted through time from more liberal to more state-centric. The regime of truth dominated the conceptions and interpretations of the information society in Greece, where strikingly similar versions of the truth were expressed. The analysis has further showed how seemingly apolitical practices and devices, such as the information society documents, have had highly disciplining effects on the Greek policy-makers’ decisions to bring about particular types of responses to the information society. Simultaneously, the
Greek policy-makers willingly subscribed to the regime of practice, reproducing similar truths about the information society and self-governing their behaviour, either because they felt that degrees of freedom were extraneously defined, or in order to pursue for instrumentally satisfactory outcomes.

The contribution of the paper lies in its theoretical conceptualisation of the way information society policies seem to have been constructed in a particular social context. The discussion of governmentality appears to have been a fruitful way to theorise the phenomenon and to provide propositions which could further lead to theoretical generalisations.

References


Soefestad, L. and M. Sein (2002). ICT and development: east is east and west is west and never the twain shall meet? IFIP TC9 WG9.4 Social implications of computers in developing countries, Bangalore.


UNDERSTANDING CITIZENS' BEHAVIOURAL INTENTION IN THE ADOPTION OF E-GOVERNMENT SERVICES IN THE STATE OF QATAR

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0420.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Adoption, Diffusion, E-government, Technology adoption</td>
</tr>
</tbody>
</table>
UNDERSTANDING CITIZENS' BEHAVIOURAL INTENTION IN THE ADOPTION OF E-GOVERNMENT SERVICES IN THE STATE OF QATAR

Shafi Al-Shafi, Brunel University, UK
School of Information Systems, Computing & Mathematics
Shafi.AlShafi@Brunel.ac.uk

Vishanth Weerakkody, Brunel University, UK
Business School
Vishanth.Weerakkody@Brunel.ac.uk

ABSTRACT:

Electronic government (e-government) has shown encouraging results in developed countries in the context of delivering electronic information and services to citizens. However, despite the many lessons that can be learned from the experiences of those e-government initiatives, developing countries are still faced with various issues pertaining to their implementation of e-government services. Like many other developing countries, the e-government initiative in the state of Qatar has faced a number of challenges since its inception in 2000. Using a survey based study this paper describes citizens’ behavioural intention in terms of applying the Unified Theory of Acceptance and Use of technology (UTAUT) model to explore the adoption and diffusion of e-government services in the state of Qatar. Analysis of the results indicates that there is significant positive relationship between performance expectancy, effort expectancy, social influence and behavioural intention to use e-government services for the citizens of Qatar. Implications for practice and research are discussed.

KEY WORDS: Adoption, Diffusion, E-government, UTAUT, Qatar
1. INTRODUCTION

The Qatari e-government initiative was launched in 2000. In global terms the UN e-government readiness report (2008) ranked Qatar’s e-government project as number 53 worldwide. As in many countries, the national e-government focus in Qatar is to achieve the highest performance in executing governmental transactions electronically, through streamlined business processes and integrated information technology solutions (IctQATAR, 2008). In 2004, ictQATAR was established to manage and develop the overall Information and Communication Technology (ICT) strategy in Qatar relating to infrastructure, service delivery and legislation of public services. A year later ictQATAR took overall control of the national electronic government initiative with an aggressive plan for e-government program parallel with e-health and e-education programs using a phased implementation plan. As part of this ongoing strategy to encourage accessibility of e-government services, free wireless internet access was introduced in 2007 in public parks. The Qatari government hopes that this concept will encourage more citizens to use e-government services and help bridge any digital divide that may exist in the state of Qatar by augmenting citizens’ participation and engagement in public services.

Given the above context, the rationale for this research is to gain a better understanding about the adoption and diffusion of ‘e-government services’ from Qatari citizen’s perspective. The relative newness of the e-government concept in Qatar and the lack of prior published research that explore the citizen’s perception of electronic government information and services offered the motivation and rationale for undertaking this research. Furthermore, initial research conducted by the authors indicates that the adoption and diffusion of e-government services has been slower than the Qatari government’s expectations (Al-Shafi and Weerakkody 2008;). Given this context, this research aims to examine the factors influencing citizens’ intention to adopt e-government services in Qatar. In order to achieve this aim a survey based quantitative research strategy is adapted. Since the primary aim of the research is to explore the intention of citizens to use e-government services in Qatar, this is achieved by examining their perceptions of ‘Ease of Use’ and ‘Usefulness’ in relation to the e-government services offered. To pursue this line of inquiry, this research uses Unified Theory of Acceptance and Use of Technology (UTAUT). UTAUT theorizes that an individual’s behavioural intention to use a technology is determined by performance expectancy, effort expectancy, social influence, and facilitating conditions (Venkatesh et al., 2003; 2008).

In an e-government context, Moon (2002) proposed that IT and web-based public services can help governments to restore public trust by coping with corruption, inefficiency, ineffectiveness and policy alienation. Conversely, lack of access to e-services (Chircu and Lee, 2005) and digital divide (Carter and Belanger, 2005; Carter and Weerakkody, 2008; Chen et al., 2006;) are challenges that can impact on trust and thereby impede the further take-up of e-government services. To bridge the digital divide, Reffat (2003) suggests that governments could help by providing computer education especially to elderly and younger people. It has also been found that younger people are accessing the internet more compared to the elderly (Kurunananda and Weerakkody, 2006). One of the main reasons is that younger people enjoy the use of the internet whereas the elderly find it difficult to learn and use. Also, the elderly do not want to be bothered with anything relating to technical matters (such as learning to use the mouse, modems, network connections and other peripheral devices) and software aspects (such as Internet Explorer).

The paper is structured as followed. In the next section a literature perspective of e-government is offered followed by an outline of the theoretical model used for the research. Then the empirical background to the research is presented. This is followed by the methodology used for the research and a presentation of the empirical results. Finally the paper concludes by analyzing the empirical results, discussing the research implications and identifying areas for future research.
2. E-GOVERNMENT IN THE STATE OF QATAR

The State of Qatar is a peninsula with a strategic position at the centre of the west coast of the Arabian Gulf. The total land area is approximately 11,437 sq km. The population estimated to be around 1,500,000 (The Peninsula, 2008; Al-Shafi and Weerakkody, 2008); however, only a minority of the population is citizens by birth, while the rest are residents who live or work in Qatar and are not Qatari’s by birth (Al-Shafi and Weerakkody, 2008).

E-government was launched in Qatar in July 2000 and the initial period of strategy formulation and implementation was laggard compared to e-government efforts during the same period in developed countries. However, with the establishment of ictQATAR in 2004 and their consequent takeover of the national e-government initiative a year later resulted in accelerated progress in the last three years. Parallel programmes were introduced in key areas such as health, interior affairs and education. The vision of ictQATAR states “Serve as independent and fair regulator, protecting consumers and businesses from unfair practices as Qatar transitions to a competitive telecoms market. And Lead the government’s ICT strategy, nurture innovative technologies to benefit those who live and work in Qatar, and help make people from all walks of life become comfortable with technology” (ictQATAR, 2008).

The Qatari e-government site offers many services, ranging from student registration and paying traffic violations to applying online for visas and permits (Al-Shafi and Weerakkody, 2008). In global terms the UN Global e-government readiness report (2008) ranked Qatar’s e-government project as number 53 worldwide, where as in 2005 it was ranked as number 62 worldwide (Al-Shafi and Weerakkody, 2008). In addition to this, the UN (2005) report considered the Qatari e-government project to be regional (West Asia) best practice. This implies that major improvements and developments have been made during recent times. As part of Qatar government’s ongoing efforts to increase accessibility to e-government services and bridge the digital divide, free wireless internet access in public parks – (iPark) initiative was launched in March 2007; this concept provides “Broadband for all” and aims to foster a knowledge based society. The primary goal of the initiative is to increase internet usage by establishing “hot spots” in public parks (IctQATAR, 2007). There are currently three designated wireless internet hotspots throughout selected public parks in the city; these parks are targeting visitors who have internet access available on their laptops, PDAs, and other internet-ready devices (The Peninsula, 2007).

3. E-GOVERNMENT ADOPTION: A LITERATURE PERSPECTIVE

With the popularity of e-government growing, various researchers have offered different definitions to explain the concept (Seifert and Petersen, 2002). However, these definitions differ according to the varying e-government focus and are usually centered on technology, business, citizen, government, process, or a functional perspective(Seifert and Petersen, 2002; and Irani et al., 2006). The definition considered to be most suitable for the purpose of this paper is one that defines e-government as making full use of the potential of technology to help put its citizens at the centre of the e-services provided and which makes its citizens its intention (Waller et al., 2001).

Like e-business, e-government promises to deliver a number of benefits to citizens, businesses and governments. The most significant benefits of e-government, according to the literature, are delivering electronic and integrated public services through a single point of access to public services 24 hours a day, seven days a week (Reffat, 2003); bridging the digital divide so that every citizen in society will be offered the same type of information and services from government (Huang, 2007); rebuilding customer relationships by providing value-added and personalized services to citizens (Weerakkody and Dhillon, 2008; Davison et al., 2005); fostering economic development and helping local businesses to expand globally; and creating a more participative form of government by encouraging online debating, voting and exchange of information (Davison, et al., 2005).

Like any other new technology or organisational concept, the introduction of e-government to a country will also result in a number of challenges for the citizens and the government alike (Seifert and Petersen, 2002; Zakareya and Irani, 2005). Lack of access to e-services (Chircu and Lee, 2005; Huang, 2007;
Carter and Weerakkody, 2008), security concerns (Harris and Schwartz, 2000), trust (Carter and Weerakkody, 2008; Welch et al., 2005; Al-Sebie and Irani, 2005), individual differences (Reffat, 2003) and digital divide (Carter and Bélanger, 2005; Chen et al., 2006) are challenges that can impact on participation and thereby obstruct the further take-up of e-government services.

From the aforementioned, this research will focus particularly on the influence that technology complexity and e-government services have on the intention to use a new technology. Similarly it will also examine the influence performance expectancy and effort expectancy has on the intention to use such services.

In terms of adoption, several studies have explored e-government acceptance in the United States (Carter and Belanger, 2005; Carter et al., 2008) and the U.K (Choudrie and Dwivedi, 2005; Carter and Weerakkody, 2008). However, no studies exist that examine factors that influence Qatari citizens’ adoption of e-government services. In this respect Lee et al., (2005) state that cross-national research on e-government is sparse in the literature and Dwivedi et al., (2006) and Carter and Weerakkody (2008) highlights the need for studies that investigate the adoption rate and behaviour of e-services. Given this context, this study attempts to address this gap in a Qatari perspective by integrating the aforementioned constructs from the Unified Theory of Acceptance and Use of Technology (UTAUT).

4. TECHNOLOGY ADOPTION: A THEORETICAL BACKGROUND

Researchers in the field of Information Systems and Technology have for long been interested in investigating the theories and models that have the power in predicting and explaining behaviour (Venkatesh et al, 2003). Various models were developed, such as the Theory Of Reasoned Action (TRA) (Fishbein and Azjen, 1975) and Technology Acceptance Model (TAM) (Davis, 1982). Each model has its own independent and dependent factors for user acceptance and there are some overlaps (Dillion and Morris, 1996).

TAM has received extensive support through validations, applications and replications for its power to predict use of Information Systems (IS) and is considered to be the most robust and influential model explaining IS adoption behaviour (Davis, 1982; Davis et al., 1989; Davis and Venkatesh, 1996; Lu et al., 2003). On the other hand, it has been found that TAM excludes some important sources of variance and does not consider challenges such as time or money constraints as factors that would prevent an individual from using an information system. Also, TAM has failed to provide meaningful information about the user acceptance of a particular technology due to its generality (Mathieson et al., 2001). Consequently, a number of modified TAM models were proposed which are applicable to contemporary technologies (Horton et al., 2001; Chau and Hu, 2001). However, researchers are confronted with a choice among a multitude of models. Hence, a new model was developed to address these limitations, which is named as the Unified Theory of Acceptance and Use of Technology (UTAUT) model and the aim of the model was to understand intention/usage as the dependent variable (Venkatesh et al., 2003).

The UTAUT model consists of eight theoretical models: the theory of reasoned action (Davis et al. 1989), the technology acceptance model (Davis, 1989), the motivational model (Davis et al., 1992), the theory of planned behaviour (Ajzen, 1991), a model combining the technology acceptance model and the theory of planned behaviour (Taylor and Todd 1995), the model of PC utilization (Thompson et al., 1991), the innovation diffusion theory (Rogers, 1995), and social cognitive theory (Compeau and Higgins, 1995). The UTAUT model combines the previous eight theoretical models and is made up of four key factors that act as determinants of behavioural intentions and usage behaviour. Also, UTAUT posits the role of four key moderator variables (Age, Gender, Experience, and Voluntariness of use). Moreover, UTAUT model has been found to be preferred to the abovementioned theoretical models as it is able to account for a high percentage of the variance (R²) in usage intention (Venkatesh et al., 2003).

Venkatesh et al., (2003) have tested the unified theoretical model in four different organizational settings for a period of six months and the study showed significant predicts intention (performance expectancy,
effort expectancy, social influence, and facilitating conditions), whereas attitude toward using technology, self-efficacy, and anxiety were theorized not to be direct determinants of intention.

5. CONCEPTUAL MODEL AND HYPOTHESIS

From the aforementioned factors that make up UTAUT, this research utilized three independent variables (performance expectancy, effort expectancy, and Social influence) and one dependant factor (intention to use) to formulate the research model used for the study.

Performance expectancy is defined as the degree to which individuals believe that using a system will help them improve their job performance and contains five variables: Performance Expectancy, extrinsic motivation, job-fit, relative advantage, and outcome expectations (Venkatesh et al., 2003). Effort expectancy is the degree of ease associated with the use of the system (Venkatesh et al., 2003). Venkatesh et al., (2003) identify three constructs from the eight models that make up the concept of effort expectancy: perceived ease of use, complexity, and ease of use. Additionally, Marchewka et. al., (2007) claimed that this construct can be significant in determining user acceptance of information technology.

Social influence is “the degree to which peers influence use of the system”, whether positive or negative, it is a very main factor in many aspects of the lives of young people and is likely to be powerful (Venkatesh et al., 2003).

Facilitating conditions are the degree to which an individual believes that an organisational and technical infrastructure exist to support the system (Venkatesh et al., 2003). Facilitating conditions are comprised of three root constructs: perceived behavioural control, facilitating conditions, and compatibility (ibid).

Due to the constraint on the length of paper, facilitating conditions and use behaviour that are originally in the UTAUT model are excluded from the proposed model in this research. Additionally, facilitating conditions is predicted to have a direct effect on actual usage instead of behavioural intention to use (Venkatesh et al., 2003).

Based on the aforementioned and the theoretical context offered, this paper will test the strength of the hypothesized relationships mentioned in the theoretical model outlined in figure-1 and the appropriateness of the model in predicting users’ behavioural intention to use e-government services in the State of Qatar.

![Figure 1. Research Model](image-url)
### Table-1: Research Hypothesis

<table>
<thead>
<tr>
<th>NO.</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1:</td>
<td>There would be a significant positive relationship between performance expectancy and behavioural intentions to use e-government services.</td>
</tr>
<tr>
<td>H1a:</td>
<td>There would be a significant positive relationship between performance expectancy and behavioural intentions to use e-government services, and this relationship would be moderated by gender.</td>
</tr>
<tr>
<td>H1b:</td>
<td>There would be a significant positive relationship between performance expectancy and behavioural intentions to use e-government services, and this relationship would be moderated by age.</td>
</tr>
<tr>
<td>H2:</td>
<td>There would be a significant positive relationship between effort expectancy and behavioural intentions to use e-government services.</td>
</tr>
<tr>
<td>H2a:</td>
<td>There would be a significant positive relationship between effort expectancy and behavioural intentions to use e-government services, and this relationship would be moderated by gender.</td>
</tr>
<tr>
<td>H2b:</td>
<td>There would be a significant positive relationship between effort expectancy and behavioural intentions to use e-government services, and this relationship would be moderated by age.</td>
</tr>
<tr>
<td>H2c:</td>
<td>There would be a significant positive relationship between effort expectancy and behavioural intentions to use e-government services, and this relationship would be moderated by internet experience.</td>
</tr>
<tr>
<td>H3:</td>
<td>There would be a significant positive relationship between Social influence and behavioural intentions to use e-government services.</td>
</tr>
<tr>
<td>H3a:</td>
<td>There would be a significant positive relationship between Social influence and behavioural intentions to use e-government services, and this relationship would be moderated by gender.</td>
</tr>
<tr>
<td>H3b:</td>
<td>There would be a significant positive relationship between Social influence and behavioural intentions to use e-government services, and this relationship would be moderated by age.</td>
</tr>
<tr>
<td>H3c:</td>
<td>There would be a significant positive relationship between Social influence and behavioural intentions to use e-government services, and this relationship would be moderated by internet experience.</td>
</tr>
</tbody>
</table>

### 6. RESEARCH METHODOLOGY

To explore the argument set out above and understand the context of the e-government services initiative in Qatar, brief informal open-ended interviews (Yin, 1994) were conducted with four citizens and three researchers who are engaged in e-government research in Qatar during August 2008. The interviews lasted around 30 minutes and provided the context to formulate a detailed survey questionnaire that was to be used to investigate the citizens’ perceptions of e-government services in Qatar. Questionnaires were selected as the primary data collection methods as it is inexpensive, less time consuming and has the ability to provide both quantitative scale and qualitative data from a large research sample (Cornford and Smithson, 1997; Miles and Huberman, 1994). Questions were compiled from IT adoption literature to represent the constructs in the proposed research model (figure 1) and wording of the questionnaire was modified to fit the research context and background information collected from the initial informal interviews mentioned above. 36 closed format questions were used limiting individual responses to multiple choice answers, for example, ranking using Likert scale (5-point scale) and ‘yes’ or ‘no’ answers (Hall and Hall, 1996; Saunders et al., 2002). This enabled the information to be grouped and analysed statistically (Leung, 2001; Hall and Hall, 1996) using SPSS V15. Since the answers can be influenced by the order the questions are presented, this was carefully planned with an introduction which explained who the researchers represent, purpose of the research and how and why the respondents were selected for the research, and the importance of their answers to the research. The main body contained topical questions ordered logically and in a manner non-threatening to respondents (Liinamaa et al., 2004). After the questionnaire was designed, a limited testing was done using six researchers and ten practitioners. This was important to improve the questions and to test respondents’ comprehension and...
clarity before the actual survey was administered (Saunders et al., 2002). The pilot testing led to the removal of two questions and modification to another one.

The survey questionnaire was distributed to a total of 250 citizens between the period of September and October 2008; 216 usable responses were obtained. The response rate was good with 86.4% responding. The protocol followed for data collection was as follows. First, one of the researchers approached students and members of academic and non-academic staff in the national university of Qatar and general citizens who use the services of two government agencies that have implemented e-government services. The researcher identified himself and provided a brief description of the research and the main purpose of the questionnaire. Then, the process of distributing and completing of the questionnaire began within the premises of the university and the two government agencies. Whilst distributing the questionnaire to users, a brief explanation of the purpose and contents of the questionnaire were offered to the respondents and participation was on a purely voluntary basis. The questionnaires were completed in an environment free from external pressures and at the respondents’ own pace. The questionnaires distributed in the two government agency premises were collected after a period of between 10 to 30 minutes, and the questionnaires distributed in the university were collected the following day.

6.1 Data analysis

The proposed research model (figure 1) consists of three independent variables, Performance Expectancy; Effort Expectancy, and social influence and one dependent variable, behavioural intention to use. To check the responses of the questions, the first stage of the data analysis consisted of checking the responses and tagging them with a unique number. The authors generated the descriptive statistics (percentage and tables) and used Linear Regression analysis by utilising SPSS (Version 15.0). Descriptive data analysis provides the reader with an appreciation of the actual numbers and values, and hence the scale that researchers are dealing with (Dwivedi and Weerakkody, 2007).

7. ANALYSIS OF THE RESEARCH FINDINGS

Of these 216 usable respondents, 41.7% were females while 58.3% were males. In terms of education, the majority of respondents (54%) hold undergraduate level qualifications degrees, 10% hold postgraduate degrees (Masters and PhD) and 36% hold equal or below secondary school certificates. As far as age, the results revealed that the majority of respondents (34%) were found in the age group of 18-24, followed by the age group of 30-44 constituting around 23.6%, then age group of 25-29 (22%), and finally age group of 45-54 composing around 13% of the total respondents. In contrast, the younger groups (less than 18) and older age groups (greater than 54) consisted together of 7.4% of the total respondents. Additionally, the study shows that the majority of respondents used e-government services for: emails (79.6%); Fun (58%); Research (62.5%); Purchasing (32%); and others (23%).

As found in the study (table 2), the average scores for respondents’ Effort Expectancy ranged from 3.70 to 4.20. Descriptive statistics show that these scores are quite high. For Performance Expectancy, the score ranged from 3.63 and 4.25, which is quite high. Concerning social influence, the score ranged from 2.98 to 3.94, indicating that the scale is average. The last score reached is 4.22 for behavioural intention to use, indicating that the scale is quite high.

<table>
<thead>
<tr>
<th>Demographic Categories</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1.42</td>
<td>.494</td>
</tr>
<tr>
<td>Age</td>
<td>3.25</td>
<td>1.258</td>
</tr>
<tr>
<td>Occupation</td>
<td>1.29</td>
<td>.456</td>
</tr>
<tr>
<td>Qualification</td>
<td>1.75</td>
<td>.632</td>
</tr>
<tr>
<td>Nationality</td>
<td>1.46</td>
<td>.500</td>
</tr>
<tr>
<td>Internet Experience</td>
<td>4.12</td>
<td>1.177</td>
</tr>
</tbody>
</table>
Performance Expectancy

PE1. Using e-government system will enable me to accomplish tasks more quickly 4.25 .773
PE2. E-government services will develop existing bureaucratic services 4.25 .721
PE3. E-government system will be useful by one click 4.11 .766
PE4. E-government system will enable me accessing it 24/7 4.25 .877
PE5. Equal chances to all citizens when using e-government 4.13 .821
PE6. Government agencies can be trusted 3.63 .995
PE7. Leadership are committed to e-government project 3.88 .863
PE8. E-government team collaborates with other government agencies 3.72 .900
PE9. E-government system integrates with other government agencies systems 3.70 .965

Effort Expectancy

EE1. Learning e-government system would be easy 4.03 .885
EE2. Interaction with the e-government system would be clear and understandable 4.00 .825
EE3. It will not be hard to be skilled to use e-government system 4.03 .806
EE4. If I got the resources, e-government system would be easy to use 4.20 .780
EE5. Overall, e-government system is easy to use 4.06 .815

Social Influence

SI1. Important people to me think I should use the online government system 3.57 .893
SI2. I would use online government services if I needed to 3.94 .807
SI3. I would use online government services if my friends and colleagues used them 2.98 1.109
SI4. People around me who use the e-government system have more prestige 3.26 1.025

Behavioral Intention

BI. I predict that I will use e-government system in the future 4.22 1.001

*Scores range from 1 to 5, where 1 = Strongly Disagree and 5 = Strongly Agree.

Table 2: Descriptive Statistics

7.1 Reliability Test

Cronbach’s coefficient alpha values were chosen to examine the internal consistency of the measure (Hinton et al., 2004) (table 3). Cronbach’s results varied between 0.56 for social influence and 0.83 for the effort expectancy. Hinton et al., (2004) have suggested four different points of reliability, excellent ranges (0.90 and above), high (0.70-0.90), high moderate (0.50-0.70) and low (0.50 and below). The previous mentioned values show that all of the constructs got high reliability except social influence which resulted to high moderate within table 3. The high Cronbach’s values of the constructs means that constructs were internally consistent and the reliability is measuring the same construct.

Table-3: Reliability of Measurements

Table-4 also shows that the correlation is significant to these key factors, Performance expectancy (0.405), Effort expectancy (0.268), and Social influence (0.247). Whereas, other moderators such as age, gender, and internet experience were found to be insignificant.

Key Factors

<table>
<thead>
<tr>
<th>Performance Expectancy</th>
<th>Pearson Correlation</th>
<th>Behavioural Intention to use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance Expectancy</td>
<td>Sig. (2-tailed)</td>
<td>.405(**)</td>
</tr>
<tr>
<td>Effort Expectancy</td>
<td>Sig. (2-tailed)</td>
<td>.268(**)</td>
</tr>
<tr>
<td>Social Influence</td>
<td>Sig. (2-tailed)</td>
<td>.247(**)</td>
</tr>
</tbody>
</table>
Table 4: Correlations

### 7.2 Regression Analysis: Factors Influencing the Intention to use e-government services.

A regression analysis was conducted with the use of e-government service channels as dependent variable and Effort Expectancy, Performance Expectancy, and social influence as predictor variables. From a total of 216 cases that were analyzed a significant model emerged \{F (3, 216) =12.325, p < .001\} (Table-5). The second significant statistic that was obtained from the analysis is the $R^2$, which ranges from 0 to 1, with 1 being a perfect fit model. It was found that $R^2= 0.185$ for this analysis. This factor explains 18.5% of the changes in the behavioural intention to use e-government services. Other unidentified factors account for the remaining 81.5%. Also, table-5 shows that of all the factors, the following have no significant impact on behavioural intention to use e-government services: Performance Expectancy ($\beta = 0.202, p= 0.029$); Effort Expectancy ($\beta = 0.173, p= 0.065$); and Social influence ($\beta = 0.166, p= 0.032$).

### ANOVA

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.084</td>
<td>3</td>
<td>6.695</td>
<td>12.325</td>
<td>.000(a)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.43(a)</td>
<td>.185</td>
<td>.170</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factors Effecting Intention To Use e-government Services</th>
<th>Beta</th>
<th>Standardised Beta</th>
<th>T Statistics</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention To Use e-government Services Factors (Constant)</td>
<td>1.071</td>
<td></td>
<td>1.996</td>
<td>.048</td>
</tr>
<tr>
<td>Performance Expectancy</td>
<td>.038</td>
<td>.202</td>
<td>2.198</td>
<td>.029</td>
</tr>
<tr>
<td>Effort Expectancy</td>
<td>.051</td>
<td>.173</td>
<td>1.857</td>
<td>.065</td>
</tr>
<tr>
<td>Social Influence</td>
<td>.056</td>
<td>.166</td>
<td>2.164</td>
<td>.032</td>
</tr>
</tbody>
</table>

Predictors: (Constant), Effort Expectancy, Performance Expectancy, and social influence Dependent Variable: behavioural Intention To Use.

Table 5: Regression analysis results

### 8. DISCUSSION AND CONCLUSION

This paper discussed the results of a survey targeted towards e-government service users in Qatar and it provides a representative account of the citizens’ perceptions of the e-government services in the state. The significant and non-significant factors found in the study and their influences on practice are outlined below.
8.1 Significant Factors
Of the adoptions factors, performance expectancy, effort expectancy, and social influence all had a significant impact on intention to use the Qatari e-government services.

8.2 Non-Significant Factors
Gender, age, and internet experience were found to be insignificant in terms of predicting the behavioural intention to use e-government services.

8.3 Implications for Practice
Citizens using e-government may benefit from the services and consequently encouraged to adopt e-government. If the government provides more benefits to its citizens in terms of convenient access and prompt services, when compared to the old and traditional means, then possibly this practice might spread the use of e-government services throughout the Qatari society.

Furthermore, the conclusions that have emerged from the analysis presented in this study are as follow:

- Although research exists that explores citizen adoption of e-government services in many countries, the authors argue that currently there is no independent studies that examine e-government adoption in the State of Qatar.
- Three constructs (namely, performance expectancy, effort expectancy, and social influence) significantly influenced the behavioural intention of citizens for adopting e-government services in Qatar.
- Empirical finding suggests that the influence of the remaining three constructs (namely, Gender, Age, and Internet experience) on citizens behavioural intention of the e-government services in Qatar were non-significant.
- Five constructs (namely, performance expectancy, effort expectancy, social influence, age, and Internet experience) were found to have a positive affect on behavioural intention. Whereas, gender was found to have a negative affect on behavioural intention.

Given these findings, it can be concluded that the e-government services initiative in Qatar has been successful initially in promoting wider access to the Internet. This is encouraging from an e-government perspective. Yet, much more can be done to raise awareness of e-government in Qatar such as advertising and promoting the national e-government website and setting up citizen service centres to assist those who are less-computer-savvy to adopt e-government services. While the research findings are encouraging from a practical perspective for the Qatari government, from a theoretical perspective these results reconfirm that technology acceptance is influenced by key constructs such as Performance Expectancy, Effort Expectancy, and Social influence aspects of the e-government services used. From a practical perspective however, one has to recognise the fact that although the survey results are encouraging, e-government has yet to mature in the state of Qatar since its inception in 2000. This is sad considering the progress other countries in the region, such as Dubai and Bahrain have made during the same period. As our survey results reflect, some of the reasons for this can be attributed to the fact that citizens still do not fully trust e-government services, are concerned about security and some are hindered by the lack of internet access. Furthermore, it can be concluded that this study extends the theoretical knowledge in the area of citizens’ adoption of technology (in this case, e-government applications and services) by testing the UTAUT model in a developing country (Qatari) context. Future research can focus on extending this study to other gulf countries and draw comparative analysis of e-government efforts in different countries in the region from a citizens’ perspective.
REFERENCES


DOES THE ANSWER LIE IN COLLABORATION? – A CASE STUDY ON E-GOVERNMENT AND SOCIETAL AGING

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0223.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-government, E-inclusion / exclusion, Exploratory framework, Empirical study</td>
</tr>
</tbody>
</table>
Abstract

E-Government can be considered the key element when it comes to modernizing public administrations as result to rising demands of citizens towards the provided services. At the same time, the demographic change posed numerous challenges to the development of new strategies. The need for inclusive E-Government, as well as human resource issues and the general reorientation of services require substantial investments. However, depopulation leads to a reduction in tax income, resulting in an imbalance of public budget. Cooperation between municipalities can be a promising approach to reduce costs and thus deal with major challenges of demographic transition in both front office and back office. Our exploratory single case study analysis introduces auspicious solutions related to these aspects and discusses in how far they might be applied to other settings.

Summary of Arguments

- Literature identifies demographic change as a key future research theme for E-Government while it is hitherto considerably under-researched.
- Aging and depopulation are two major, often inseparable elements of the demographic change.
- The demographic change creates a dilemma for public administrations as it, on the one hand, increases the need for reforms and IT investments in inclusive Age-Aware E-Government solutions, and, on the other hand, decreases the financial means available.
- Cooperation both in front as well as back office E-Government – in the case study setting in terms of a shared service center, inter-organizational collaboration, and public-private-partnerships – is a possible mean to cope with demography-induced challenges.

Keywords: E-Government, E-Inclusion, demography, societal aging, depopulation, elderly.
1 INTRODUCTION

E-Government is gaining in importance due to rising demands of citizens towards their local administrations in terms of service improvement. In addition, cost pressure forces public administrations to look for innovative ways of service provision. As a consequence of the Lisbon-Agenda, all EU (European Union) member states committed to implement an E-Government-oriented strategy as an effort of public administration modernization. On the other hand, demographic change is a topic of greatest practical and theoretical relevance. Not only academic literature (Birg, 2005; Coleman, 2006; Kaufmann, 2005), also every bigger newspaper talks about critical demographics in industrialized countries and population projections tend to receive a more dramatic paint every time they are released. Almost all of Europe is subject to a vast reduction of fertility and an increase of live expectancy, resulting in both depopulation and demographic aging (Birg, 2005). Thus, inclusive and especially age-aware E-Government is gaining in importance in all industrialized countries and is discussed as one of the most current topics in research and practice (Kaplan, 2005; Becker et al. 2008a; Niehaves et al. 2009).

Especially senior citizens are still very much excluded from participation in electronic (European Commission 2006). Against this background, the EU initiative i2010 set up a comprehensive strategy to strengthen citizen-centric inclusive E-Government services, dedicating future E-Government measures to a set of broad policy guidelines. Amongst others it is focused upon an inclusive European information society. This reflects that socio-demographic change and its consequences for E-Government are discussed as one of the key future E-Government research themes (Wimmer, Codagnone and Janssen, 2008). Against this background, we seek to address the following research questions within this paper:

What are major demographic trends in industrialized societies – taking the example of Germany –, what are possible consequences for public administrations, and what could be possible means in E-Government to react to these challenges?

In order to address the research questions, the remainder of the paper is structured as follows: Section 2 builds the theoretical foundation of the paper by analyzing related work, specifically on the demographic transition, E-Inclusion, and the effects of aging in the E-Government literature. Section 3 elaborates the research methodology applied, here: explorative qualitative single case study analysis based on nine expert interviews. Background information on the demographic situation in the case of Alphaville (pseudonym is used for reasons of anonymity), Germany, as well as the selected qualitative interview data will be presented in Section 4 and discussed and interpreted in Section 5. Our paper concludes with a discussion of limitations, a summary of results and an argument on potentially fruitful avenues for future research.

2 THEORETICAL FOUNDATION

2.1 Demographic developments

According to Hauser & Duncan (1959), demography can be defined as „the study of the size, territorial distribution, and composition of population, changes therein, and the components of such changes, which may be identified as natality, mortality, territorial movement (migration), and social mobility (change of status)” (Hauser and Duncan, 1959, p. 2). Here, literature specifically identifies three major factors underlying the development of demography: a) fertility, b) mortality, and c) migrations (Kaufmann, 2005). Accordingly, demographic transition can be understood as the progressive alteration of these determinants. Especially fertility and mortality have undergone significant changes in most industrialized countries. On the one hand, fertility has been declining due to, for instance, changed life models or family planning and the possibilities of birth control and abortion (Hill and Kopp, 2000). Morgan & Hagewen (2005) state that fertility transitions “[…] are complete in many developed countries and are in progress in much of the rest of the world. The transition model has three stages: relatively high and stable fertility, followed by a period of fertility
decline, and then followed by relatively low and stable fertility.” (Morgan and Hagewen, 2005, p. 231)

Mortality on the other hand has been declining as well, mainly due to advances in medical care (Höhn, 2000). Many industrialized countries experienced this process of demographic transition in the past and are now only able to prevent a population decline by retaining a positive net migration. Therefore, those countries with a fertility rate below the replacement level and low external migration are most likely subject to depopulation. Additionally, structurally weak regions are frequently losing population due to selective emigration (Kröhnert, Medicus and Klinkholz, 2007). Especially young highly mobile people move out of these areas in order to have better chances to get a job. Furthermore, while life expectancy is rising and populated experiences normal aging, less young people are following up which results in an overall aging of the society in these countries, and here in specific regions. The old-age dependency ratio that sets the number of people aged over 60 years in relation to the number of people aged between 20 and 60 years dramatically increased by two to three times its size in some industrialized countries (Birg, 2005). Therefore, depopulation and aging can be identified as the two major consequences of demographic transition in industrialized countries.

2.2 Effects of Aging

Even though the demographic transition has an influence on most of the described gaps, aging is certainly one of the most important aspects to consider when creating or adapting E-Government strategies. According to Davis’ Technology Acceptance Model (TAM), the two main dimensions of technology acceptance are perceived usefulness and perceived ease of use (Davis, Bagozzi and Warshaw). The perceived usefulness on the one hand has the biggest influence on the actual usage. Older people are lacking the awareness of advantages related to ICT usage because they are easily satisfied with their current possibilities and cannot imagine how ICT could improve their lives (Morris and Venkatesh, 2000). They have a critical attitude towards the usage of new technologies because they did not get used to them throughout their working lives. They appreciate the face-to-face contact and fear that new media will contribute to their isolation.

The perceived ease of use, on the other hand, is usually declining with an increasing age. An impairment of vision implies difficulties in recognizing details especially if they are presented on a computer screen (Shirley, 2004). Furthermore, it is harder for the elderly to adapt necessary skills because they have problems to understand complex new processes. Recent studies show that elderly more often experience problems with ICT than younger persons (European Commission 2004a, 2004b). They are still very much excluded from participation in electronic services (for the case of Germany see Table 1, Niehaves & Becker 2008).

<table>
<thead>
<tr>
<th></th>
<th>Total Population</th>
<th>Senior Citizens (55-74)</th>
<th>Ratio (SeniorCit/TotalPop)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet usage</td>
<td>69 %</td>
<td>37 %</td>
<td>0.53</td>
</tr>
<tr>
<td>E-Commerce usage</td>
<td>38 %</td>
<td>15 %</td>
<td>0.39</td>
</tr>
<tr>
<td>E-Government usage</td>
<td>28 %</td>
<td>12 %</td>
<td>0.43</td>
</tr>
</tbody>
</table>


Table 1. Online Service Usage (Germany)

However, the degree of technology aversion differs between individuals within the group of senior citizens. Browne states in this context that “once elders become avid users, their online skills and activities do not differ much from other age groups” (Browne 2000). Bavec’s (2008) correlation analysis between E-Government use by individuals and selected socio-economic indicators revealed "regular use of internet services", "interest in science and technology", and "trust (social capital)" as major factors for E-Government usage. All these indicators, however, have relatively low values among the elderly. While the Internet usage is comparatively low within the group of senior citizens, major reasons for not using E-Government services where missing personal contact, the complexity of E-Government services, and concerns about additional costs (for the case of Germany see Table 2). All these reasons for non-usage have been mentioned more often by senior citizens than they have been mentioned by the entire population (Niehaves et al. 2008).
<table>
<thead>
<tr>
<th>Reason</th>
<th>Total Population</th>
<th>Senior Citizens (55-74)</th>
<th>Ratio (TotalPop/SeniorCit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal contact missed</td>
<td>48 %</td>
<td>52 %</td>
<td>1.08</td>
</tr>
<tr>
<td>Too complex</td>
<td>24 %</td>
<td>30 %</td>
<td>1.24</td>
</tr>
<tr>
<td>Concerned about additional costs</td>
<td>13 %</td>
<td>17 %</td>
<td>1.27</td>
</tr>
<tr>
<td>Other reasons</td>
<td>16 %</td>
<td>14 %</td>
<td>0.89</td>
</tr>
</tbody>
</table>


Table 1: Reasons for not Using E-Government Services (Germany)

2.3 E-Inclusion

These trends pose several challenges to the development of E-Government strategies. The societal aging is one factor that leads to what is called the digital divide, in this context understood as an emerging polarization phenomenon in society, creating a gap between those who do have access to and use the potentialities of information and communication technologies, and those who do not (European Commission, 2004a). Apart from the described demographic gap referring to the fact that senior people often do not use ICT on a regular basis (E-Aging) (Becker et al. 2008b), several other gaps can be identified. Disabiliites can debar people from actively using information technology. For the usage of online services the most important disabilities to consider are visual handicaps, cognitive defects and limitations of motor skills. Geographical differences refer to gaps in ICT usage between different regions. Socio-economic gaps include differences in occupation, income and education whereas ethnical and cultural gaps identify barriers in the ICT usage of migrants and ethnical minorities.

Here, E-Inclusion focuses on the elimination of these barriers for the use of ICT. The declaration of Riga gives the following definition of E-Inclusion: “‘eInclusion’ means both inclusive ICT and the use of ICT to achieve wider inclusion objectives. It focuses on participation of all individuals and communities in all aspects of the information society. eInclusion policy, therefore, aims at reducing gaps in ICT usage and promoting the use of ICT to overcome exclusion, and improve economic performance, employment opportunities, quality of life, social participation and cohesion.” (Ministers of the EU, 2006, p. 1) The main focus of E-Inclusion is on creating accessible services over ICT. This effort can be divided into accessibility and usability aspects (Kraner, 2004). It is stated that the design and delivery of key services and public service policies shall be user-centric and inclusive, “using channels, incentives and intermediaries that maximise benefits and convenience for all so that no one is left behind.” (European Commission 2006b) Finally it also proposes to ensure “that electronic documents are available in such a way that they can be used by people with disabilities in an appropriate and, where possible, EU-wide recognised” (European Commission 2006a) format. With these statements, declared by 34 member countries, eInclusion in eGovernment or inclusive eGovernment becomes a key issue in many EU countries.

The discussion of related work shows that current research on E-Government strategies mainly focuses upon front end issues to deal with the process of demographic aging. As of today, back office improvements in the area of E-Government in order to cope with the demographic challenge are thus far under-investigated. Furthermore, challenges and possible solutions regarding depopulation are under-researched as well.

3 RESEARCH METHODOLOGY

In order to investigate into our research question, we chose an exploratory case study as method of analysis. The following classification and methodological discussion refers to Yin (2003) and Benbasat et al. (Benbasat et al., 1987). Our research goal was to collect qualitative data as a basis for building a preliminary theory of demography-aware E-Government (initiatives). A holistic single case design in a local public administration was chosen in order to allow for a rich understanding of the
potentially various (stakeholder) perspectives. Thus, 9 semi-structured expert interviews were conducted, pre-informed and structured with the help of a prior literature analysis. The resulting interview guideline included questions on demography, E-Government, the local government website, E-Inclusion and human resource management. Each interview partner was addressed with a selection of those questions corresponding to his or her specific working area. Besides such specific questions, the interviews also included an open discussion on all of the above given topics and were as well open to other related aspects that the interviewee brought up.

The experts interviewed held the organizational roles of, for instance, head of IT department, E-Government responsible, webmaster, head of human resources, coordinator of demographic program and the governing major. All interviews were performed in German and later translated into English. The case setting Alphaville was chosen mainly for the reason that our pre-analysis of potential case candidates indicated that the city was heavily affected by depopulation and aging, the identified two major demographic trends in industrialized countries, including Germany. In order to frame the qualitative case study findings and to discuss the question of generalizability, we commence the case analysis with a brief case description, including available studies and quantitative data on the specific case demographics and their embeddedness in the overall societal demographic developments. Moreover, we conclude our case discussion addressing Lee’s (1989) key questions of in how far a) the setting variables and b) the setting conclusions might potentially be i) repeatable and ii) generalizable (Lee, 1989).

4 CASE STUDY RESULTS

4.1 Demographic Trends in Germany

Regarding the demographic development in Germany and specifically in terms of the case under investigation – the city of Alphaville –, two major demographic trends can be identified: depopulation and aging.

A) Germany – as most industrialized countries – is subject to significant depopulation. In 2006, the average number of births per women reached a historical low with 1.33 children (German Statistical Office, 2008). Already in the late 1960s, this index fell below the limit of 2.1 needed to remain a constant population level. For quite a while, immigration was able to compensate the resulting population loss. Nevertheless, this trend stopped in 2003 and Germany’s population is shrinking ever since (Kröhnert et al., 2007). According to the 11th coordinated population projection, the total number of inhabitants could reach a level of 68.8 million in 2050 (Eisenmenger, Pötzsch and Sommer, 2006). This will be down from around 82 million today. The trend of depopulation worsens especially in those areas that suffer from economic and structural problems. For example, regions in Eastern Germany and the Ruhr area – the case city of Alphaville is part of the latter – are facing depopulation related to regionally declining work opportunities. While young and often more mobile inhabitants are following the job market and emigrating, immobile senior inhabitants often stay in the specific regions. Such development contributes to a process of societal aging with a specific regional component.

B) Aging can be regarded as the second major trend of demographic change in Germany. Over the last centuries, life expectancy at birth has been significantly rising in Germany. From 1900 to 2006 it increased by more than 30 years for both boys and girls (German Statistical Office, 2008). Nevertheless, life expectancy of the 60-year-olds also went up within this timeframe. Along with the described lack of fertility this lead to a process of aging within society because “[…] improvements in survival mostly benefit the older population, making the population itself older.” (Coleman, 2006, p. 63) While fewer children are born, the elderly age-groups with high birth-rates are getting older and older. This trend will continue in the future. Recent projections show that the number of people older than 65 years will increase from 15.7 million in 2005 to 22.9 Mio in 2050 (German Statistical Office, 2008). Already today, Germany is the country with the third highest percentage of this elderly age group in an international comparison. By 2050, the share of the elderly will take up more than 30
percent of the German population. Already in 2010, the number of older people will exceed the number of younger ones.

4.2 Demographic Trends in Alphaville

A) Deindustrialization led to significant depopulation in Alphaville. Currently, Alphaville has around 64,200 inhabitants (09/2008) on an area of 37 square kilometers. In contrast to this, in 1975, a total of 70,645 people were living in Alphaville. The comparison of these figures states a loss of nearly 7.6 percent within the three decades (Strohmeier, Bader, Melzer, Schulz and Wunderlich, 2006). With the closing of coal-mines, Alphaville – the city with the former biggest coal-mining industry in the Ruhr area – experienced massive deindustrialization. While the number of available jobs decreased, the overall population started to decline (Kröhnert et al., 2007). According to recent projections, cities in the district of Alphaville will encounter a population loss of 8 to 15 percent between 2004 and 2020. Alphaville will be one of the most affected cities with a forecasted depopulation of 14.5 percent.

B) Selective migration and depopulation led to an aging of the population in Alphaville. While the loss of workforce needed led to general migration, many senior citizens were often not willing to move away from the area in which they have lived the greatest share of their lives. On the other hand, they were often not looking for new jobs and thus stayed in the region. This led to a trend of aging within the society of Alphaville. With the population percentage of children and juveniles aged under 18 years hitting a high of 26 percent in 1975, the same age group today only makes around 18 percent of the inhabitants. Moreover, the population share of citizens of age 65 years and older went up from 12 percent to around 22 percent within the very same timeframe (Strohmeier et al., 2006). This demographic trend has led to the situation that in Alphaville elderly today make a greater share of the population than children and juveniles. Most recent projections show that the age structure will continue to change dramatically. By 2020, the 30 to 49-year-olds will lose their majority of number to the 50 to 64-year-olds. Moreover, the share of people older than 80 years will in 2020 have almost doubled its size compared to 2003 (Bertelsmann Stiftung, 2008). Thus, Alphaville will continuously be subject to significant depopulation and aging in the future.

As a result, the city can be regarded as a case of rather dramatic economic change and demographic transition. While it is aligned with the national trends of decrease of fertility and increase of life expectation, Alphaville has suffered, more than many other cities, from an economic downturn and the selective migration (emigration) of younger inhabitants. In that regard, Alphaville can be seen as a rather typical case for those cities in Europe that have been affected by deindustrialization and resulting depopulation and aging.

4.3 Qualitative Case Data

The qualitative interview data shows that the two major demographic trends – depopulation and aging – confronted Alphaville with numerous challenges. The city therefore sees dealing with the demographic change as a cross-divisional task that affects the overall strategic orientation and decisions. An Alphaville public administration department head stated:

‘We see coping with the demographic change as a cross-functional and strategic task that affects all actions of the public administration.’

In Alphaville, a decrease in population led to a decrease in tax revenue. The city is not able to reduce costs to the same degree. Infrastructure, including public administration, is laid out for a specific number of citizens and contains a great share of fixed costs which cannot be cut back in the short- or mid-term. An Alphaville public administration department head stated:

‘It is impossible to close 10 percent of the schools if the number of students decreases by 10 percent. Neither are we able to close down a couple of traffic lights if the population declines.’
This is similar for most of the other infrastructure related expenses. Therefore, the tense financial situation of Alphaville has intensified in the move of demographic transition, here depopulation (see Roberts 2003 for a general analysis of aging-related financial stress). Discussing possibilities to cope with this financial strain, solely reducing the public administration personnel (costs) is not a possibility: it would mean less manpower for a rising number of tasks. An Alphaville public administration department head stated:

‘We won’t be able to cut back personnel in the same extend that we are losing population. [...] Considering that today we do have to take a credit of 50000 € per day in order to cover our over-the-budget expenses – and this is about the annual salary for an employee – we would have to fire one person a day in order to be ‘efficient’.’

Furthermore, population aging in Alphaville evokes the need for rethinking the existing E-Government services offered. Against the background of a larger share of senior citizens in the population, E-Inclusion becomes an important topic. The Alphaville.de webmaster stated:

‘Focusing on the elderly is important because it is evident that more and more seniors use the internet. For them it is a good possibility to participate in daily life even if they are not mobile anymore.’

This shift in focus leads to a rising number of tasks for the city. Providing services for the elderly requires additional thinking on usability and accessibility and therefore leads to a higher work load within the public administration. An Alphaville public administration department head stated:

‘In our city, this situation will intensify. Focusing on elderly people will become a crucial task in public administration’s work. [...] There are many additional things that have to be considered: Barrier-freeness, easy access, adaptation of consultation hours, re-formatting printed materials and so on.’

However, the city of Alphaville is turning their back on a purely E-Inclusion focused strategy for elderly service provision. Instead of getting the elderly to use E-Government services directly, establishing a mobile citizen service that can provide the needed services and assist the elderly is considered a promising approach. An Alphaville public administration department head stated:

‘We don’t know yet in how far the people that are currently 50 or 60 years old will adopt new media in order to retain their mobility when they turn 70, 80 or 90. Currently, we are planning to implement a mobile citizen service that brings the services to the elderly rather than taking IT to the care centers.’

In addition, the demographic change also has an internal effect. The aging population implies an aging workforce within the public administration. In Alphaville, there has been evidence that this development leads to a reduction in performance and especially in innovation potential. An Alphaville public administration department head stated:

‘We have seen problematic examples where employees aged somewhere around the mid fifties or late fifties would slowly surrender from what I would call ‘production process’ and were just looking to ‘sneak out’.’

This leads to a decreasing degree of knowledge transfer between the elderly and younger employees. Knowledge management systems are considered a possible solution to this problem. These, as well as needed changes for E-Inclusion measures, necessitate substantial IT infrastructure investments. However, the city of Alphaville lacks the financial power to push such reforms. Therefore, the demographic change poses a dilemma to the city. An Alphaville public administration department head stated:

‘We would have to do a lot more in order to cope with the challenges of the demographic change. This, however, is not possible with intensifying financial problems. [...] Depopulation has a very high negative influence on municipalities like us that are already facing severe financial problems.’
On the one hand, the aging population calls for an adaptation of online services offered and IT investments. On the other hand, depopulation intensifies the critical financial situation making needed investments impossible. As an attempt to cope with this situation, Alphaville uses innovative forms of cooperation to share costs, as the Alphaville CIO stated:

>'These [the cooperations] start with IT services that we operate together with other municipalities. Our payroll accounting is carried out in cooperation with the neighboring city [Betaville]. We receive IT services from a regional service center that is shared with other municipalities. Those are cross-city efforts that we do focus on in order to be efficient.'

Especially concerning IT cooperation, the city still sees numerous possibilities for improvement. Since calculating capacity can be shared across long distances, Alphaville looks over local boundaries in order to find cooperating partners. An Alphaville public administration department head stated:

>'IT opens up a lot of possibilities for cooperation. Instead of cooperating with our neighboring city [Betaville] we could also share IT capacity with, for example, Munich. This is no problem today.'

In addition to such models of cooperation with other cities, Alphaville also focuses on public private partnerships (PPP) as a mean of serving both cost reduction/efficiency and effectiveness (inclusive online services). The new website that was mainly developed in order to create accessible E-Government services has been implemented in cooperation with two partly city-owned service providers. Thus, Alphaville uses cooperative elements in both front office (PPP) and back office (shared service centres). In the future, the city plans to extend these models, especially in E-Government, in order to offer more efficient services to the citizens.

5 DISCUSSION

In the case under investigation, depopulation and aging have been identified as major demographic trends. They were caused by both societal demographic transition, as to be found in most industrialized countries, and regional developments, including economic downturn. While depopulation resulted in an immense pressure to reduce costs, the provision of age-aware inclusive E-Government services necessitates additional IT investments. Cooperation both in the back office and front office in terms of a Shared IT Service Center or a Public Private Partnership in E-Government were considered as feasible means to cope with such (demography-related) investment dilemma (for a depiction of case study results see Figure 1). In order to discuss the question of generalizability of the results, we conduct a brief case study discussion following Lee’s (1989) set of analysis questions (Lee, 1989):

1.) What is the initial case setting and in how far is it bound to specific situational and historical circumstances?

A core issue in the case study was the need for reducing costs within the public administration. Deindustrialization in Alphaville led to a decrease of jobs in the area and therefore to an emigration especially of young people. This resulted in a process of aging within the society. We identified that depopulation and aging were two demographic aspects which came hand-in-hand. Depopulation reduced the direct tax income of the city and worsened its financial situation. The inability to cut back infrastructure in the necessary extent puts the public administration in front of only two possible ways of reaction. Either the oversized infrastructure is maintained resulting in higher per-capita expenses, or the relevant institutions are closed down and rebuild in an adapted size which also creates high costs (Kröhnert et al., 2007). Both ways, decreasing revenues are offset by relatively steady costs. In addition, the aging workforce within the public administration requires a reorientation of human resources as well as an improvement of knowledge transfer to younger employees and thus investments in e.g. knowledge management and human resource systems. Bill (2007) stated in that context that strategic human resource management will be one of the most important adjustment screws to turn in order to create a future-proof and efficient public administration. All this leads to an imbalance of public budget. Population based algorithms for allocating federal taxes to municipalities
intensify this trend. Aging contributes to the financial aggravation because tax revenues from the senior citizens tend to be comparatively low (Bach, Borg, Krimmer, Raffelhüschen and Schulz, 2002). Furthermore, a rising number of elderly evokes the need for an adaptation of the existing E-Government services and necessitates the availability of financial means. Cities like Alphaville, that are subject to a distinct process of demographic change, therefore face a dilemma between the need for creating a new E-Government infrastructure due to aging on the one hand and less financial power due to depopulation on the other hand.

Figure 1: Demographic trends and their consequences for E-Government

2.) Which conclusions are drawn from the case study data and in how far are these conclusions bound to specific situational and historical circumstances?

Based on the case study observations, conclusions include that dealing with the challenges has to be considered a cross-divisional task that affects the strategic orientation of a public administration. Furthermore, focusing on cooperation with both other municipalities could be regarded as a possible way to cope with the demography-related E-Government dilemma. Cooperation is a considerable option to reduce costs for public services. High fix costs for specific E-Government services and infrastructure can be divided and distributed among the cooperating partners. Parishes of those regions have to look over their individual municipal boundaries and work in cooperation on both front and backend matters. These models do not have to be entirely reinvented but are already present in some public administrations and can be extended to the sector of E-Government (Stopper, 2008). Role models for this can be Shared Services Centers (SSC) as used in private enterprise cooperation. These centers are used to combine and centralize administrative tasks of different parishes while at the same time retaining decentralized core units in the participating municipalities. The concept therefore combines the advantages of centralization and decentralization (Hensen, 2005). Furthermore, the case study data revealed an orientation towards alternative channels. When thinking about how to cope with aging-related challenges, literature oftentimes assumes the principle value of information
technology to the elderly (e.g. Namazi & McClintic 2003). The i2010 EU action plan on inclusive E-
Government calls for “no citizen to be left behind” (European Commission 2006c). Thus, asking the
questions if IT really helps the elderly and if the provision of services for elderly could easier and
cheaper be achieved using a different channel is important. A promising approach of establishing E-
Government services for the elderly is the provision of Mobile Citizen Services. In this model,
administrative services are provided through citizen consultants on public places such as e.g. hospitals
or senior residences. Thus, it “[…] combines aspects of e-government services as well as the idea of a
central administrative contact point (‘one-stop office’)” (Buß et al. 2003). A possibility of E-
Government cooperation combining front office and back office aspects is IT cooperation. The key
goal of this concept is bundling of competences rather than technological redundancy (Arendt and
Biel, 2004). This can include projects for creating E-Government portals for a whole area or sharing
IT infrastructure like storage and application servers.

3.) In how far do other settings show similar features and, thus, in how far is the case study setting
generalisable?
The majority of those cities suffering from deindustrialization faces similar demographic trends and,
therefore, is confronted with equivalent E-Government challenges. Nevertheless, also rural and
structurally weak areas are subject to comparable tendencies. For instance, many rural municipalities
in Eastern Germany face excessive depopulation and correlated aging. When looking at the
development in other industrialized countries, it is to say that the identified trends are not only specific
to Germany but can also be found in most other European countries (24 out of the 25 oldest countries
of the world, measured by the population share of above-64yr-olds, are to be found in Europe) and
Japan (the oldest country in the above mentioned ranking (Population Reference Bureau, 2006)).
Projections show that such demographic transition will continue in the future and affect an increasing
number of countries. Depopulation and aging will therefore be challenges to many public
administrations, not only in Europe.

4.) Are the conclusions drawn from the case study setting transferable to other settings?
The main conclusions address the establishment of different forms of cooperation and pursuing a
strategy of participation within E-Government services. Especially the cooperative aspect was a
central element in the case study setting in order to cope with the demographic challenge. Reference
models for cooperative approaches as identified in literature (Schroth, 2008) might help to overcome
existing boundaries that prevent seamless interoperability. Cooperation might become increasingly
important for public administrations because depopulation and aging will widen (see discussion point
3). In addition, such cooperative attempts, e.g. in terms of Shared Service Center or Public Private
Partnership, might not only be feasible means to address demography-related challenges, but might
help to address general ambitions to decrease costs and increase effectiveness. Cooperation, as
observed in the case study, might thus have demographic transition as one, but not the only reason.

6 CONCLUSIONS & FUTURE RESEARCH

Due to the demographic transition two major trends can be identified in Germany as well as in most
other industrialized countries, especially in Europe: a) Depopulation affects rural and structurally weak
areas as well as cities that are subject to deindustrialization. b) Aging often comes hand in hand with
depopulation as a consequence of selective emigration of younger inhabitants in search for jobs. While
depopulation led to a decrease in financial means for the affected public administrations, the aging
society demands for a change in the public service infrastructure. The needed reforms in both front
office and back office often cannot be taken due to the financial situation. One possible way out of this
dilemma is focusing on cooperating with other municipalities in order to share costs. Especially
Shared Service Centers are a promising approach. Limitations of the presented study include that a
single case was taken into account and that it did not focus on age-aware E-Government services
(front-office), but followed a rather broad perspective allowing back-office organization to form a
major part of the analysis (instead of E-Inclusion and Technology Acceptance where there lies the
greater literature body of knowledge).
Hence, future research could contribute a more front-office oriented perspective and connect such to the above given findings. Moreover, comparative research is necessary, taking into account multiple cases – not only from Germany, but also from other demography-transition affected countries such as Japan, Italy, or Sweden. Here, analyses could explore other means applied in order to cope with the demographic challenge. Furthermore, future research could address an evaluation of alternative channels for public service provision. For instance, it could incorporate the development of a cost-benefit perspective on channels, both online and offline. In addition, concerning the aging workforce, the development and implementation of suitable and cost-efficient knowledge management systems to ensure conservation of knowledge could be an important topic. These aspects indicate that the study area of demography and E-Government offers greatest potential for future research. Demographic transition constitutes one of the most severe challenges to our societies and economies and it is against this background, especially from an E-Government perspective, certainly under-researched. This reflects the fact the given case study directly and indirectly addresses two major future E-Government research themes: a) governance of public-private-civic sector relationships, and foremost b) E-Government in the context of socio-demographic change.

Acknowledgement

The authors gratefully acknowledge the financial support of the Volkswagen Foundation (Schumpeter Fellowship of second author). The title of this paper has been inspired by the article of Sorrentino & Ferro (2008): “Does the Answer to eGovernment Lie in Intermunicipal Collaboration? An Exploratory Italian Case Study”.

References

European Commission (2004a) eInclusion@EU: Strengthening eInclusion & eAccessibility across Europe. Analytic framework - eInclusion and eAccessibility priority issues.

Proceedings ECIS 2009
European Commission (2006c) Communication from the commission to the council, the European parliament, the European economic and social committee and the committee of the regions, i2010 eGovernment Action Plan: Accelerating eGovernment in Europe for the Benefit of All, Brussels.
The (Missing?) Value of IT in Public Organizations - The Case of the Swedish Rescue Services

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0687.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Evaluation, Business value of IT, IT/IS management, IT investment</td>
</tr>
</tbody>
</table>
THE (MISSING?) VALUE OF IT IN PUBLIC ORGANIZATIONS -
THE CASE OF THE SWEDISH FIRE RESCUE SERVICES

Frisk, Elisabeth, Gothenburg University, Dept. of Applied IT, 412 96 Gothenburg,
elisabeth@ituniv.se

Ljungberg, Jan, Gothenburg University, Dept. of Applied IT, 412 96 Gothenburg,
ljungbej@ituniv.se

Abstract

To actually gain organizational value from investments in Information technology (IT) is a widely known problem that many businesses today struggle with. This paper presents an interpretative case study of three public fire rescue service organizations in Sweden. The aim of the study is to investigate what problematic issues that could be raised in achieving organizational value of IT in public organizations. The study indicates that the absence of strategies and IT evaluation methods create fragmented, uncoordinated IT-investment activities, which in turn creates a frustration at all levels of the organization. The public sector invests heavy in IT, seemingly with poor management control. The consequences will be misuse of resources and unnecessary high IT costs. We argue that the use of formal evaluation methods would help the situation, but that economic methods are not enough as the sole solution. A combination of the economic and the Interpretative IT Evaluation approach seems more promising in order to capture both efficiency and effectiveness aspects.

Keywords: IT investments, IT value, IT management, Public sector, Evaluation.
INTRODUCTION

That IT investments not always deliver value or meet business and organizational goals is still a problem that concerns many businesses. An enquiry in Sweden 2006, answered by 307 IT managers within private and public sector concluded that 82% of all IT projects fail in delivering expected value, in 2005 the result was 72% (IT management, 2006). Furthermore, it is often not clear what the expected value is. This problem of achieving value of IT has in prior research been framed as the productivity paradox, originating from studies during the 1980s that found no connection between IT investments and the productivity (Solow, 1987; Brynjolfsson, 1993). Lately, the productivity paradox has been questioned, since there is also evidence that IT in fact provides positive impact on productivity (Dedrick et al. 2003). It has been argued that much of the uncertainty concerning the IT pay-off relates to weaknesses in measurement and evaluation in practice (Willcocks and Lester, 1996).

In this paper we report on an interpretative case study of the Swedish Fire Rescue Services where problematic issues are identified concerning organizational value of IT-investments among different stakeholder groups. We discuss what implications these findings will have on the concept of IT-value at firm level and on the evaluation of organizational value of IT. Thus, two questions are raised:

• What issues and problems are raised concerning organizational value of IT investments?
• What implications will these issues have on the organizational value of IT and how to assess it?

This study focuses on non-profit organizations in the public sector, from a management perspective, and a single IT investment. The research is conducted as an interpretative case study based on interviews in three Fire Rescue Service organizations. An IT investment is here defined as “any acquisition of hardware and software which is expected to expand or increase the business benefits of an organization IS and renders long-term benefits” (Apostolopoulos & Pramataris 1997).

The paper starts with a discussion of the concept of value and how it can be assessed. Then the research approach is outlined, followed by findings from the interviews. The findings are then used to further discuss the raised questions and finally a conclusion is given.

THE ORGANIZATIONAL VALUE OF IT

Since the role of IT initially, during the 1960s, was primarily to increase efficiency it is not strange that early evaluation methods (Hamilton and Chervany, 1981a, 1981b) mainly focused on measurable quantitative economic effects. Avgerou (2000) noted that since the role of IT has changed, the potential value of IT has also changed. The early computerization projects that substituted computer data processing for manual data processing had more or less clear efficiency objectives and intending to cause significant organizational change. These projects were assessed by using a cost-benefit analysis, and the value was related to an economic perspective. In the 1990s the value of IT was linked to the perception of IT as an enabler of organizational transformation, and assessing value of IT by a benefit-cost analysis seemed too limited (Symons 1991). Therefore in prior research economic methods based on economic theory has been criticized and instead evaluation methods based on the interpretative IT evaluation approaches have been suggested (Walsham 1999, Symons 1992, Jones and Huges 2001, Stockdale and Standing, 2006; Ward and Daniels 2006) Thus, the value of IT has moved from purely efficiency to an effectiveness enabler. But in practice management seems most concern about methods from an economic perspective (Ward and Daniel 2007).

Organizational value of IT as increased efficiency and effectiveness

The task of Management is according to Lewis et al. (2007) to “administrating and coordinating resources effectively and efficiently in an effort to achieve the goals of the organization”. Both effectiveness and efficiency are relevant measures to consider since the role of the involved organizations is to deliver services (pre-determined goals) to citizens in an efficient way. Efficiency
could be described as using the fewest inputs (people, material and money) to generate a given output and effectiveness concerns the degree to which goals are achieved (Lewis et al. 2007). Further Fitzgerald (1998) describes efficiency as doing things right. Productivity is a related concept that also has been widely discussed. According to Lewis et al. (2007) productivity is a measure of the efficiency with which the firm transforms inputs into goods and services. Effectiveness is a concept that can be seen as complementary to efficiency. Common to the different explanations is that effectiveness is related to the fulfilment of goals (Lewis et al. 2007) and doing the right things (Fitzgerald, 1998). These goals can be related to formal goals that have been stated in strategy documents, plans etc. (Modell and Grönlund 2006). One problem with this definition is that many organizations can survive long periods despite the fact that formal goals are not achieved. Within public organizations one common way to define effectiveness is whether formal goals are fulfilled or not. For a commercial organization, effectiveness may be judged by profitability, but for a public organization it would rather be judged by the fulfilment of mission and goals. But efficiency and effectiveness are often related to an economic perspective where efficiency includes measures such as productivity and effectiveness includes measures such as profitability (Lewis et al., 2007). Several researchers claim that an economic perspective is too narrow since the organizational value of IT by its nature is pluralistic (Guba and Lincoln 1990; Bannister 2001), multidimensional (Cronk and Fitzgerald 1999) and including a plethora of social issues.

2.2 Organizational value of IT as multi-dimensional and pluralistic

Value has been described by Guba and Lincoln (1990) as pluralistic and involves the view of different stakeholder groups. Cronk and Fitzgerald (1999), describe “IS business value” as the sustainable value added to the business by IS, either collectively or by individual systems, considered from an organizational perspective, relative to the resources required. Incorporating different stakeholders’ views, takes the organizational value of IT a step further compared to the pure economic focus.

Bannister (2001) argues that value of IT is multi-dimensional in its nature, encompassing economic, psychological, cultural and political aspects of value. According to Bannister (2001) we could distinguish three different concepts related to IT value: Values (with capital “V”, i.e. core values), value and benefit. Values are norms or modes of behavior that individuals, groups or organizations hold are right. They are visible in different cultural manifestations, in attitudes and beliefs, and in behavior. Value is then a “quality applied to a good, service or outcome which supports, meets or conforms with one or more of an individual or group’s Values” (Bannister 2001, p. 3). A corporation whose primary Value is to keep costs down and work in an efficient way, would value an IT-system that reduces head count. An organization like a fire rescue service, whose Value is to save life and property, may also value an IT-system supporting the operational work of fire fighters, i.e. increased quality. Benefits can then be seen as an operationalization of the Values. “Value is what we perceive, benefit is what we receive” (ibid.).

Given this brief discussion of value, it becomes evident that different stakeholder groups related to one organization will have different views on IT-value. Berghout and Remenyi (2005) note that the interpretative IT evaluation approach, considering different stakeholders views, has received the widest attention in the IT evaluation research field. However, in practice the economic approach is still the most popular (Ward and Daniel 2006; Frisk and Plantén 2004).

2.3 Organizational value of IT as economic and social

The economic approach takes an objectivist perspective, i.e. one presumes that there exists an objective reality, and that there is a subject-object dualism that implies that the observer can distance him from the phenomenon studied (Serafeimides, 2001; Guba and Lincoln, 1990). Thus, multiple interpretations of value, ambiguity, or other than economical dimensions cannot be addressed. Value in the economic approach can consider efficiency in terms of reduced cost but also effectiveness in
terms of profitability. This implies that both efficiency and effectiveness can be considered but only from an economic perspective. Also, these methods are summative i.e. they are not reiterative. In the research field IT evaluation the economic approach has received a lot of criticism. Serafeimidis (2001) argues that the economic stream suffers from a number of deficiencies such as limited consideration of organizational context and neglect of human aspects of evaluation. Jones and Huges (2001) are of the opinion that evaluation is a complex, multifaceted, difficult, continuous and social process. Further, economic evaluation methods have not kept pace with the shift in use of IS (Symons, 1991).

Since IT-evaluation is a social process, with multiple potentially conflicting views on the value of IT related to a specific context the interpretative IT evaluation approach has been put forward (Jones and Huges, 2001; Symons, 1991; Stockdale and Standing, 2006). The interpretative IT evaluation approach has been applied in varying shapes and forms: interpretive evaluation (Walsham, 1999), situated hermeneutic evaluation (Jones and Huges, 2001), CCP (concept, context and process) frameworks (Symons 1991; Stockdale and Standing, 2006), Benefit management (Ward and Daniel, 2006; Bennington and Baccarini 2004 Ashurst and Doherty 2003; Thorp 2001; Remenyi and Sherwood-Smith 1999) critical approach to evaluation (Klecun and Conford, 2005), to mention a few. In contrast to the economic approach, these IT evaluation approaches have been sparsely adopted by practitioners (Serafeimidis, 2001). The starting point of interpretative IT evaluation approaches is different stakeholders’ perception of reality (Walsham, 1995; Guba and Lincoln, 1990; Symons, 1991; Stockdale and Standing, 2006) and considers IT evaluation as a social process (Jones and Hughes, 2001) and value as an output from different stakeholders’ perceptions.

3 RESEARCH APPROACH

3.1 Organizational Setting: The Fire Rescue Services

On the national level the Swedish Rescue Services Agency (SRSA) is a government authority with expertise in different fields, including fire prevention. On the municipal level, the Fire Rescue Service is responsible for providing the population with services such as prevention, preparation, and response. The Fire Rescue Service is structured either as a Fire Rescue Service (FRS) or as a Fire Rescue Alliance (FRA). The main difference between FRS and FRA is that the former involves only one municipality. The relationship can be described as the FRS works in collaboration with the municipality it which the FRS acts. A FRA involves the collaboration of several municipalities but is structured and acts as its own municipal community. The economic pre-condition for FRA is given by the confederation (Förbundsdirektion) that is composed of politicians from the different municipalities that are part of the collaboration. The political goals are implemented in the responsible government agency (SRSA), and further by the municipal fire rescue services (FRS or FRA) in terms of:

1. The Legislated National Vision: "There must be fewer deaths, fewer injuries and less destroyed properties (10 years)
2. Political Priority, plan of action for the municipalities: Strategies for Prevention, to prepare, to carry out and to follow up (4 years)
3. Balance Scorecard, The Fire Rescue Services, Indicators within the four perspectives (year of plan)

The organizations involved in this case study differ in several aspects, see table 2. Also, organization C has outsourced the operation of IT to the municipality.

<table>
<thead>
<tr>
<th>Organization</th>
<th>Type</th>
<th>Employees</th>
<th>Fire Stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, FRA</td>
<td>Alliance with five municipalities’</td>
<td>1000</td>
<td>9</td>
</tr>
<tr>
<td>B, FRS</td>
<td>Large city</td>
<td>650</td>
<td>9</td>
</tr>
<tr>
<td>C, FRS</td>
<td>Middle sized municipality</td>
<td>150</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 1. Characteristics of the involved organizations
The main focus in this study was to interview the respondents with experience from IT investments used in the operative work. One example is Deadalos/IKAROS includes different modules such as: object register, resource handling and operative performance. Another system is RIB, a decision support system that includes information such as: how to deal with an accident, how the work of prevention can be planned, the risks when an accident has happened, and where the teams are located. Most of this kind of system can be related to decision support systems.

IT investments can be initiated in the budget process or spontaneously by employees or outside the organizations. Once a year a budget process takes place and then requests can come from different departments in the organization. The requests can be anything from a running band to a new IT system. Initiatives inside can also come spontaneously during the year from employees. Initiatives outside can be a new law or a system suggested from SRSA. The decision-making on IT investments is decentralized, and are often taken on the departmental level. But, if the costs exceed a specific amount, the decisions are taken by the officer of the FRS. Before such decision, the relevance of the IT investment is often discussed by the board. If the IT investment costs exceed regular budget restrictions, it becomes a political issue and the decision will be taken by the local government committee, appointed by the local municipality. Factors supporting the decisions are benefits, cost and techniques. Benefits are mostly argued from the requester’s perspective and not from an organizational point of view. The cost calculation consists mainly of costs for IT hardware, since costs for personnel are already a cost for the FRS/FRA. So, using the organizations own resources are not part of the cost calculation for an IT investment. Also, costs for running the system are seldom reflected upon.

### 3.2 Method

This is an interpretative case study (Walsham 1995) where data collection was done by semi-structured interviews and a workshop. Three organizations within the Fire Rescue Services in Sweden are involved. The study is based on the respondents’ perceptions and then the interpretation of the researcher (Klein and Myers, 1999). The study took place in 2005 and 2006, during a period of six months. It involved six to seven people in each organization. Eighteen persons from different organizations, levels and with different roles were interviewed in total (see table 2). In addition, one of the researchers participated over three days with a fire crew. During this time the researcher talked and interviewed the team members, approximately 20 persons. Managers at the top level and functional level selected the respondents attending the workshops. The managers had awareness of who had the experience from IT investments or assessing IT investments. The study involved at each organization six to seven managers from different levels. Examples of different roles attending the study are: At top level, Chief Manager, Vice Chief Manager and IT Manager; at functional level, Managers of department; at the operative level, Managers of the operative work force. Here means management someone who has an overall responsibility for the organization, the department, or for a team alternative the turnout.

<table>
<thead>
<tr>
<th>Strategic responsible</th>
<th>Organization A</th>
<th>Organization B</th>
<th>Organization C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development-responsible IT-responsible</td>
<td>Chief officer IT-strategist</td>
<td>Responsible for operational work</td>
<td></td>
</tr>
<tr>
<td>Functional responsible Projects-responsible</td>
<td>Department-responsible</td>
<td>Department-responsible IT-responsible</td>
<td></td>
</tr>
<tr>
<td>Operational responsible Front-line responsible Co-workers Firemen</td>
<td>Front-line responsible Co-workers</td>
<td>Front-line responsible Co-workers</td>
<td></td>
</tr>
</tbody>
</table>

*Table 2. Different roles of the respondents participating in the interviews.*

Each interview lasted for approximately two hours. The questions concerned their perceptions of the value of IT investments, how evaluation of IT investments is managed, why value of IT investments is
not achieved and how IT evaluation of IT investments might be improved. The interviews were recorded, transcribed, and finally analyzed. The analysis was performed by first reflecting upon the respondents’ answers and then clustering these into similar themes. First the responses within an organization were compared and then the findings were compared between the different organizations. After that the results were reported back to each organization in the form of text reports that was emailed to all respondents. A follow-up meeting was arranged where the results were discussed. The purpose with the report was to create understanding for if the findings were correctly interpreted. The respondents agreed on the results. No coding tool has been used. Walsham (2006) notes that using a coding tool can take too much attention itself so it is not always a necessary tool.

4 FINDINGS: CHALLENGES IN ACHIEVING VALUE FROM IT

From the interviews we did find a set of issues that tended to come up in various forms. We have here clustered them in five themes.

4.1 Diversification of IT value at different organizational levels

The perception of value of IT seems to differ between different organizational levels. On a strategic management level, IT is described as a cost rather than delivering value:

“…even if you try to rationalize with IT, some other costs pop up and you do not get the expected value”.

On a tactical level, IT is described as a valuable tool that gives information for better analyses. On an operational level, some respondents have difficulties in seeing the value of IT and one opinion is:

“you can’t fight fires with computers”.

However, several respondents also emphasize the importance of seeing value of IT from an organizational point of view and also related to the citizens. Today the perceptions were that IT value too often is seen from an individual point of view.

4.2 Narrow perspective on the value of IT

If IT investments are assessed from mainly a cost and a technical perspective and not from an organizational perspective, the risk is that decisions on IT investments are based on poor facts. It was considered important to pay attention to how the IT investment will affect the organization. Further, a “citizen perspective” is considered to be important in order to avoid only the internal focus, since most of the initiatives seem to have a cost, technical and/or a “gadget” focus:

“It is important to assess IT-projects initially if the IT-investment should benefit the organization and not only some individuals.”

“The internal discussion often concerns techniques, technical platforms and systems, but questions should be raised such as, what needs should be fulfilled, what do we want to achieve, and how should we proceed?”

“When we talk about IT we should talk about what we really would like to achieve in the public sector and not only technique.”

4.3 Lack of co-ordination

4.3.1 Lack of co-ordination on national level

NRSA do not give imperative guidelines to the Swedish Fire Rescue Service on what kind of IT, or how IT should be used in order to achieve legislation, national visions, or organizational goals. Instead
Each FRS/FRA is responsible for its own choice of IT solution. This has been criticized in the interviews since the perceptions are that each FRS/FRA has its own ad-hoc development workshop, and that this is both expensive and ineffective. Some collaboration exists but it could, according to the respondents, be room for more.

4.3.2 Lack of co-ordination on municipality level

Also decisions on departmental level are taken without any organizational co-ordination. This has contributed to an ad-hoc development of IT systems since the same problem or possibility can be solved by different technical solutions. One opinion is:

“Unfortunately we give priority to individual desires instead of the total picture. We can’t, for example, agree on one brand for digital cameras or digital calendars”.

Also, historically, the FRS/FRA’s are not so constrained by money. This has contributed to an ad-hoc development of IT since there are limited demands to show the benefits and costs of IT-investments. The respondents’ perceptions are that decisions on new IT investments are in favor of: “those who shout the loudest”. This brought about a culture that permits individuals to demand IT investments for their own interests, resulting in limited co-ordination and no alignment with the need of the organization.

4.4 Absence of strategy

Respondents from all organizations perceived that their organizations are lacking a visible strategy for the organization as well as for IT. According to the respondents this is one important cause for not achieving value from IT investments. Instead they suggested that organizational goals should direct the choice and decisions of new IT investments. One respondent stated that:

“We don’t know the plans for the coming two years, which means that the persons who are most anxious for new information systems or information technology get their requests approved.”

The respondents are of the opinion that the lack of organizational strategies as well as IT strategies, have contributed to an ad-hoc development of IT. This in turn has lead to a very difficult situation for the IT-department, since they need to be experts on many different systems. Also, the costs for IT operations are increasing fast.

Several of the respondents claimed that IT should contribute more to the organizational development by considering IT as a strategic resource. Today IT’s main function is to rationalize, and no discussion of IT’s strategic role is taking place.

4.5 IT evaluation issues

4.5.1 No consideration of the total costs

Actual cost for an IT investment should be further clarified. Costs for hardware and software were calculated, but seldom the cost for operational work, staff, education etc. Several respondents had observed that within their organizations there was a tendency to buy new IT instead of upgrading the old. The opinion was that a proper calculation of actual costs would gain a better understanding of the costs of used resources and give a better ground for comparing the benefits between investing in new IT or upgrade existing IT.

4.5.2 No follow-up of the IT investments

Today, IT investments are rarely followed-up and the perceptions are that there exist no guidelines for what criteria to consider, how to perform and when to evaluate. Some practical examples from buying IT systems that were not so successful, was when one of the organizations bought a Geographic
Information System, and a system called “Managing Risk and Analyses”. None of the systems are in use. The perceptions were that, despite a lot of good ideas, there were no process in place to take care of them:

“We had a lot of good technical ideas and solutions, but what do we do with it”?

Using a formal evaluation process were called for by several respondents:

“... not for catching the bad guys just to be able to tie up, to do a reflection and to learn.”

Another argument is to be able to do better priorities between the different IT investments and also between different kinds of IT investments.

4.5.3 Too few stakeholders in the evaluation process

The respondents perceive that no guideline that makes explicit who should be involved when evaluating IT investments exists. It was considered important to include the opinion of the users and the IT department in the evaluation process. The IT department will secure that the IT investment will fit into the infrastructure and existing systems, and also to eliminate the risk of unnecessary high operational costs for IT. A relevant question for several respondents were:

“Could the new system communicate with other systems or is it a different system which is built upon a different platform”.

Also, the respondents advocated that different stakeholders’ view should be involved in the evaluation process. Stakeholders’ meaning those who will be affected by the system, but also the “citizen perspective” is mentioned as an important perspective to consider.

4.6 IT affect power relations

Several respondents considered IT to be “charged”, i.e., IT gives people power. This could contribute to the development of IT on wrong grounds, being motivated by individual power interests instead of the need of the organization. The decisions-makers should be aware of these aspects, and therefore take an organizational view when evaluating new IT-investments, and not only a cost or technical perspective.

5 DISCUSSION: THE MISSING VALUE OF IT

We will now discuss the findings in terms of four themes: efficiency and effectiveness; the pluralistic nature of IT-value; the evaluation of IT-value; and IT-value in public organizations.

5.1 IT value as increased efficiency and effectiveness

One relevant question to ask is whether value of IT investments in terms of efficiency can be evaluated without any connections to the effectiveness i.e. overall goals and strategy of the organization?

Effectiveness has in overall been described as fulfilment of goals, doing the right things (Fitzgerald 1998). According to the respondents, effectiveness was not considered when evaluating value of IT investments. A risk when evaluating value of IT from only an individual perspective and neglecting the overall goal and context of the IT system, (Symons 1991; Jones and Huges, 2001; Stockdale and Standing 2006), is that the IT development becomes ad-hoc, which is a fact at particular one of the involved organizations. This ad-hoc development and lack of visible organizational goals could in turn contribute to transmission of resources (Modell and Grönlund 2006) and outputs that have no connection to the overall organizational goal and strategy. In the long run that give less valuable outcome, effects to the taxpaying citizen.
Organizational value of IT should include both a dimension of efficiency and a dimension of effectiveness i.e. do things right and do the right things. Also, since an IT investment affect different levels differently within the organization, such as the top-level, the tactical level and the operational level, and one level can receive benefits while another department or level can receive disadvantages, it is important to make the value visible for all levels. Another important factor to take into consideration is that a public organization cannot always dismiss resources, just allocate resources. Therefore effectiveness and efficiency need to interplay in the IT evaluation process, none should be considered by their own since the public organization is foremost driven by pre-determined political goals that should be achieved, and not by the “bottom-line”. Further efficiency should include both an economic and a qualitative dimension.

When respondents spoke about the need for organization they meant that IT investments should be related to overall organizational goals and strategy. But since no strategy was in place for the organizations, several respondents argued that IT investments have too much individual and internal focus instead of the mission for the citizen. When no strategy is in place there is a risk that IT investments mostly will be used for changing the use of resources’ and outputs without any connection to the organizations overall goal. Then, how is efficiency/productivity to be measured if the outputs are not related to the outcomes? This is an important question and in particular since there is a great uncertainty what could be consider as proper targets and measures in public sector (Modell and Grönlund, 2006). If no strategy is in place there may be a risk that core values may come in the background for individuals’ interest.

5.2 IT-value as pluralistic

The respondents describe value of IT from their own perspective and role in the organization. One exception was IT management, who related value to the overall goals and strategies of the organization i.e. effectiveness. On top level, the respondents (not IT management) relate IT-value to reduced costs, i.e. efficiency. On tactical level several respondents relate value of IT to better analysis of the organization and to improve the decision-making i.e. quality. On operational level first-line managers have difficulties to relate how IT has improved their work. Instead they perceive IT as time consuming since today the IT systems require more input of information than previous and the respondents were uncertain to what extent that information contributed of value to the organization. For example, the respondents have to give the same input to accident reports, irrespective of the size of the accident. Also the firemen seem more interested in investing in better vehicles that could contribute to both a more efficient and qualitative work. Most of the respondents also mention the citizen as an important stakeholder that should be considered in the IT evaluation process. But stakeholders such as government, politicians, community, union, suppliers, and other collaborators etc, received less attention when it comes to the value of IT investments. In the case of public sector organizations the pluralism of IT-value becomes especially complex due to the many stakeholders outside the organization. The value or effects of an IT investment needs also to be visible for all levels in the organization since value (output alt outcome) and input and output value is not always acting together at the same organizational level. Potential conflicts of interests are inherent in this pluralism both among different levels and groups in the organization, and between the organization and the stakeholders outside it, as well as between management and other stakeholders.

5.3 Implications for IT-evaluation

The identified problems such as strategic match, coordination of resources, proper calculation of costs, involving different stakeholder groups in the IT evaluation approach indicates that both the economic and the interpretative IT evaluation approaches complement each other since efficiency and effectiveness needs to be considered from an economic and stakeholders view. In prior research the economic approach has received a lot of criticism, e.g. for limited consideration of organizational context and neglect of the social and human aspects of evaluation (Serafeimidis, 2001, Jones and
Huges 2001; Stockdale and Standing 2006) However, the economic approach should not be abandoned rather it should be complemented. The economic methods are tools that provide valuable economic information such as costs and cash flow that is existential for many companies. The purpose is not to regard the organization as a social process it is to consider economic information. These methods can indicate future costs, cash flow, and must, continually be updated as other information. So, evaluating IT investments from an economic perspective will continue to be an important perspective in the IT evaluation process if it is used properly. The danger to such methods is that they can manipulate the cost items and the user can choose what items to add. Therefore, an Interpretative IT evaluation approach can complement such approach and create understanding from different stakeholder groups increased understanding for the organizational value of the IT investment.

In order to create understanding for organizational value, the use of resources must be related to the output and the outcome. The output compared to the use of resources can be described in a cost and benefit analysis. But the value must also relate to the outcome of the organization i.e. the goals, mission and strategy of the organization. This could be done by using an interpretative IT evaluation framework like CCP (Context, Concept, Process). CCP focuses on explaining: why (context) this IT investment should be done; according to the stakeholders what (content) should be followed up; the IT evaluation process (Serafeimidis, 2001; Jones and Huges, 2001; Guba and Lincoln, 1999; Remeny and Sherwood-Smith, 1999). The interpretative IT evaluation approach could therefore improve evaluating organizational value of IT investments in the public sector since the approach considers the context (which relates to strategy and co-ordination), content (involves different stakeholders view) and an IT evaluation process that follow up the IT investment in use (Symons 1991; Jones and Huges, 2001; Stockdale and Standing 2006).

5.4 IT-value in public organizations

The question of organizational value of IT-investments is in much a question of communication: how could one demonstrate and articulate the value for money from IT investments? If it is not communicated, it does not exist. With regard to public sector this often becomes a question of value for taxpayer money. In UK for example, a legislation of Best Value (BV) was introduced in year 2000. BV states that all public sector organizations (e.g. the police and fire services) must ensure that their IT-systems perform well and that they are delivered in a cost efficient way. Efficiency and economic aspects as value for taxpayer money is however not the only important question. Too much focus on economic issues will potentially downplay other important dimension such as the strategic value of IT. In the context of a public organization this may be related to outcome, not just output. In the case of the fire rescue services an example of outcome is “fewer deaths, fewer injuries, and less destroyed properties”. Relating national and regional visions to organizational value of IT is not a straightforward process, but without this discussion, the value of IT will remain unclear.

The general discussion of IT-value in public sector tends to be mainly focused on efficiency (do things right), and we have argued for a stronger orientation towards effectiveness (do the right things). However, if the value achieved is low in any of these two dimensions, the total value will be low as well. In order to create value for the public organization both internal and external factors needs to correlate since the organizations are driven by political goals. Output must correlate to outcome, strategy with the operational level. In summary it is important that IT-value is related to both the core operations (outcome) and to the operative work (output/efficiency) on different levels such as the top management, tactical and operational.
6 CONCLUSIONS

This paper has addressed issues in the achievement of the expected organizational value of IT, and what implications this may have for the assessment of that value.

Two points are important regarding IT value. First, to take the concept seriously, to problematize and define what is meant by IT value. Secondly, the usefulness of an IT-investment to the organizations should be in the foreground, not the value of technical features of the system. Furthermore, organizational value of IT should include a consideration of both efficiency and effectiveness. This is especially important in the public sector since it is driven by political goals and not by the “bottom line”.

Regarding IT-evaluation, the main point is that both economic and interpretative approaches are needed. Evaluating IT investments from an economic perspective with monetary items will continue to be an important perspective. The problem here is that the economic approaches are general, and not developed for specific use in IT-investments. The calculation of cost in these methods must be further developed. The interpretative IT evaluation approach could contribute positively to the IT evaluation process in the public sector since it considers strategic issues, co-ordination, stakeholders’ involvement and views evaluation as a formative process. Thus, it relates to several of the problematic issues identified in the study. Another issue that needs attention is how to take care of power and politics in the IT evaluation process.

The findings described above are particularly important for the public sector since no traditional income statement will validate the choice of used resources to the outcome. In a profit organization the income statement will sooner or later give indications if business is on the wrong track. Who will have this insight in a public organization? If a public organization does not co-ordinate its output to the outcome there could be a redistribution of resources to non-productive output, i.e. output not related to the expected outcome of the public organization. This could in turn be a serious waste of tax money on wrong output. Thus, achieving value is not only an evaluation problem it also demands management to address issues such as strategy, co-ordination, and cost control.

References


A DELPHI STUDY ON COLLABORATIVE LEARNING IN DISTANCE EDUCATION

Journal: 17th European Conference on Information Systems

Manuscript ID: ECIS2009-0134.R1

Submission Type: Research Paper

Keyword: Computer-mediated communication (CMC), Computer supported collaborative work (CSCW), E-learning, Collaboration
Abstract

Collaborative Learning (CL) is increasingly being used in Distance Education (DE), as it has been identified as an effective solution to known weaknesses such as high average rates of dropout and low quality of learning attainment. Information Technology is a core component of this type of learning as it not only provides the means to collaborate over distance but also has the potential to enable higher learning outcomes. There are a rapidly growing number of technologies in use today and the importance of these to collaborative learning initiatives, and the role they play, is an area of active research in the Information Systems (IS) community. IS educators and practitioners face an increasing challenge therefore to successfully implement CL in DE, precipitated not only from technical advances but also from wider social and organisational concerns. Using a Delphi study, this research is the first to investigate the factors that influence collaborative learning in distance education by surveying the opinions of an expert panel in this area. The aim was to produce an integrated list of the most important implementation factors and to investigate the role technology is perceived to contribute. The findings identified seventeen of the most important factors. These factors cover a range of themes including course rationale and design, instructor characteristics, training, group dynamics, the development of a learning community and technology. The potential of technology however does not seem to be fully realised and newer technologies such as multi-user environments would seem to be of limited use in practice, according to the expert panel.

Keywords: Collaborative Learning, Distance Education, CSCL, Delphi Method, IS Education.
1 INTRODUCTION

Distance Education (DE) is a broad term that refers to delivering a curriculum to learners who are not physically present on campus. Recent technological advances, along with changing learner demographics have triggered a significant rise in the popularity of this type of education. As distance education has evolved, it has become inextricably linked to technology (Garrison 1985) and as such information systems involving communication and information technology have become the underlying core of current and future DE innovations and trends (Lockwood 2001).

In conjunction with the advances in technology, DE is becoming increasingly popular throughout the world (Brandon and Hollingshead 1999; Oblinger 2001; Beldarrain 2006) with its market share growing rapidly (Salas, Kosarzycki et al. 2002). However there are still problems with its practice, including high average rates of dropout and low quality of learning attainment (Bernard, Rojo de Rubalcava et al. 2000). Various causes have been attributed to these problems including: feelings of isolation; lack of two or three way communication; procrastination; and difficulties associated with self-regulation (Bernard, Rojo de Rubalcava et al. 2000). New internet-based learning environments offer ways to overcome some of these shortcomings (Beyth-Marom, Chajut et al. 2003) for example, by providing means for asynchronous and synchronous learning with new technologies being used to foster student interaction and collaboration (Beldarrain 2006). Changes in the pedagogical approach are also happening with a move from instruction-based learning models to a constructivist (rather than being transmitted, knowledge is created, or constructed) or collaborative learning approach (Leidner and Jarvenpaa 1995). Numerous studies have demonstrated the superiority of collaborative learning (CL) over traditional modes of learning (for examples see Hiltz 1988; for examples see Alavi 1994; Leidner and Fuller 1997) and it has been identified as a potential solution to the weaknesses of traditional distance education courses (Bernard, Rojo de Rubalcava et al. 2000).

Collaborative learning emerges through the interaction of individuals with other individuals, knowledge is created through these interactions as individuals ‘exercise, verify, solidify, and improve their mental models through discussion and information sharing’(Leidner and Jarvenpaa 1995). Collaborative tasks can include decision-making, problem solving, report production, or experimental projects. When students work together collaboratively, they not only learn themselves, but they are also contributing to the development of the group (Salas, Kosarzycki et al. 2002). The collaborative learning model assumes that individuals’ learning is enhanced when they create knowledge by actively constructing a representation of the material being taught working on the assumption that ‘individuals learn better when they are forced to discover things themselves, rather than when they are told or instructed’(Leidner and Jarvenpaa 1995). Working in groups is not just a valuable way of learning but also develops the abilities for cooperative work, which is essential in the modern working place. Incorporating collaborative activities into modern distance education courses should produce graduates who can work effectively and efficiently with others, while also understanding the role of modern information technologies in collaboration, communication and knowledge creation.

Computer-Supported Collaborative Learning (CSCL) deals with how technology can be used to support collaborative learning and draws primarily from two contributing fields of inquiry; collaborative learning theory and Computer Mediated Communication (CMC) theory and research (Brandon and Hollingshead 1999). CSCL focuses on frameworks and tools to assist in collaborative learning along with the implementation of actual educational systems, collaborative learning activities and new pedagogical approaches. The use of technology therefore is more than a mere supporting infrastructural component. Technologies that promote communication and interactive can add value to the learning process by enabling the development of higher-order thinking skills, increased involvement, interest and motivation and overall the attainment of higher learning outcomes (Piccoli et al., 2001). Although emerging technologies offer a vast range of opportunities for promoting collaboration, distance education programs face challenges that may influence the implementation of these technologies (Beldarrain 2006).
Clearly, without technology the use of collaborative learning in distance education would be severely restricted however, technology is not the only influencing factor and the organisational processes and human interactions surrounding it are also of crucial importance in the implementation of CL in DE. A number of studies have investigated the factors which are relevant to CL in DE, mostly focusing on specific areas of interest, for example; CL and computer supported groups (Brandon and Hollingshead 1999); student preferences (Beyth-Marom, Chajut et al. 2003); social interaction (Kreijns, Kirschner et al. 2003); issues with CL in DE (Bernard, Rojo de Rubalcava et al. 2000); Computer Supported Collaborative Learning (CSCL) (for examples see Silverman 1995; English 1999; for examples see Beldarrain 2006); success factors of CMC technologies (Tolmie and Boyle 2000); system characteristics (Pituch and Lee 2006) and emerging themes in distance learning (Salas, Kosarzycki et al. 2002). The literature reviewed indicated that several factors impacted CL in DE, although each independent study considered only a limited number of factors. In general, the factors can be grouped into themes such as; course rationale and design; instructor characteristics; learning community; student characteristics; group dynamics; support and training; assessment and technology. Each of these themes have a number of factors associated with them, however an integrated list of a full range of the contributing factors does not seem to have been previously researched.

The aim of this research therefore was to develop an integrated list of the most important factors, with the aim of establishing the key issues involved in the implementation of CL in DE. Using the factors identified from the literature as a basis (see table 2.3), a panel of experts in the area ratified, expanded and ranked what they believed to be the most important factors. This study further elicited the view of the expert panel regarding their perception of the role technology plays in this type of education in order to explore whether the potential of information technology is recognised and fully exploited in practice.

2 RESEARCH DESIGN

A Delphi survey was chosen as the methodology for this study, as it is a technique for collecting opinions that aim to overcome the weaknesses implicit in relying on a single point of view. By involving a group of experts in the area, the study ensures that the results are not based on a single experience of collaborative learning in distance education. Delphi surveys have the benefits of group interviews without the need to gather the experts together in one location (Clayton 1997). They also allow the participants to express their opinion without undue pressure from others in the group, while the iterative nature of the technique provides the opportunity to clarify or change views based on others perspectives (Clayton 1997; Schmidt 1997). There are a number of permutations of the Delphi method and a ranking type Delphi study, designed to elicit the opinion of a panel of experts through iterative controlled feedback, was chosen for this study. The framework used was based on non-parametric statistical techniques, as outlined by Schmidt (1997), and aimed to answer the following research questions:

RQ1: What are the most important factors that influence the effective use of collaborative learning in distance education?
RQ2: What is the perceived role of technology in this form of education?

2.1 Panel Description

Rather than focus on a homogenous group (such as lecturers) a cross section of expertise was sought. This was to ensure that the factors identified as most important considered a range of perspectives and not just the views of a particular group.

The panel selected was composed of 18 panellists from three geographical regions: Ireland, USA, and UK. They represented collaborative learning through three distinct groupings; nine programme
directors using collaborative learning in distance education courses; five lecturers who are currently using collaborative learning techniques in distance education courses; and four academics with experience in the area of computer supported collaborative learning. Because the success of the Delphi technique relies upon the use of informed opinion, random selection was not considered when selecting the Delphi participants (Wicklein 1993). However, educational qualifications were taken into account, all of the panel were educated to a minimum of Masters Level, with fifteen having Doctorate qualifications.

Twelve universities were represented: Carnegie Mellon University, (United States); Middlesex University, (United Kingdom); New Jersey Institute of Technology (United States); National University Ireland, Galway (Ireland); Oscail, Dublin City University (Ireland); Penn State University (United States); University College Cork (Ireland); University of Edinburgh (United Kingdom); University of Leicester (United Kingdom); University of Hawaii (United States); University of Limerick (Ireland); and The Exploratorium, San Francisco (United States).

Criteria used in selecting the participants were based on their involvement with collaborative learning, distance education and computer supported collaborative learning. The average number of years experience in the area of Collaborative Learning was 7-10 years, with 50% of the panel having over 11 years experience. The panel were also highly qualified in the field of distance education with the average number of years experience 7-10 years and 44% of the panel having over 11 years experience. The average number of CL courses managed or taught was 4-7, as was the average number of DE courses managed or taught. The selected participants are considered to be well informed, leading authorities in their field by their colleagues, supervisors and peers. Overall the panel can be considered highly qualified and well equipped to provide opinions on the factors relating to CL in DE, as qualified by the following section.

2.2 Identifying Experts

The identification of ‘experts’ for the Delphi panel was based on a multiple-step approach suggested by Olaki & Pawlowski (2004). Initially a knowledge resources nomination worksheet (KRNW) was prepared which included details of the desired background or skill sets required. The KRNW was populated by going through each of the desired skill sets and identifying a key contact in the area. Personal contacts of both the researcher and supervisor were used as the initial contact point in each of the categories. Using the ‘snowball’ sampling method (Skulmoski, Hartman et al. 2007) the initial contacts were asked to provide recommendations for other potential participants in the study. The qualifications and skill sets of the named experts were ranked according to exposure and experience with collaborative learning initiatives and distance education. As a variety of perspectives were required the experts were ranked based on their experience and to ensure a selection of viewpoints. This provided a list of 46 experts to be invited to partake in the study.

These potential candidates were contacted, by email, and invited to participate in the study. The subject of the study was explained along with the procedure and the commitment required. To minimise non-response, one of the initial contacts introduced both the researcher and the topic of research to a number of the potential candidates, as suggested by Hsu and Sandford (2007). Delphi group size depends on group dynamics rather than statistical power and panels of 10-18 experts are recommended (Okoli and Pawlowski 2004). Of the forty six invitations, eighteen candidates accepted and this was considered ideal, as it would allow the study to start with the top end of the range. During the study two panel members dropped out, leaving a panel of sixteen who completed all stages of the study.

2.3 Survey Rounds

The initial survey was sent on the same day that an expert agreed to serve on the Delphi panel, as recommended by Okoli and Pawlowski (2004), using email as the communication channel. Following
the approach used by Kasi et al. (2008) rather than have the panellists participate in a brainstorming session, a list of potential factors identified from the literature was provided (table 2.3). This list provided the panellists with a structured instrument to begin the Delphi process and contained 28 factors for the initial ratification and discussion.

<table>
<thead>
<tr>
<th>Factor ID</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>Students should have prior experience of collaboration technology</td>
</tr>
<tr>
<td>F2</td>
<td>Promotive interaction should be encouraged within groups</td>
</tr>
<tr>
<td>F3</td>
<td>Students learning style should be conducive to group-work</td>
</tr>
<tr>
<td>F4</td>
<td>Tutor should assume facilitator role</td>
</tr>
<tr>
<td>F5</td>
<td>Tutor teaching style should encourage involvement and participation</td>
</tr>
<tr>
<td>F6</td>
<td>Group members should have adequate interpersonal skills</td>
</tr>
<tr>
<td>F7</td>
<td>Technology used should enable synchronous communication e.g. MSN chat and teleconferencing</td>
</tr>
<tr>
<td>F8</td>
<td>An appropriate rationale for collaborate learning should be developed</td>
</tr>
<tr>
<td>F9</td>
<td>Group work should promote individual accountability</td>
</tr>
<tr>
<td>F10</td>
<td>Course content should encourage interaction with both tutor and peers</td>
</tr>
<tr>
<td>F11</td>
<td>Tutors should have institutional support for their role</td>
</tr>
<tr>
<td>F12</td>
<td>The technology used should be accessible to all participants</td>
</tr>
<tr>
<td>F13</td>
<td>Course subject matter should include problem based tasks</td>
</tr>
<tr>
<td>F14</td>
<td>A consistent user interface should be provided</td>
</tr>
<tr>
<td>F15</td>
<td>The development of a learning community should be encouraged and nurtured</td>
</tr>
<tr>
<td>F16</td>
<td>Pre-course evaluation of learner profiles and learner needs should be carried out</td>
</tr>
<tr>
<td>F17</td>
<td>Tutors should prepare students to work collaboratively</td>
</tr>
<tr>
<td>F18</td>
<td>Social environments should be provided for non-project communication</td>
</tr>
<tr>
<td>F19</td>
<td>Course subject matter should encourage opinion diversity</td>
</tr>
<tr>
<td>F20</td>
<td>An appropriate rationale for use of Computer Mediated Communication technologies should be developed</td>
</tr>
<tr>
<td>F21</td>
<td>Technology used should enable asynchronous communication e.g. e-mail and bulletin boards</td>
</tr>
<tr>
<td>F22</td>
<td>Effective technical support should be provided to both tutors and students</td>
</tr>
<tr>
<td>F23</td>
<td>Group work should promote positive interdependence</td>
</tr>
<tr>
<td>F24</td>
<td>Students learning style should be conducive to sharing information with others</td>
</tr>
<tr>
<td>F25</td>
<td>Group size should be kept small (e.g. 4-5 students)</td>
</tr>
<tr>
<td>F26</td>
<td>Course subject matter should be discussion based</td>
</tr>
<tr>
<td>F27</td>
<td>Tutors should be trained for their role</td>
</tr>
<tr>
<td>F28</td>
<td>Group processing discussions should be encouraged</td>
</tr>
</tbody>
</table>

Table 2.3 Initial List of Potential Factors

Each participant was asked to select the factors that they deemed to be important to CL in DE. They were also asked to provide details of any factors that they considered to be important but were not on the original list. Along with the list of potential factors the panel were also sent a short survey outlining a number of technologies currently in use in the area. They were asked to identify those technologies that they currently used for CL initiatives and add any technologies that they use but were not on the original survey. The purpose of this additional questionnaire was to gain a snapshot of the technologies currently in use by the panel.

The results of the initial survey involved the panel removing 9 factors from the original list and adding a further 26. The additional items were reviewed and a new consolidated list of 45 factors was developed. As the target size for the ranking of the factors was around 20 items (Okoli and Pawlowski 2004) a second survey was required to narrow down the consolidated list. The questionnaire was randomly ordered to cancel out bias in the order of listing of the items (Brancheau, Janz et al. 1996). Each panellist was asked to select (but not rank) at least 10 of the most important factors (Schmidt 1997). Due to the nature of the topic and the diversity of the panel, a wide range of opinion on the factors was expected. An analysis of the results of the second survey suggested that opinion was
spread over a number of factors. A cut-off point of 40% was determined the most appropriate level for paring down the list as this ensured that the list contained a suitable number for ranking (less than 20, as suggested by Schmidt (1997) and also ensured that important factors were not dismissed at this stage. This list highlighted the 17 most important factors, from an initial list of 54 potential factors (original 28 plus the 26 added by the panel).

Once this list of most important factors was identified, the first of the ranking rounds was sent out. The aim of this phase was to determine the level of consensus on the ranking of the relevant factors. The pared list, from the second survey, was arranged in random order, and the respondents were asked to rank all the items (Schmidt 1997). Four different randomised versions were produced and divided among the panellists, as suggested by Branchan et al. (1996). The questionnaire also asked experts to submit comments explaining or justifying their rankings as it is suggested that experts ‘should arrive at consensus more quickly if provided with some sort of feedback about the panellists’ reasoning’ (Okoli and Pawlowski 2004).

The ranked lists were measured using Kendall’s $W$ coefficient of concordance, as it is recognised as one of the best ways for measuring non-parameter rankings (Schmidt 1997; Okoli and Pawlowski 2004). The values of $W$ range from 0 to 1, with 0 indicating no consensus, and 1 indicating perfect consensus. The value of $W$ obtained from this first ranking round was 0.148, which suggested weak agreement on the rankings and thus a second ranking round was necessary. As suggested by Okoli and Pawlowski (2004) the second ranking round was listed in order of the mean ranks obtained in the first round. For each item, the following information was provided to the panel: (1) an indication of the current level of consensus (based on the value of $W$) (2), the mean rank of the item; and (3) a column providing details of the views of the other panel members on each factor. Based on this, each expert was asked to revise their rankings for each item, again asking them to explain their rankings and revisions. A separate column was provided for these additional comments.

The response to the second ranking round indicated that the majority of panellists did not wish to change their opinion, with only four of the panel members revising their rankings. However, a number of additional comments were obtained and Kendall’s $W$ improved to 0.221. At this stage it was decided that further ranking rounds would not be required. Dissensus, or lack of agreement is a valid finding for a Delphi study (Skulmoski, Hartman et al. 2007). The diversity of the panel, and the subjective nature of education, highlighted by the detailed comments obtained in the survey, clearly identified a number of viewpoints on the factors considered most important. The panel had agreed on the most important factors, the lack of strong consensus was on the priority, or ranking, of these factors.

The first ranking round also included an additional questionnaire, which detailed the technologies currently being used by the panel (identified in the first stage of the survey) and asked each member to rate (using a likert type scale) how useful each technology was to collaborative learning.

3 FINDINGS

3.1 Most Important Factors

This study identified the top 17 of the most important factors from a comprehensive list of 54. The following table provides the results of the final ranking round and outlines the factors in ranked order, along with their mean rank and interquartile range (IQR). The IQR shows the range of opinion on the ranking of the factor; the higher the IQR the greater the range of opinion.
<table>
<thead>
<tr>
<th>Rank</th>
<th>Description</th>
<th>Mean Rank</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Instructional design of the activity, activity structure and assessment needs to promote CL</td>
<td>5.80</td>
<td>7.00</td>
</tr>
<tr>
<td>2</td>
<td>Tutor teaching style should encourage involvement and participation</td>
<td>6.20</td>
<td>4.00</td>
</tr>
<tr>
<td>3</td>
<td>The development of a learning community should be encouraged and nurtured</td>
<td>6.47</td>
<td>8.00</td>
</tr>
<tr>
<td>4</td>
<td>The technology used should be accessible to all participants</td>
<td>6.47</td>
<td>10.00</td>
</tr>
<tr>
<td>5</td>
<td>Tutor should assume facilitator role</td>
<td>6.67</td>
<td>7.00</td>
</tr>
<tr>
<td>6</td>
<td>Personalised, detailed and quality-controlled feedback on assessment work should be provided</td>
<td>7.20</td>
<td>7.00</td>
</tr>
<tr>
<td>7</td>
<td>An appropriate rationale for collaborate learning should be developed</td>
<td>7.47</td>
<td>9.00</td>
</tr>
<tr>
<td>8</td>
<td>Tutors should be trained for their role</td>
<td>7.53</td>
<td>8.00</td>
</tr>
<tr>
<td>9</td>
<td>Promotive interaction should be encouraged within groups</td>
<td>9.13</td>
<td>4.00</td>
</tr>
<tr>
<td>10</td>
<td>Group work should promote positive interdependence</td>
<td>10.00</td>
<td>7.00</td>
</tr>
<tr>
<td>11</td>
<td>Learning environment should be user friendly and kept simple</td>
<td>10.47</td>
<td>9.00</td>
</tr>
<tr>
<td>12</td>
<td>Prior design of collaborative tasks is essential: i.e. design for learning, then e-moderate for participation</td>
<td>10.60</td>
<td>7.00</td>
</tr>
<tr>
<td>13</td>
<td>The development of teamwork skills should be explicitly build into the instructional design</td>
<td>10.93</td>
<td>8.00</td>
</tr>
<tr>
<td>14</td>
<td>Technology used should enable multiple means of communication</td>
<td>11.07</td>
<td>8.00</td>
</tr>
<tr>
<td>15</td>
<td>There should be lots of opportunity for social communications in the early part of the course</td>
<td>11.27</td>
<td>4.00</td>
</tr>
<tr>
<td>16</td>
<td>Technology used should enable asynchronous communication</td>
<td>11.87</td>
<td>8.00</td>
</tr>
<tr>
<td>17</td>
<td>Tools should support multiple learning styles</td>
<td>13.47</td>
<td>5.00</td>
</tr>
</tbody>
</table>

Table 3.1 – Results of Delphi Ranking Round

By reviewing the comments and rankings for each factor received, it was possible to group them into the themes originally identified from the literature and classify the most important factors within these themes.

- **Course rationale and design** - the design of the activity needs to promote collaborative learning and (#1) there needs to be an appropriate rationale behind its use (#7). Prior design of the collaborative tasks is also considered important (#12) along with the development of teamwork skills explicitly built into the design (#13).

- **Instructor characteristics** – the instructor teaching style should encourage involvement and participation (#2) with the tutor assuming a facilitator role (#5).

- **Learning community** - the development of a learning community should be encouraged and nurtured (#3) and while there should be opportunity for social communications in the early part of the course (#15), this is not seen as high a priority.

- **Assessment** – personalised, detailed and quality-controlled feedback on assessment work is seen as of high importance (#6), but considered relevant to all forms of education and not just CL.

- **Support and training** – tutors should be trained for their role (#8) and this is seen as an important factor. Support, both technical and institutional, is not deemed to be as high a priority and did not make it to the list of most important factors.

- **Group dynamics** - Promotive interaction refers to the quality of interaction among group members and is seen as important (#9) along with positive interdependence, which means that group members must see value in working together for collaborative learning to occur (#10). Group size and the interpersonal skills of the group, are not perceived to be as high a priority and while ratified from the literature did not make it through to the most important factors.
Technology – of high importance is the fact that the technology should be accessible by all students (#4), and while this may be considered obvious, is an essential component to CL in DE. Also the learning environment itself should be user friendly and kept simple (#11). Other technological factors are also considered important: it should enable multiple means of communication (#14), including asynchronous communication (#16) and should support multiple learning styles (#17). While technology factors did come in near the bottom of the final rankings, it is clear that they are still considered highly important as they made up almost 30% of the final list of ‘most important’ factors.

3.2 Collaborative Learning Technology

The technology questionnaire sent out with the Delphi study further provides a snapshot of the technologies currently in use for CL in DE, and their perceived usefulness. The following table provides a summary of the results.

<table>
<thead>
<tr>
<th>Technology</th>
<th>% of Panel using this technology</th>
<th>Average Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>VLE / Online Forums / Bulletin Boards</td>
<td>100%</td>
<td>Extremely Useful</td>
</tr>
<tr>
<td>Chat Functions / Synchronous Discussion</td>
<td>94%</td>
<td>Moderately Useful</td>
</tr>
<tr>
<td>Computer / Audio Conferencing / VoIP (e.g. Elluminate, Skype)</td>
<td>83%</td>
<td>Moderately Useful</td>
</tr>
<tr>
<td>Collaborative document tools (e.g. Google docs, Word comment)</td>
<td>78%</td>
<td>Extremely Useful</td>
</tr>
<tr>
<td>Email / Email List Server</td>
<td>72%</td>
<td>Moderately Useful</td>
</tr>
<tr>
<td>Wiki Spaces</td>
<td>67%</td>
<td>Moderately Useful</td>
</tr>
<tr>
<td>Social Networking Software (e.g. Ning, del.icio.us, wiki, facebook)</td>
<td>61%</td>
<td>Limited Use</td>
</tr>
<tr>
<td>Blogs</td>
<td>56%</td>
<td>Moderately Useful</td>
</tr>
<tr>
<td>Calendars, Agendas or Schedules</td>
<td>56%</td>
<td>Moderately Useful</td>
</tr>
<tr>
<td>Voting</td>
<td>50%</td>
<td>Limited Use</td>
</tr>
<tr>
<td>Multi-User Virtual Environments (e.g. Second Life)</td>
<td>44%</td>
<td>Limited Use</td>
</tr>
<tr>
<td>Podcasting</td>
<td>44%</td>
<td>Moderately Useful</td>
</tr>
<tr>
<td>Group conferencing (with synchronous audio / video) / Video Conferencing</td>
<td>40%</td>
<td>Limited – Moderate Use</td>
</tr>
</tbody>
</table>

Table 3.2 Technology and Collaborative Learning

It would seem that Virtual Learning Environments, including online forums and bulletin boards are of most use to collaborative learning in distance education. The entire panel uses this technology and it received a high level of support on its usefulness. Collaborative document tools are also considered ‘extremely useful’ by the panel, with 78% of them using this technology. The majority of the technology is considered ‘moderately useful’ including audio conferencing and email, even though these had a high percentage of usage. It is interesting to note that some of the newer technologies such as Multi-User Virtual Environments and group conferencing appear to have limited usefulness in practice according to the expert panel.

4 DISCUSSION

The Delphi study enabled a group of experts in the area to identify and discuss the most important factors contributing to CL in DE. Using a base list of 28 factors identified from the literature, the panel
suggested an additional 26 factors based on their experience in the area. These factors were ratified and discussed during the rounds of the Delphi exercise, cumulating in agreement on 17 of the most important factors, albeit with differing opinions on the rankings of these factors.

4.1 Most Important Factors

In line with the literature (Brandon and Hollingshead 1999; English 1999; Tolmie and Boyle 2000; Kennedy and Duffy 2004) course rationale and design is considered highly important. Personalised, detailed and quality-controlled feedback on assessment work should be provided as it is also seen as of high importance. The literature also suggests that the rationale behind the use of Computer Mediated Communication (CMC) technologies is considered important (Tolmie and Boyle 2000), however this study did not find it to be a high priority, perhaps because in DE it is necessary to use CMC technologies. While there are a number of suggestions around course content discussed in the literature (Silverman 1995; Brandon and Hollingshead 1999; Bernard, Rojo de Rubalcava et al. 2000) course subject matter was not found to be particularly important by this panel of experts.

The role of the tutor, or instructor, is significant to CL, with ‘teaching style’ considered to be a ‘most important influence’ on involvement and participation (Salas, Kosarzycki et al. 2002). The expert panel would seem to agree with this as one of the highest-ranking factors directly related to instructor teaching style, suggesting that it should ‘encourage involvement and participation’. Learner-centred courses require the instructor to assume a facilitator role, and again this is suggested as important in the literature (Silverman 1995; English 1999; Bernard, Rojo de Rubalcava et al. 2000). The panel concurs with this agreeing that the tutor should assume a facilitator role, and also receive training for the role. Support, both technical and institutional, is not deemed to be as high a priority and did not make it to the final list of most important factors.

It would also seem that, in line with the literature (Hiltz 1998; Kreijns, Kirschner et al. 2003), the encouragement and development of a learning community is considered highly important to the effective use of CL in DE. While there should be opportunity for social communications in the early part of the course, this is not viewed as high a priority. The importance of the group to collaborative learning has been regularly discussed in the literature (for examples see (Hiltz 1988; Brandon and Hollingshead 1999; English 1999; Tolmie and Boyle 2000; Kreijns, Kirschner et al. 2003); the results of this study would seem to concur with the literature, at least with regard to the importance of promotive interaction and the fact that group work should promote positive interdependence. However, other aspects deemed important in the literature, for example, group size (Tolmie and Boyle 2000; Kreijns, Kirschner et al. 2003), were not supported by this panel. This may be due to fact that the technology available today allows large groups to work quite well. As one panelist pointed out ‘perhaps bigger groups enable the work to be sustained when some of the group are inactive for long periods’. Group interpersonal skills were also not deemed to be of major importance, while individual accountability was seen as important but not ‘most important’ and did not go through to the final list.

Student characteristics are regarded as important in the literature with Bernard et al. (2000) suggesting that that ideally ‘developing a profile of the learner’s knowledge, skills and experience, as well as their perceived needs’ will aid in the design and implementation of effective DE courses. Learner differences involve both the way that students will interact with the technology as well as affect the degree to which they will participate in online collaboration activities (Salas, Kosarzycki et al. 2002). However, the panel did not concur with the literature; they did not consider student characteristics an important factor in the effective use of CL in DE. It was agreed that while students may have a preference for certain learning styles these can be overcome and adapting learning styles to suit course requirements is part of the learning experience itself.
4.2 Role of Technology

The results of this study highlight the importance of technology by identifying five factors as being among the most important factors. Accessibility to the technology has been highlighted as important in the literature (Bernard, Rojo de Rubalcava et al. 2000) and this has been upheld by the panel, placing this factor among the top five most important factors. Enabling asynchronous communication (Silverman 1995) was also considered important enough to make it through to the final list. The panel members themselves added the three other technological factors in the initial phase.

The lower ranking of the technology factors would seem to be due to the fact that it is deemed a supporting role and less critical than good design and tutor characteristics. While not considering the technology unimportant, the point was made that:

‘with a good plan, and buy-in from teachers, the technology should not matter. Of course there is a need to match the technology to the task – but perhaps we are getting to the position that we are doing this in what seems like an intuitive and natural way. The technologies are (relatively) mature and powerful, so we can achieve our ends with a range of different technologies’.

While the above comment recognises that the technologies are now relatively mature and powerful, it would seem that technology is not being fully utilised. One factor that had a wide discordance concerned the ability of technology to cater for different learning styles. While the majority of the panel did not consider this to be a high priority, one panel member fully supported this and felt that the rest of the panel was overlooking it.

‘Tools need to reflect the multiple styles for learning and not assume that students should adapt to purely linguistic ones. This doesn’t mean we have to test and understand every student’s primary learning style, only that we have to design environments that appeal to multiple styles in a variety of ways’.

Alavi and Leidner (2001) point out that ‘the role of IT in enabling individualized learning methods, while not new, has received strikingly little attention’; this study supports this view and further indicates that the potential of IT has still not been recognised in practice. The potential of technology to provide an individualised, effective learning environment is not recognised or utilised by the majority of this panel of experts. A recent study (Menchaca and Bekele 2008) identified that ‘the availability of multiple tools added flexibility to the learning environment’ which helped ensure a successful DE programme, as did the use of technology tools that appeal to multiple learning styles. Perhaps as more research identifies the usefulness of these technological tools they will be utilised more fully in practice. From an IS perspective it would seem that in practice, there is suboptimal use of technology in this educational environment. In particular, newer technologies such as multi-user environments, group conferencing and social networks are perceived to be of limited use in practice. These technologies have the potential to enable collaborative learning to take place over distance and as such their perceived lack of usefulness is of concern. If these technologies are to be fully optimised as an enabling factor in collaborative distance education then their educational benefits need to be more strongly highlighted to practitioners.

5 CONCLUSIONS

The motivation behind this study was to develop for the first time a comprehensive list of the most important factors that influence the effectiveness of collaborative learning in distance education. As an Information Systems study the role that technology plays was of particular interest and part of the study included pulling together a snapshot of the technologies currently in use in the area. A panel of experts worked together to ratify and consolidate the factors and the study achieved its aim of producing the integrated list. While the actual rankings of the factors did not achieve strong group consensus the lack of agreement highlighted the diverse views and opinions on which factors are of highest priority. The study highlighted the 17 most important factors and also provided an understanding of the perceived role of technology in this educational area.
5.1 Implications for IS Research

While the study established that technology is among the most important factors, it also highlighted the fact that it is viewed more in a supporting role than as an enabling factor. In particular, the use of technology to support multiple learning methods is an area that is not currently optimising the potential of technology. Further research into the role of technology in collaborative learning might consider how the technology is being used and why it is not being fully utilised in practice.

The perceived usefulness of the technologies could also be further explored, in particular the lack of support for some of the newer technologies. The fact that Multi-User Environments are used by less than 50% of this panel, along with the suggestion that it is of limited use warrants further research. As these new technologies can actually enable collaborative learning to take place, rather than just support it, it is important to establish why they are not being considered particularly useful in practice.

As with any Delphi study, the results are based on a limited number of subjects. While these subjects were chosen following rigorous guidelines, one must be cautious in generalising. The sample is relatively diverse in terms of perspectives on CL in DE and this diversity may have influenced the lack of consensus on the priority of the items. Further research could be carried out with panels of similar perspectives to determine if the results would hold.

5.2 Implications for practitioners

The main aim of this research was to develop an integrated list of the most important factors, relevant to the effective use of collaborative learning in distance education. The opinion of an expert panel provides practitioners with a comprehensive, expert opinion on the factors that are deemed highly important in the area. This final list of 17 factors provides details of the areas to focus on when implementing such an initiative.

Practitioners can also use this study to gain an insight into the role technology plays and gain an understanding of its importance, based on the opinion of experts in the field. The snapshot of the technologies currently in use can provide details of which technologies are perceived to be most useful along with those which are considered of limited use.

Overall the study provides a practical guide for those considering implementing collaborative learning in distance education, along with some motivation for future research for the IS community regarding the suboptimal utilisation of technology in practice.

6 REFERENCES


A ‘USES AND GRATIFICATIONS’ APPROACH TO UNDERSTANDING THE ROLE OF WIKI TECHNOLOGY IN ENHANCING TEACHING AND LEARNING OUTCOMES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0184.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Wikis / Wikipedia, Technology adoption, Learning, Constructivism</td>
</tr>
</tbody>
</table>
A ‘USES AND GRATIFICATIONS’ APPROACH TO UNDERSTANDING THE ROLE OF WIKI TECHNOLOGY IN ENHANCING TEACHING AND LEARNING OUTCOMES

Zixiu Guo, School of Information Systems, Technology and Management, University of New South Wales, Sydney NSW 2052, Australia, z.guo@unsw.edu.au

Ying Zhang, School of Information Systems, Technology and Management, University of New South Wales, Sydney NSW 2052, Australia, zhang.ying@student.unsw.edu.au

Kenneth J Stevens, School of Information Systems, Technology and Management, University of New South Wales, Sydney NSW 2052, k.stevens@unsw.edu.au

Abstract

The use of the Wikis in both post-graduate and undergraduate teaching is rapidly increasing in popularity. Much of the research into the use of this technology has focused on the practical aspects of how the technology can be used and is yet to address why it is used, or in what way it enhances teaching and learning outcomes. A comparison of the key characteristics of the constructivist learning approach and Wikis suggests that Wikis could provide considerable support of this approach, however research into the motivations for using the technology is required so that good teaching practices may be applied to the use of Wikis when utilized in the higher education context. This study articulates a research design grounded in the Technology Mediated Learning (TML) paradigm that could be used to explore teachers and students’ motivations for using Wiki technology to enhance teaching and learning outcomes. Using the ‘Uses and Gratification’ approach, a popular technique used for understanding user motivation in technology adoption, a two-stage research design is set out. Finally, the paper concludes with a discussion of the implications for both information systems researchers and higher education.

Keywords: Wiki technology, constructivist learning, uses and gratifications approach (U&G), technology-mediated learning (TML), motivations
1 INTRODUCTION

The introduction of Web 2.0 technologies into the university teaching environment is transforming the traditional e-learning world (O'Reilly 2006). The shift from static Web 1.0 technologies to the interactive potential of Web 2.0 provides the opportunity of improving the learning behaviors of the learners as they are no longer just consumers of preconfigured content, but participants in the creation, collation and sharing of the learning content, playing a much more important role in the entire learning context (Ebner & Walder 2007).

Of the various applications that fall under the Web 2.0 umbrella, Wiki is the one that has most rapidly proliferated in higher education (Choy & Ng 2007, Elgort 2007). As “a social networking adaptive technology that emphasizes a more task-oriented collaborative editing of content and development of ‘collective’ interlinked knowledge” (Duffy & Bruns 2006), Wikis have been put to a wide range of uses in a broad range of contexts. Research into Wikis use in academia is relatively new (Parker & Chao 2007) and although some research has examined Wikis in learning (Bruns & Humphreys 2005, Cook 1998, Duffy et al. 2006, Elgort 2007, Elgort, Smith & Toland 2008), these studies have tended to focus on characterizing Wiki learning activities or establishing guidelines to aid in the implementation of Wikis (Forte & Bruckman 2007).

Two significant gaps are evident in the research. Firstly, there is a gap in the understanding of the use of Wikis in learning as it relates to people’s motivations with respect to that use (Wagner 2004). This deficiency mirrors that found in most research into Technology-Mediated Learning (TML), in which there is seen to be little consideration of the important internal social and psychological process through which learning occurs (Alavi & Leidner 2001). These authors call for greater depth and breadth of TML research in the Information Systems (IS) domain in order to understand how technologies can enhance learning. In this regard, the ‘Uses and Gratifications’ (U&G) perspective to the study of technology use has been found to offer some insight into the reasons why people adopt an emergent technology (Flanagin & Metzger 2001), although the level of granularity used when studying the Internet may have damped the results. The other ‘gap’ relates to understanding the way in which Wikis enhance a learner’s learning as, despite its use as a teaching tool, it is not clear how, and in what contexts, a learner’s use of Wikis enhances his/her learning (Forte et al. 2007).

It is understandable that these areas are yet to be investigated, given the newness of Web 2.0 technologies in university teaching, but as Flanagin and Metzger (2001, p154) point out, “this rate of change, however, only underscores the importance of a rigorous examination of new communication technologies’ development and use.” This study sets out a research design by which these two knowledge gaps can be investigated and hopefully addressed. It employs the U&G approach to examine (1) what motivates educators and students to use Wiki technology in teaching and learning, and (2) which groups of motivations does Wiki technology fulfil best? The paper commences by defining Wiki technology and discussing its adoption in the education sector. An overview of good teaching practices then follows. The potential role of Wikis in that practice is then explored followed by an explanation of U&G perspective. The research questions are then posed and the research design by which the proposed research questions could be investigated is then articulated and discussed. The paper concludes with an outline of the anticipated implications of the proposed research.

2 WIKI TECHNOLOGY: DEFINITION AND USE IN EDUCATION

Wiki, a term that means “quick” in Hawaiian, was originally developed by Ward Cunningham and further defined by Leuf and Cunningham as a “freely expandable collection of interlinked Web pages, a hypertext system for storing and modifying information—a database where each page is easily editable by any user with a forms-capable Web browser client” (Leuf & Cunningham 2001, p14).
According to Duffy and Bruns, (2006), the key characteristics of Wikis are:

- A website that allows a user to add content, and allows that content to be edited by any other user.
- Can be personal, but are usually open to collaborate.
- Involve the creation of documents without the user having a detailed knowledge of HTML.
- Tend towards expressing ideas as relationships between pages, thus creating a network of interrelated topics that is based on a topical approach.
- Are a-temporal, that is, the nodes (or interlinking textual references) change not according to time, but by way of development of the evolving and edited text.
- Track the changes to individual pages over time and allow users to browse the history of a page.
- Encourage cross-linking and are dominantly spatial in structure.
- Provide a space where knowledge becomes networked (situated, contextualized) but remains ephemeral; it changes, and can be changed and mediated by the community.

Wiki technology has been found to be useful in teaching, especially where collaboration and knowledge sharing are important (Raman, Ryan & Olfman 2005) and has been used to support activities such as project management (Xu 2007), developing and maintaining software projects (Malani & Dwyer 2005), supporting writing instructions (Lamb 2004), arranging information and sharing knowledge (Elgort et al. 2008), online teaching and assessment (Bruns et al. 2005), and online collaboration in the e-learning environment (Raitman, Augar & Zhou 2005).

3 GOOD TEACHING PRACTICES

In seeking to understand the role that Wikis can play in higher education, this study assumes a constructive perspective of learning, with the deep approach to learning and student-focused approach to teaching being important contributions of the constructivist paradigm of learning.

A student’s approach to learning is typically categorized as either a surface learning approach or a deep learning approach. The surface approach to learning involves the students seeking to avoid failure but having no inherent desire to work too hard. This approach usually leads to satisfactory performance in assessment but to very poor learning outcomes over the longer term (Biggs 2003, Gibbs 1992, Ramsden 1992). The characteristics of the surface approach to learning include: the intention to complete task requirements; memorizing information needed for assessment; the failure to distinguish principles from examples; treating task as an external imposition; focusing on discrete elements without integration; and unreflectiveness about purpose or strategies (Entwistle 1987). In contrast, the deep approach to learning involves students being intrinsically motivated, seeking integration between components and between tasks, playing with ideas (Gibbs 1992), engaging with the materials of the subject and re-constructing their own world view as a consequence of what they have learned (Biggs 2003, Gibbs 1992, Ramsden 1992). The major characteristics of the deep approach to learning include intention to understand; vigorous interaction with content; relating new ideas to previous knowledge; relating concepts to everyday experience; relating evidence to conclusions; and examining the logic of the arguments (Entwistle 1987). The deep learning approach is seen as superior as it is considered to produce better performance in assessment and deliver better learning outcomes over the longer term (Biggs 2003, Gibbs 1992, Ramsden 1992).

An educator’s approach to teaching can be classified into five different types (Trigwell, Prosser & Taylor 1994):

1. A teacher-focused strategy with the intention of transmitting information to students;
2. A teacher-focused strategy with the intention that students acquire the concepts of the disciple;
3. A teacher/student interaction strategy with the intention that students acquire the concepts of the disciple;
4. A student-focused strategy aimed at students developing their conceptions;
A student-focused strategy aimed at students changing their conceptions.

The teacher-focused strategy focuses on teachers and represents teaching as being mainly about passing on knowledge, whereas student-focused strategy focuses on students and represents teaching as helping the students develop their own knowledge (Trigwell et al. 1994).

The teacher-focused strategy has been found more likely to encourage students towards a surface learning approach than the student focused strategy (Trigwell & Prosser 1996), hence suggesting that the student-focused approach to be the preferable approach. Trigwell, Prosser and Waterhouse (1999) give the key characteristics of student-focused teaching strategy as:

- The teacher believes that what the student does and not what the teacher does determines what the student learns;
- The teacher is one who encourages self-directed learning;
- The teacher is one who makes time for students to interact and to discuss the problems they encounter;
- The teacher is one who assesses to reveal conceptual change;
- The teacher is one who provokes debate and raises and addresses the taken-for-granted issues;
- The teacher is one who uses a lot of time to question students’ ideas, and to develop a “conversation” with students in lectures.

These views are consistent with Biggs’s writing on how a learning environment can be transformed to achieve quality outcomes: good teaching as the encouragement of a deep approach to learning (Biggs 2003). The key component in this transformation is the transition from teacher-centered practice, where student learning is seen as a result of what the teacher does, to student-centered practice, where student learning occurs as a result of what the student does (Housego & Freeman 2000). Four key elements being identified as a good teaching are (Biggs 1989):

- Motivational context: deep learning is more likely when students’ motivation is intrinsic and when the student experiences a need to know something;
- Learning activity: students need to be active rather than passive since deep learning is associated with doing;
- Interactions with others: it is often easier to negotiate meaning and to manipulate ideas with others than alone; and
- A well-structured knowledge base: without existing concepts it is impossible to make sense of new concepts.

Both the deep approach to learning and student-focused approach to teaching are important contributions of constructivist paradigm of learning. In this paradigm knowledge as well as meanings are seen as constructed rather than given (Jonassen 1999, Parker et al. 2007) and are seen to embody a number of key principles (Miers 2004, p.4):

- Active and manipulative: engaging students in interactions and explorations with learning materials and providing opportunities for them to observe the results of their manipulations;
- Constructive and reflective: enabling students to integrate new ideas with prior knowledge to make meaning and enable learning through reflection;
- Intentional: providing opportunities for students to articulate their learning goals and monitor their progress in achieving them;
- Authentic, challenging and real-world (or simulated): facilitating better understanding and transfer of learning to new situations;
- Cooperative, collaborative and conversations: providing students with opportunities to interact with each other to clarify and share ideas, to seek assistance, to negotiate problems and discuss solutions.

Constructive perspective of learning has become the dominant paradigm of learning and its implications for how educators teach and learn to teach are seen as considerable, but necessary if the students within each educator’s charge are to succeed (Jonassen, Hernandez-Serrano & Choi 2000).
4 USE OF WIKIS TO FACILITATE CONSTRUCTIVE LEARNING

Information technologies have been used extensively to enhance learning outcomes and a great deal of research has examined the influence of various technologies on instructional methods and learning outcomes (Alavi et al. 2001, Jonassen et al. 2000). Predominately these studies have focused on the influence of technology features on learning outcomes by comparing various forms of technology mediated learning with traditional learning environments (Alavi et al. 2001) and while these studies have answered questions of ‘does the technology enhance learning?’, they have not readily provided answers on ‘how the technology enhances learning?’ The framework proposed by Alavi and Leidner (2001) (set out below in Figure 1), addressed this issue by shifting the attention to explicit examination of the relationships between technology features, instructional strategy, and psychological processes that impact learning outcomes in a given learning context. As they argued, "it is not the technology features, in and of themselves, that matter. It is the mutual influence of technology features, instructional strategy, and psychological processes that impact learning outcomes.” (p.6)

From a student-focused, constructivist perspective, learning technologies are tools for mediating the practice of learning (Jonassen et al. 2000). Jonassen (1997) emphasized that “technologies should be used to keep students active, constructive, collaborative, intentional, complex, contextual, conversational, and reflective. Current popular online course delivery and management systems, such as WebCT, do not support constructivist perspective of learning (cf: Miers 2004, p.9).

Wiki characteristics and features make it possible to facilitate constructive learning (Bruns et al. 2005). The left hand columns of Table 1 below present the key characteristics of a constructivist learning environment (Miers 2004) as well as learning resources, activities, or supports required to facilitate such a constructivist learning environment (Oliver 2001). And the right hand column of Table 1 provides a list of wiki features found in the literature. We propose a possible link between these two parts. In other words, in order to empirically examine how Wikis can be used to enhance learning outcomes, we attempt to identify various reasons (motivations) for teachers and learners to use various Wiki features in the learning contexts. If motivations for using Wikis are consistent with characteristics of the constructivist perspective of learning and associated activities (left hand columns in Table 1), it will help answer our questions of how and why Wikis can enhance learning outcomes.

5 USES AND GRATIFICATIONS APPROACH

Derived from the mass communication literature, the U&G approach provides a user-centered perspective on the relationship between users and technology. The U&G perspective focuses on explaining the social and psychological motives that shape why people use technologies and that motivate them to select certain technologies in order to gratify a set of psychological needs behind those motives (Katz, Blumler & Gurevitch 1974, Rubin 1994). According to Katz et al (1974), one basic assumption of this approach is that media users are goal-directed in their behavior, and the personal use of media is an active choice made to satisfy needs. The second assumption of this approach is that media users are aware of their needs and select the appropriate media to gratify those needs.

The U&G approach has found to be a useful vehicle to explore people’s motivation for engaging one specific mediated technology over another (Newhagen & Rafaeli 1996, Ruggiero 2000). Technology studies that have taken a U&G approach have focused on a number of technologies, such as television, VCR, telephone, cable TV, and the Internet (Ruggiero 2000). Indeed, the U&G approach has been used to investigate users’ motivations for using a particular mediated technology whenever a new technology becomes available (Elliott & Orosenberg 1987).
Figure 1: A framework for technology-mediated learning research (Source: Alavi et al. 2001)

<table>
<thead>
<tr>
<th>Constructivist Learning Environment Characteristics (Miers 2004)</th>
<th>Learning Resources, Activities, or Supports that Facilitate Constructivist Learning Environment (Oliver 2001)</th>
<th>How does Wiki Technology Support These Resources, Activities, and Supports?</th>
</tr>
</thead>
</table>
| **Active and manipulative** | - Learning objects, animations and simulations that allow students to explore content, test their hypotheses and obtain feedback  
- Websites that facilitate exploratory, non-linear learning  
- Intuitive navigation that between activities and resources that guides and scaffolds but enables learner control | - Wikis versioning capability can show the evolution of thought processes as students interact with the site and its contents (Duffy et al. 2006, Wagner 2004).  
- Wiki pages are highly flexible in structure. Pages can cite other pages based on a powerful back link function (Wagner 2004, Xu 2007).  
- Navigation varies. Users can organize Wiki pages in any predetermined order to enable control (Duffy et al. 2006). |
| **Constructive and reflective** | - Concept mapping software such as inspiration to record prior knowledge  
- Chat and forums to share and discuss prior knowledge and reflect on learning and receive feedback  
- Online journals/weblogs to share learning  
- Note taking form on a webpage that enables notes to be saved, viewed, added to and reflected on form any location | - Once a Wiki has been created it will persist and can be built up as a knowledge repository (Hester 2008, Raman et al. 2005).  
- Wiki is a discussion medium for group work. Students can easily get instant feedbacks from peers or instructors (Wagner 2004).  
- Online journals are provided (Rauschmayer 2008).  
- Wikis are being used as e-portfolios, a tool for collection and reflection at any time, from anywhere (EduCause Learning Initiative 2005) |
| **Intentional** | - Microsoft Word or online rubric generators to construct rubrics to record negotiated learning outcomes and monitor progress  
- Calendar a tool to enable teachers to inset important dates | - The ability of Wiki to export notes to Microsoft Word makes reporting easy (EduCause Learning Initiative 2005)  
- Calendar tools are provided or integrated with external applications which help students to keep track of their own learning process |
and allow students to add their own dates to keep track of their learning
- Concept mapping software such as Inspiration to record learning goals and reflect on them
- Collaborative opportunities to discuss learning goals and obtain constructive feedback from the teacher and peers by means of private email and email listservs
- Online self-assessment surveys or quizzes to evaluate and reflect on learning
- Teacher and peer assessment opportunities through the ability to upload, edit and exchange files as well as make and attach notes

(Richardson 2008, Schwartz, Clark, Cossarin & Rudolph 2004).
- The asynchronous written mode of Wiki web pages enables self-paced study by students (Elgort et al. 2008).
- Wiki as a knowledge repository can support reflection on previous learning (Elgort et al. 2008, Hester 2008).
- Email and group discussion is provided to facilitate efficient communication between instructors and students (Buffa 2006, Schwartz et al. 2004).
- Online self-assessment surveys or quizzes can be set up by instructors. After students’ response, instructors can easily add comment (Liu, Chen & He 2008).
- The functions such as uploading, editing and exchanging files and attaching notes are provided (Challborn & Reimann 2005).

Table 1: Key characteristics of a constructivist learning environment and corresponding Wiki characteristics, and features

| Authentic, challenging and real-world (or simulated) | Problem-solving, task-based activities in which learners are actively engaged in dealing with open-ended questions inquiries and tasks such as: online projects, slam-dunk digital lessons, learning quests, and ask an expert. |
| Cooperative, collaborative, and conversational | Discussion groups to pose and discuss questions, share ideas, seek support and negotiate issues |
| | Forums to pose questions, provide opinions and gain exposure to other ideas |
| | Group collaboration on a task |
| | Chat to discuss and share ideas informally |
| | Online journals/weblogs to record and share thoughts, ideas and learning |
| | Interactive whiteboard for collaborative sharing |
| | Online notice-boards that enable both the teacher and the learner to share ideas |
| | Online polling tools to question, canvass and survey student opinions on a variety of issues |
| | Video-conferencing tools that enables students to see the instructor or other students in a chat session |
| Wiki pages can be linked to external information sources such as web sites relevant to a particular topic/problem (Duffy et al. 2006). |
| Wikis support exploratory learning by providing problem manipulation space (Wagner 2004). |
| Wikis enable collaborative content creation, peer assessment; individual as well as group reflection on learning experiences, user-centric up-to-date information regarding changes in collaborative spaces which all facilitate knowledge sharing and group collaboration (Duffy et al. 2006). |
| Online journals/notice-boards are provided to enable feedback and ideas sharing among peer or between students and instructors (Duffy et al. 2006). |
| Collaborative projects enabled by Wikis can help promote “pride of authorship” and ownership in the team’s activities (EduCause Learning Initiative 2005, Raitman et al. 2005). |
| Online polling tools can be integrated with Wikis to support survey and assessment (Yang, Wu, Koolmanojwong, Brown & Boehm 2008). |
The characteristics of active choice of technologies and user-centered nature make the U&G approach particular useful for understanding motivations to use Internet-based technologies (Kuehn 1994, Morris & Ogan 1996, Ruggiero 2000) and studies that have focused on this area have found a range of motives at work. Garramone and Anderson’s (1986) investigation of electronic bulletin boards in politics found that the needs for surveillance, personal identity and diversion were equally explanatory in that context. Korgaonkar and Wolin (1999) established five motivation factors for the web users: escapism, information control, interactive control, socialization, and economic motivations. Papacharissi and Rubin (2000) also developed a scale of Internet usage motives that consisted of five primary dimensions: interpersonal utility, pass time, information seeking, convenience, and entertainment. Other new gratification dimensions have included: problem solving, persuading others, relationship maintenance, status seeking, and personal insight (Flanagin et al. 2001). Collectively, the U&G perspective has been very useful in understanding motivations and needs for using the Internet.

These studies did however look at the Internet in very general terms, seeing it more as a single technology rather than as the bundle of applications and functions (Parker & Plank 2000). These studies also tended to use previously defined mass media gratifications items rather than developing and using the gratification items uniquely associated with Internet technological context. Understanding the teachers and students’ motivations associated with using Wikis, through the use of the U&G perspective, coupled with the consideration, from a constructive perspective, of the learning context in which this Wikis will be deployed, should lead towards a better appreciation of how Wikis can be best used to enhance learning. The following research questions formally operationalize the research objectives of this study.

RQ1a: What motivates university teachers to use Wikis in their teaching?
RQ1b: What motivates university students to use Wikis in their learning?
RQ2a: Which groups of teachers’ needs do Wikis fulfill best?
RQ2b: Which groups of students’ needs do Wikis fulfill best?

6 RESEARCH DESIGN

Due to the lack of research in identifying users’ motivations for using Wikis in the learning context, this study will adopt Kuehn’s (1994) two-stage research approach for U&G profile development for a comprehensive examination of users’ motivation. Specifically, an initial study will be conducted using interviews to identify the different ‘need statements’ of both university teachers and students. These ‘need statements’ are motives that shape both teachers and students’ Wiki usage patterns in the learning contexts. Then, both teachers and students’ Wiki use motivations identified in the first step will be further examined separately in a large scale survey in order to group these statements into profiles representative of specific motivation dimensions for technologies of each user group.

6.1 Stage 1: Understanding the motivations for using Wikis

In order to identify motivations for teachers to implement Wikis to facilitate their teachings and motivations for students to use Wikis in their learning, 15-20 teachers and students will be interviewed using the Repertory Grid Technique (RGT). These staff and students will be purposefully recruited from the Engineering and Business schools at a large Australian university.

RGT was developed by Kelly (1955) to study personal construct systems and is predicated on the idea that individuals use their own personal constructs to understand and interpret events and that these constructs are influenced by each individual’s background, personal experiences, beliefs and value systems (Napier, Keil & Tan 2007). RGT involves the generation of a list of concepts (elements) about things or events to be studied and the forming of attributes (constructs) based on the list of concepts (Zhang & Chignell 2001). It is a structured interview process with procedures for uncovering the
cognitive constructs of individuals (Tan & Hunter 2002), and has been widely used in organizational and IS research (e.g., Hunter 1997, Napier et al. 2007, Phythian & King 1992).

The output of this stage includes two parts:

- A comprehensive list (elements) of current technologies (including Wikis) utilized by teachers and students in the learning contexts.
- A set of motivation statements (constructs) unique to both teachers and students in the university context for these technologies (including Wikis).

Content analysis will be used to interpret the RGT data, as content analysis allows for the creation of thematic categories from the constructs described in the interviews (Neuendorf 2002). For the categorization of constructs, an adjusted generic core-categorization procedure outlined by Jankowicz (2004) will be used, thus allowing research questions 1a and 1b to be addressed.

6.2 Stage 2: Investigating the relationships between motivations and Wiki uses

Second stage aims to: 1) provide empirical verification and validation of the stage one results; and 2) explore the relationships among learning technologies, teaching strategies, and learners’ psychological learning process, as depicted in Figure 1.

Participants will be asked to complete a questionnaire designed to assess their usage of technologies for satisfying their various needs/motivations (as identified in stage one). Respondents will be asked to rate their level of agreement with the motivations for using each of the technologies identified in stage one in learning contexts on a scale of 1-9 (where 1 = “Strongly Disagree”, 5 = “Neutral”, and 9 = “Strongly Agree”). For each of included technologies, participants will be asked to report their levels of expertise, level of accessibility, frequency of access and weekly usage. Both teachers and students who have experience of using those technologies will be invited to participate.

Research questions 1a&1b will be further validated by using a principal component factor analysis with varimax rotation to extract and interpret potential motivation dimensions (factors) (Papacharissi et al. 2000). Research questions 2a&2b can be then assessed by using MANOVA with the motivation dimensions of teachers and students separately, as the independent variables, and the mean satisfaction ratings as the dependent variables.

7 IMPLICATIONS AND CONCLUSION

A number of limitations are noted for this proposed study. First, this project may be limited to participants who are studying at one university with the majority of them majoring in business and engineering. Students’ majors and university technology use culture may affect their experience with and motivations for using them. Thus, generalizability of the results and conclusions drawn from this study will need to consider the demographics. The U&G perspective has been criticized as being too individualistic by providing little explanation on the formation of social and psychological needs or ignoring the social implications of technology use (Elliott, 1974, Ruggiero 2000, Zhu 2004). To address this issue, a study investigating the psychological and social factors that affect individual’s motivations for using technology and the consequences of technology-related behaviors will be conducted in the future.

The proposed research will be of significance to practice in a number of ways. In general, an improved understanding of the social contexts of technology use in universities, from both a teachers and students’ perspective, should help the implementation of technologies and flexible teaching and learning models into the learning environment. From a teaching perspective, a better understanding of both the motivation for using Wikis and the efficacy of using Wikis should provide a useful prompt for the rethinking of good teaching practices in the pursuit of supporting socially constructed learning practices (Duffy et al. 2006). When educators understand the motivations that guide student
interactions with the technology, they will be able to accommodate those needs more responsively in their teaching strategies.

This proposed research will contribute to the research effort regarding technology use in education in a number of ways. First, it will demonstrate the usefulness of the U&G approach in this area of research, especially in regard to Web 2.0 technologies. It will be amongst the first studies to conceptualize how both teacher-specific and student-specific technology use motivation scales are constructed by employing a multi-method approach. Most importantly, this proposed study will make progress towards understanding why technologies can be used to enhance learning outcomes by examining the links between technology features and instructional variables that might influence learning outcomes.

University students are among the most computer-savvy and “connected” users of the Internet (Aiken, Vanjani, Ray & Martin 2003). Using various technologies has become so pervasive in the lives of this young generation, that it has become a natural extension of themselves (Hoffman, Novak & Alladi 2004). These students will simply expect that new technologies, such as Wikis, will be an integrated part of their learning. It is important then that we understand how to best harness these technologies so that they enhance good teaching practices and allow us to prepare these future engineers, software developers, business leaders and the like, with the critical, creative, collaborative, and communicative capabilities that are required for their professions (Duffy et al. 2006). Understanding how technologies can be used to facilitate and development of those skills is of considerable importance to educators.

References


Oliver, R. (2001). In S. Stoney & J. Burn (Eds.), Developing e-learning environments that support knowledge construction in higher education. We-B Centre (pp. 407-416). Paper presented at the 2nd International We-B Conference, Churchlands: Australia.


E-LEARNING ACCEPTANCE IN WORKPLACE TRAINING: THE CASE OF A GREEK BANK

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0398.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-learning, Case Study, Adoption, Longitudinal data</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
E-LEARNING ACCEPTANCE IN WORKPLACE TRAINING: THE CASE OF A GREEK BANK

Borotis, Spiros, Athens University of Economics and Business, 76 Patission Str., 10434 Athens, Greece, borotis@aueb.gr

Poulymenakou, Angeliki, Athens University of Economics and Business, 76 Patission Str., 10434 Athens, Greece, akp@aueb.gr

Abstract

This study presents particular issues concerning the acceptance of e-Learning in a major Greek bank. A generic unified analysis framework is developed based on dominating theories of individual information systems acceptance, integrating also the organizational and work-context dimensions. Concerning the organizational level, management of employees’ development, training culture, and alignment with corporate strategy raise particular significance. On the work-social context dimension, Training Department, branch managers and corresponding supervisors, as well as colleagues, consider the main key stakeholders. Additionally, time resources and work overload appear as critical when designing this type of training interventions. Last, in the individual level, perceived usefulness and ease of use, personal innovativeness with information technology, computer anxiety, self-efficacy, and intrinsic motivation to learn affect in broad terms the attitudes employees formulate towards e-Learning in this organization. All these issues, in the various aforementioned levels, must be taken into account when designing and maintaining technology-supported training interventions, in order to mitigate resistance and maximize the potential benefits.

Keywords: e-Learning, information systems acceptance, case study, longitudinal data.
1 INTRODUCTION

The advent of the knowledge economy and related technologies has created a wave of IT interventions in the area of human resource development (HRD). In this novel information systems adoption context, critical adoption drivers shift from the pursuit of operational and/or managerial efficiency and effectiveness to the challenging areas of human capital development and employee motivation and engagement. As with other IT systems, the introduction and implementation of HRD systems into contemporary organizations agendas necessitates the recognition, acceptance and support from various stakeholders in order for them to provide the full benefits and functionalities they carry.

As Gartner (2008) says¹, in the reality of commoditized banking products and increasing global regulation, IT professionals and business leaders in the banking industry must invest in the right technologies to support growth, preserve margins and aid compliance of human capital with corporate objectives. Human resource systems contribute significantly in all these priorities, as they sustain the human capital improvement and effective utilization of personnel, reduce costs, and - especially the training programs - consider a valuable aid towards achieving competitive advantage.

Workplace training concerns a systematic approach towards learning and human resource development, in order to improve individual, team, or organizational effectiveness (Goldstein and Ford, 2001). Nowadays, this activity is supported by the various types of “technology-enabled learning” interventions, one of which is e-Learning. This training delivery mode exploits the power of web networking and capitalizes on corporate technology infrastructures in order to deliver instruction and achieve particular objectives (Rosenberg, 2000).

This study unfolded from a research project which intended to the total improvement of the development of employees working in a large Greek Bank, as well as to surface particular deficiencies of current training operations. The bank is a part of a large financial group, operates in the Greek market and has more than 650 branches worldwide, employing more than 7000 people. Its strategy focuses on becoming the leading bank in the wider region of South Eastern (SE) Europe, a rapidly developing and profitable market of 60 million people, with profound implications to human capital development. Generally speaking, the small size of the Greek market, the maturity levels it has approached, the fierce competition, mergers, acquisitions, and the last global economic deficiencies, have motivated several banks to expand to the SE Europe and Turkey. Those markets provide them particular opportunities for growth, for differentiating their sources of incomes and portfolios, and for increasing their size, which is relatively small to their competitors operating in the EU. All these conditions, as well as new products, services, business process reengineering projects towards establishing customer-oriented philosophies, and the compliance with National legislation and European Central Bank, create particular training needs for banks’ geographically dispersed employees.

Although the primary objective of the project was to produce a training needs assessment methodology in the context of technology-supported learning, soon the interest of both researchers and the Bank was shifted to the wide human capital development practices generally, and to e-Learning specifically. To this end, the specific project protocol and the particular needs of the Bank, specified the goal of identifying the particular factors that affect e-Learning operation and adoption. To this end, a double research agenda unfolded. The first concerned the organizational acceptance of e-Learning, which had to do with the management of human capital development. The second concerned the acceptance of e-Learning from employees, who operated the same time as users of e-Learning courses and learners. The planning and implementation of the project spanned more than four years, from March 2003 to July 2007. However, at several junctures during the implementation, the evolution of the project was hindered by various internal management changes.

¹ http://www.gartner.com/it/products/research/industries/industries.jsp#bank
The starting point of the changes that raised the interest of the research team was a huge Business Process Reengineering project, which was implemented in 2001. It generally included the business transformation from a process-oriented to a customer-oriented approach. Just after the reengineering, an innovation occurred to the training function of the bank. E-Learning was established as a mode of training in the form of blended learning (Trasler, 2002), covering the gaps of traditional, on-the-job, and self-training. A well known Learning Management System used to deliver the asynchronous self-based e-Learning courses was acquired. Off-the-self courses as well as other outsourced and custom developed for the specific needs of the bank were also acquired. As e-Learning adoption requires a lot of resources, willing, and careful design, various worries raised to the researchers forming the initial research questions, at least in broad terms (Mintzberg, 1979).

- What issues were affecting the acceptance of the initiative both from the organization and the individuals?
- Were there any particular context-specific characteristics that were facilitating or inhibiting the initiative?

These queries formed a priori the units of analysis, which were the organization and the employees, and the method of work, the case-based research approach. In the following section we specify potential important variables based on extant literature and the experience of the research team.

2 THEORETICAL FOUNDATIONS

2.1 Information systems acceptance

Information systems are introduced to an organization for various reasons: automate procedures, improve efficiency and effectiveness, and increase workforce performance. Some systems are adopted in order to establish a particular way of doing business, through making their use mandatory and inevitable. Other times, using the system is not obligatory and depends on the voluntary use of employees. To this end, people tend to use an application or not based on their beliefs whether it will enhance their job performance (Davis, 1989). Between these two border lines, there is a continuum characterized primarily by the level and the degree of actual effective and efficient use of the system. Thus, the case observed in practice is that, the more particular systems are used, the greater impact they can have (Trice and Treacy, 1988). As these systems usually demand significant resources to be implemented and maintained, organization have to employ aggressive tactics to encourage themselves and users to expand and take advantage of the installed technological systems (Jasperson et al., 2005).

Upon organization’s decision to adopt a new technology, individuals’ perceptions and attitudes affect the success of the initiative (Leonard-Burton and Deschamps, 1988). Generally speaking, people are sometimes unwilling to accept and use systems, even if they would improve their job performance or relieve them from demanding and time-consuming tasks (Mathieson, 1991). Literature indicates that people tend to resist to MIS because of their own internal factors (people-determined), because of poor system design (system-determined), and due to the interaction between specific system design characteristics and the related (organizational) context of use (Marcus, 1983). Focusing on the first of the aforementioned dimensions, as systems become useless when people do not use them “the right way”, i.e. effectively an efficiently, it worth to examine how they decide whether they will use them or not (“the right way”). This is examined in the literature of information systems under the umbrella of Information Systems Acceptance. Moreover, the organizational context of use, motivate us to examine the particular work-related characteristics that may affect acceptance and use.

In this aim, Information Systems acceptance research has developed, empirically tested and shaped many alternative models to the research community and practitioners, focusing on the user/individual level. The main focus of various theories and models appear in the area consider the intention or actual usage as the dependent variable. The Theory of Reasoned Action (Fishbein and Ajzen, 1975), the
Theory of Planned Behaviour (Ajzen, 1991), and the Technology Acceptance Model (Davis, 1989) dominate in this area.

According to the Theory of Reasoned Action (TRA), Behavioural Intention is the primary determinant of an individual’s behaviour. Moreover, his intention to adopt an innovation is influenced by a personal factor (Attitude) and a social influence factor (Subjective Norm). Attitude is an individual’s positive or negative belief about performing a specific behaviour. Subjective Norm is the perceived influence of social pressure on an individual to perform or not perform the behaviour. The TRA has been used in practice (e.g. Becker and Gibson, 1998; Chen and Chen, 2006), but not appropriately some times (Sheppard et al., 1988).

The Theory of Planned Behavior (TPB) (Ajzen, 1991) is an extension of the TRA, which was developed in order to overcome the limitation of dealing with behaviours over which people have incomplete volitional control, introducing the concept of perceived behavioural control. As many behaviors present difficulties in execution that may limit volitional control, perceived behavioral control predicts also directly the behaviour, serving as a “proxy” for actual control (Armitage and Conner, 2001). The concept is is strongly interrelated with the concept of self-efficacy (Bandura, 1977), which seems to be a crucial parameter of motivation (Hodges, 2004). The aforementioned ideas are supported adequately by various studies in various domains (Courneya et al., 1999; Mathieson, 1991; Trafimow et al., 2002; Sutton, 1998). Alongside, various implementations and alterations of the TPB have been proposed in the literature (e.g. Taylor and Todd, 1995; Crespo and del Bosque, 2008; Pavliou and Fygenson, 2006).

TAM originates from the two aforementioned theories, including only a proportion of them. Since its invention, TAM has been used extensively, in various contexts, and has been enhanced with various other factors (e.g. Davis and Wiedenbeck, 2001; Legris et al., 2003; Lu et al., 2008; Ngai et al., 2005; Van Raaij and Schepers, 2008). Additional variation of the aforementioned model include the TAM2 (Venkatesh and Davis, 2000), the Unified Theory of Acceptance and Use of Technology (Venkatesh et al., 2003), and combinations of the aforementioned models (Taylor and Todd, 1995; Yi et al, 2006).

2.2 Scrutinizing human behaviour

The common and main as well concept of the aforementioned theories is attitudes, which attract continue research in various disciplines (Ajzen, 2001). Attitudes reflect the individual’s positive or negative evaluations of performing a particular behavior, and their formation is based on the Expectancy-Value Model (Fishbein and Ajzen, 1975). In the context of information systems, literature has indicated a number of beliefs and associated outcome expectations, like the perceived ease of use and perceived usefulness (Davis, 1989), the personal innovativeness with IT (Agarwal, and Prasad, 1998; Lian and Lin, 2008), trust (Pavlou and Fygenson, 2006), etc. Subjective norm refers to the individual’s perceptions of general social pressure to perform (or not) the particular behaviour. The claim here is that, if the individual perceives that significant “referent” others – whose beliefs may be important to the individual – approve or disapprove the behavior, it is more (or less) likely to perform it. “Referent others” are people.

Additionally, other issues also have been proposed as facilitating or inhibiting factors in the creation of intentions. Computer anxiety, intrinsic motivation, gender, and personal IT innovativeness to learn can be categorized as individual factors that affect intentions to use information technologies through attitudes and perceived behavioural control (Harris, 1999; Taylor, 2007; Thatcher and Perrewé, 2007; Thatcher et al., 2007; Venkatesh and Morris, 2000; Ong and Lai, 2006; Tharenou, 2001; Davis et al., 1992). Additionally, work specific factors as work overload (Ahuja et al., 2007; Moore, 2000), and time resources (Pavlou and Fygenson, 2006) have been also proposed in the literature as issues affecting indirectly the intentions.
2.3 e-Learning

E-Learning is typically employed through online learning courses, in synchronous (e.g. virtual classrooms) and asynchronous formats, and in the context of Web 2.0 through informal learning tools, like wikies and blogs. This technology can be used stand-alone or complementary to other training modes, in order to mitigate the skill shortage that several banks are recently facing as a result of particular organizational transformations, e.g. from bureaucratic to customer-oriented (Neirotti and Paolucci, 2005; Regini et al., 1999).

Study of various cases demonstrates that there are particular characteristics of e-Learning used in corporate training. First, organizations seldom use “pure e-Learning”, but they prefer the so called “blended learning” which consists a new and promising learning approach (Williams, 2003), as well as the dominant way of using e-Learning effectively and efficiently (Kim and Choi, 2004; Finn and Bucceri, 2006), maximizing also the ROI in HRD interventions (Oakes and Green, 2003). Blended learning is the effective integration of various learning techniques, technologies, and delivery modalities to meet specific communication, knowledge sharing, and informational needs (Finn and Bucceri, 2006). The most comprehensive approach of the different blended learning dimensions has been proposed by Patti Shank (2004) and many studies indicate the right mix as one of the critical success factors (Trasler, 2002; Williams, 2003; Oakes and Green, 2003).

3 RESEARCH METHODOLOGY

3.1 Research method

Given our interest in the continuous alterations as well as the hazy process initiation and management concerning the training and e-Learning function in the Bank, an exploratory, qualitative, longitudinal case-based approach with participant observation was selected and followed (Yin, 1994). Typically, a case study investigates a phenomenon in its natural setting, using multiple methods of data collection in order to gather information from various units (people, teams, organizations, etc.). Case studies also allow the researcher to answer “how” and “why” questions, in order to understand and decode the complexity of the processes taking place. These critical questions appear both in interpretive and positive studies (Walsham, 1995). Last, it appears to be quite useful when there are few studies on the issue in the particular context, as happens in this case (Benbasat et al., 1987). The case study was conducted in order to understand the dynamics presented in the case organization (Eisenhardt, 1989), primarily by the two authors, and supplementary by five other persons, with various backgrounds, forming an insider-outsider research team as recommended by Eisenhardt (1989).

The primary sources of information for the case study were stakeholders’ semi-structured interviews with the upper and line management of the human-resource related departments, the operation team of e-Learning, tutors, and employees working both in branches and corporate headquarters. Moreover, internal documents were reviewed (e.g. strategy and process-related documents), reports, and meeting minutes. The interviews were conducted from May 2003 until March 2007. Twenty two interviewees sat for a total of fifteen interviews, with the time frames varying between 1.5 to 2 hours. These semi-structured interviews were based on predefined research agendas already delivered to interviewees before the meetings, and reflected the key issues of research. Depending on the issues, the interviewees varied between the upper management and staff of the Training sub-department and related corporate functions, to employees working in the branches. Moreover, the same day after each interview, the research team was conducting an internal meeting to clear and agree on the new information acquired and lessons learned. The interview data were used to uncover the particular hidden aspects of the interviewees’ words, create a common understanding between the research team, and to identify critical issues for research. Frequent discussions between the researchers were also taking place in order to reflect on the newly acquired knowledge, especially after studying related
scientific material on the topics analyzed in section 2. The meeting minutes and internal documents were additionally used to understand the issues identified during the case analysis. Moreover, the meeting minutes were also reviewed by the interviewees in order to reduce bias.

3.2 Data collection and analysis protocols

In this study, a multiple levels of analysis perspective were adopted (Yin, 1994; Lee, 1989). First the authors aimed at gaining deeper understanding about the actions of the main stakeholders concerning the management and operation in human capital development in the Bank, as their experiences were important and critical in the particular context of action (Benbasat et al., 1987). Moreover, as e-Learning consists the most promising system in corporate training, the main characteristics, as well as the way it was introduced and used (“pushing” vs. “pulling” activities) was reviewed. To this end, a second level of analysis was created, focusing on the perceptions of individual individuals towards the technology intervention in their learning. In other words, the authors aspired to the double nature of employees, i.e. users of IT and learners, in order to develop meaning on this innovative behaviour.

In this research an inner context and process driven approach through time was pursued, consistent partially with the work of Pettigrew (1990) on contextualism. Outer context was willingly excluded from this study as training the employees concerns an internal topic of each organization, although affected by the economic, social and sectoral changes. Moreover, particular emphasis was given to provide meaning on the various emerging issues. To this end, the study emphasized towards the importance of embedding e-Learning in corporate training agenda, and studying the change occurred and the various reactions of different stakeholders. Moreover, it located changes in past, present and future time (occurred and prospective changes), and explored the relation between context and action.

Due to particular project constraints, a single site analysis was ensued. Although contextualism research sustains the comparative method, this approach (i.e. single site) has also been widely applied in practice in order to develop deep insights (Boudreau et al., 2001; Orlikowski, 1996). Alongside, a triangulation methodology was applied in practice, through the in-depth interviews with various stakeholders, document reviews and the multiple researchers’ perspective.

4 FINDINGS AND DISCUSSION

The issues analyzed for the need of the research were developed around the questions established before. In order to mitigate complexity, a multiple levels of analysis approach was selected for studying the inner context (Pettigrew et al., 2001). These include the managerial context, the work-social network and particular work-setting characteristics, and the individuals. For presentation reasons, the analysis framework is presented in this point, and is used for case analysis and discussion in the following sections.

![Figure 1. The proposed framework of analysis](image-url)
4.1 The Human Resource Development function

Human resource management and development in the Bank was the first complicated issue identified. A legacy Personnel department was responsible for all traditional HR-related activities apart those included in the developmental agenda, e.g. training and development, organizational development, human resource planning, and career planning (Wilson, 1999). This department was supervising two separate sub-departments, the Human Resource Development and the Training department. Whereas the former was conducting all the aforementioned issues except training, as well as some training needs analyses, the latter was responsible for developing, managing and delivering training and e-Learning to the employees, with the main objective of serving efficiently the branch network training demands. This corporate function appeared to be the most promising and active in relation to the others, as it was also conducting training needs assessment and proactive training. Actually, it was aspiring at undertaking all employees’ developmental activities, formulating a new corporate function with the HRD department. The aggressive nature of tactics was sustained by its brilliant leader, who was talking in passion about his (team, not personal) objectives,

“We operate under fierce competition ... the branch employees need up-to-date information and training in order to enhance customer satisfaction. On the other hand, the HRD department is small and slow, and usually we have to search in employees records for the eligible ones to participate in training. Moreover, we continually watch the competition, we try to learn from the achievements and failures of others, and formulate accordingly our unit’s strategy. Moreover, we often discuss with other unit’s managers in order to learn in advance about the new products and services, or even organization-wide initiatives, and develop proactively the supplementary training interventions. Last, we conduct official surveys and unofficial discussions with branches’ managers and employees, in order to improve our services, and convince them about the significance and benefits of training generally, and e-Learning in particular.”

Obviously, the overlapping and the same time partially dispersed human resource-related functions, operated in a context of internal competition. On the one hand we meet a creative Training department, which introduces and use modern technologies to accomplish its tasks. On the other hand we see a legacy Personnel department, which does not even have in place a modern human resource information system to manage employees (Ceriello and Freeman, 1998). Actually, the information applications used by the latter were mainly spreadsheets and legacy applications formulating islands of automation (Applegate et al., 1999).

Apart from that, the Training department was continuously trying to establish its strategic position in the HR agenda to the upper management committee of the organization, with which it cooperated frequently, and sustained its short (one year) and long-term (three years) strategic plan. The departmental strategic plan was tied up to the corporate business plan.

Examining the case, particular grains of training culture surfaced. Apart from the technological innovations, the Bank was undertaking particular activities that contributed to the total improvement of its personnel, e.g. hiring knowledgeable employees, implementing job rotations, and undertaking actions towards career development (Gilbreath, 2008; Preskill and Donaldson, 2008). Additionally, all the subject-matter experts were educated in international business schools and accredited from international organizations. These people operated also as an unofficial knowledge management system, as they contributed both to the corporate strategy planning as well as the facilitation of daily work duties of employees through answering their questions.

The aforementioned activities were conducted in order to create and sustain a corporate culture towards the uptake of the human capital as well as the establishment of a concrete and knowledgeable workforce. Apparently, and as the Training Manager underlined,

“Our objective is to transform the Bank to a learning organization (Senge, 2006). Apart from developing the line staff, we have also to train the branch managers, as their role in people
management is crucial for daily operations. Unfortunately, they usually don’t know what their subordinates know. To this end, we try to convince them to participate in training, even join a university to acquire a masters degree or attend an HR conference!”

Apparently, the branch managers and directors, as well as the subject matter experts, consisted the main stakeholders in the motivation of employees to attend e-Learning. The former were forming a critical role in the work-social network of employees, approving their request for participation in training and e-Learning, and motivating them to join it. On the other hand, subject matter experts were responsible both for training content, and the traditional training in the classroom, which complemented the e-Learning modules.

4.2 Social network and other work settings issues

Despite the recognized need for training, particular issues inhibited the effectiveness, efficiency, and finally active and concentrated participation of employees in training. The first related issue appeared back to 2003 and concerned the desire of the Training Department to foster employees to attend e-Learning outside the work context (e.g. from home). Before that, the bank had created the so-called “e-Learning kiosks” inside many branches, where each employee could go to attend e-Learning during the working hours. Soon the idea was abandoned. As a teller mentioned, “I did it some times when there was not much workload and after the agreement of the branch manager but, when customers came, they were not served adequately fast and started screaming “why don’t you call an additional employee to serve us”? I had to stop immediately”. 

Although attending e-Learning from home seemed to be a very good idea for the upper management, employees expressed different opinions. As a female branch network employee stated, “We are in a panic all day. Who has time to attend e-Learning during working hours? They offer access from home also but, I am a mother and although I’m tired when I get back home, I must help my children to study and take care of the household. Sorry, no time for e-Learning”. 

Obviously, both persons raised the issue of the limited time resources for attending e-Learning during working hours (Pavlou and Fygenson, 2006). Moreover, the lady focused on the issue of creating a work-family conflict, a characteristic of inadequate psychosocial healthy work environments (Gilbreath, 2008) and critical antecedent of work exhaustion (Ahuja et al., 2007). 

To this end, a survey was conducted from the Training department in order to motivate employees to attend e-Learning outside the work context. The study was conducted in 2003 with the use of a questionnaire filled in by 479 employees (58% response rate). The participants were 60% female, another 60% were working at the bank more than 10 years, and were employed in all levels and roles. 

Despite the previous negative concerns, employees indicated their positive attitudes towards e-Learning (> 80%), the moderate perceived usefulness (55%), the need to attend e-Learning outside the work context (70%) due to heavy work load and limited time resources, and need to avoid interference of work to their personal lives (27%). Actually, the 75% denoted the family affairs and the 63% the fatigue due to heavy work load as the main inhibiting factors for attending e-Learning from home. Additionally, there were indications of lack of training culture in the branches, an issue that was defined through the reduced participation in e-Learning in branches, the moderate attitudes towards technology, and the lack of connection between learning and career development. Interestingly, employees had already participated in e-Learning were denoting their intention to re-use it. Additionally, only 22% of trainees declared that they preferred the classroom-based training, in order to maintain physical contact with other trainees and the tutors. The main outputs of this study implemented in practice through marketing activities towards e-Learning benefits.

Another main stakeholder of the work-social network of the employees were their supervisors and branch managers. These people were challenged daily to serve customers and achieve corporate objectives utilizing fully the capabilities of the staff they managed. Although most of them presented
significant experience, they usually did not have to attend the courses that their subordinates had. They didn’t even know the particular knowledge of each employee! Moreover, as some clerks mentioned, they didn’t motivate them to attend e-Learning.

4.3 Employees and e-Learning: the sovereign in a contest

The results of e-Learning acceptance from employees appeared to be quite interesting. Reactions of employees towards e-Learning courses were quite positive, requests for access were continually increasing and direct and indirect training costs were reduced. Moreover, the absence of employees from the position was decreased too due to e-Learning. Alongside, an internal survey conducted in the end of 2003 revealed the significant usability of existing e-Learning courses (Nielsen, 2000). Actually, data indicated the increased content completeness (72%), visual design (93%), interactivity (62%), navigation (82%), moderate to easy access (68%), and moderate control (58%). Main findings also included the decreased motivation for searching for additional material (30%), moderate lack of additional learning resources (49%), and limited capabilities for collaboration with other learners (30%). Totally, the 65% of the trainees appeared to be satisfied with the existing e-Learning courses, an issues that raised inquires about the remaining 35%. As the design of e-Learning courses and the content were satisfactory, what issues could hinder the total acceptance? To this end, next we focus to – negative or positive – issues presented by the interviewees that may explain this discrepancy.

Literature indicates the need for IT competencies for bank employees, as well as managerial practices that facilitate the adaptation of human capital to technological change (Neirotti and Paolucci, 2005). In this regard, we assume that e-Learning necessitate the aforementioned skills also. Thus, acceptance of e-Learning from employees should be affected by the knowledge of related skills. Moreover, as e-Learning is considered as an innovation in corporate training agenda – due to its technology dimension – people should present adequate innovativeness in using this IT. To this end, the issue of Personal Innovativeness with Information Technology emerged (Agarwal and Prasad, 1998). This claim is also documented by the words of an employee,

“I like e-Learning. I can learn on my own time, even from the office – although a little difficult some times – or usually from home during the weekends. Generally speaking, I like very much the computers and the Internet, my friends call me “technology freak”!”

Apart from the dimension of technology, this employee place the issue of controllability, setting the route towards the construct of perceived behavioural control in the Theory of Planned Behaviour.

On the other hand, literature indicates the anxiety that various people feel when they use computers (Thatcher and Perrewé, 2007). As an employee stated,

“I have to use computers for my work, I can say that I feel comfortable, although a little strange some times, as I learned to operate computers in the bank. But, I don’t use them extensively at home”

Last, various studies recognize the importance of intrinsic motivation to learn in participating in training sessions or using related systems (Davis et al., 1992). This issue appears also in this case,

“I like to learn new things. I read books, and I want to do a part-time masters! I know to handle computers and navigate to the internet, and I use it daily. I have used e-Learning to learn, and I think it is easy for me to do that!”

To this end, we see that the issues of perceived ease of use, perceived usefulness, personal innovativeness with information technology, computer anxiety, and motivation to learn have been mentioned as critical from the employees in the bank. To this end, these issues may formulate particular individual characteristics that may affect attitudes and control of the behaviour of employees in the formation of e-Learning acceptance.
5 CONCLUSIONS AND LIMITATIONS

Although the comparative and a multiple-site approach is encouraged by Pettigrew (1990) and is applied in other studies (e.g. Edmonson et al., 2001), this study was conducted in one organization an issue that implies particular limitations for the generalizability of results (Lee, 1989). Though, this research site was fulfilling the demands of the particular research topics and the questions being posed. Moreover, as it was presented in the previous section, it included particular critical incidents and dramatic glimpses into the current human resource development processes, which may act and support adequately the meaning elicitation through an in-depth case study. Additionally, the selected organization employs high experience levels of the phenomenon under study as banks constitute knowledge-intensive organizations, and the particular bank constituted a pioneer in e-Learning in the Balkans (i.e. extreme case (Eisenhardt, 1989)).

A limitation of various studies in the area is that, they are conducted in academic settings, where students are the participants (Venkatesh et al., 2003). Only a few studies are based on data acquired from employees (Plouffe et al., 2001), and to this end, the presented study differentiates from other works.

This study reviewed the organizational, work-related and individual characteristics surrounding the training generally and e-Learning specifically in a Greek bank. Initially two research questions were deployed; the first was focusing on the issues affecting the acceptance of e-Learning as a corporate training means both from organizations and employees. The second question imposed concerned the identification of context factors that were facilitating or inhibiting acceptance. Data indicated interesting finding in the aforementioned issues, and were presented with the formulation and use of a particular framework.

Generally speaking, the e-Learning could be characterized as effective and well accepted in this organization. Data proved partially this claim, but pinpointed also particular deficiencies that hinder its total acceptance. The main findings and strenghts of this research – in-progress – are as follows; first, the analytical proposed framework integrates both the organizational, work-social context and individual aspects, formulating a unified approach for studying information systems acceptance generally, and e-Learning in particular. The common frameworks used by now cover only the last dimension. Second, the key organizational characteristics identified in the organizational level – including the management functions which were operating in a complicated way, the particular training culture characteristics, and the alignment of corporate and training strategy – consist an issue that facilitate the initial adoption and future expansion of e-Learning efforts of an organization, leading to the issue of e-Learning readiness (Borotis and Poulymenakou, 2004). Third, work related issues arised, including time resources and work overload, as well as the organization of social context which could affect the acceptance and operation of e-Learning contribute to the formulation of contemporary psychological healthy working environments with respect to the balance between working and personal life (Gilbreath, 2008). Main stakeholders in this context that must be taken into account when designing these environments include the training department, the branch managers and corresponding supervisors, and the pure employees. Last, the particular individual characteristics identified and could affect attitudes formation included the perceived ease of use and usefulness of the e-Learning programs, personal innovativeness with IT, computer anxiety, intrinsic motivation, and self-efficacy of learning through e-Learning.

The aforementioned issues, as presented in Figure 1, may formulate an extended version of the Theory of Planned Behaviour, which takes into account the organizational level of acceptance. Future work should test with quantitative data the aforementioned proposed model, in various organizations coming from different industry sectors. The authors believe that the particular work-context issues emerged in this case, are common in various others work domains, due to the fundamental characteristics of contemporary organizations.
References


IT IMPACTS ON OPERATION-LEVEL AGILITY IN SERVICE INDUSTRIES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0192.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business value of IT, Dynamic Capabilities / relationships / perspective, IT capability, Services operation and management</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
IT IMPACTS ON OPERATION-LEVEL AGILITY IN SERVICE INDUSTRIES

Lee, One-Ki (Daniel), University of Massachusetts Boston, College of Management, 100 Morrissey Blvd., Boston, MA, USA, Daniel.Lee@umb.edu
Xu, Peng, University of Massachusetts Boston, College of Management, 100 Morrissey Blvd., Boston, MA, USA, Peng.Xu@umb.edu
Kuilboer, Jean-Pierre, University of Massachusetts Boston, College of Management, 100 Morrissey Blvd., Boston, MA, USA, Jeanpierre.Kuilboer@umb.edu
Ashrafi, Noushin, University of Massachusetts Boston, College of Management, 100 Morrissey Blvd., Boston, MA, USA, Noushin.Ashrafi@umb.edu

Abstract

This study examines the relationship between operation-level agility and firm performance in service industries. The study is augmented by investigating the role of IT resources and competence to achieve this specific type of agility. As of to date, most of the published literature in this stream of research has focused on manufacturing industry. This research is an early attempt to examine the strategic value of IT-enabled operational capabilities in service industry. We propose a theory-based model of the positive relationships among IT service competence, operation-level agility, and firm performance. Survey data of medium to large-size enterprises in service industries in the United States were used to validate the proposed model. The results indicate that operation-level agility is a significant driving force of firm performance in the service industries and that IT service competence significantly determines the operation-level agility. The results emphasize that IT-supported operation-level agility significantly leads to a better performance.

Keywords: Operation-Level Agility, IT Competence, IT Resources, Firm Performance
1 INTRODUCTION

Globalization and the advancement of information technology (IT) have produced an intensely competitive, dynamic, and unstable business environment. This competitive landscape is often shaped by escalating competition and strategic maneuvering based on price-quality positioning, attempts to establish market advantage, and pressure of new knowledge creation (Hitt & Keats & DeMarie 1998, McNamara & Vaaler & Devers 2003, Shi & Kunnathur & Ragu-Nathan 2005). For example, industries such as electronic goods and finance face challenges by short lifecycle of products and rapid technology diffusion.

Agility, the ability to rapidly respond to external changes, has been proposed as a solution to such changing environment (D'Aveni 1994, Sambamurthy & Bharadwaj & Grover 2003). Agility has been examined at various levels in the literature, such as enterprise-level agility, business function-level agility, project-level agility, and system-level agility (Lee & Banerjee & Lim & Kumar & van Hillegersberg & Wei 2006, Overby & Bharadwaj & Sambamurthy 2006, Prewitt 2004, Sambamurthy et al. 2003). In this study, we focus on agility at the operational level, the ability of a firm to achieve speed, accuracy, and cost economy to realize emerging opportunities for operational innovations.

During the last two decades, operation-level agility has received considerable attention in the manufacturing industry (Narasimhan & Swink & Kim 2006). In manufacturing settings, agility has been emphasized in terms of leanness and flexibility of operational processes, such as resource procurement, manufacturing, quality control, and product delivery (Burgess 1994, Ettlie 1998, Gerwin 1993, Llorénsa & Molinaa & Verdúb 2005). This capability has been discussed as a driving force for firms' exploitation of changing market opportunities, thus leading to competitive market position. The capability of people to exploit their knowledge to cope with market demands, internal process changes, and interactions with suppliers has been addressed as another factor impacting business process agility (Grover & Cheon & Teng 1996, Hamel 1994, Pavlou & El Sawy 2006). Research on agility in service industries, however, remains sparse while there has been a pressing need to cope with growing and evolving service sector of the economy (Roth & Menor 2003).

The service industry is different from manufacturing in many ways. As pointed out by Roth and Menor (2003), business processes in manufacturing firms affect consumers mainly through their products which are usually physical goods. Hence, in the manufacturing settings, operation-level agility is usually restrained by physical constraints such as locations, resource availability, and delivery time. In contrast, the service offerings and delivery involve processes enhanced by support amenities, facilitating information, and implicit services, e.g., psychological benefits (Menor & Roth & Mason 2001). Moreover, while the direct interaction with customers in the service settings makes it easier to collect market intelligence, the customers' demands are more complex and likely to change due to the dynamics of direct interaction with individual customers (Roth & Menor 2003). Due to the rapidity of change in competition, market dynamics, and customer preferences, the breadth and pattern of responses required in the service settings are much broader, more frequent, and sometimes more unpredictable (Beidleman & Ray 1998, Menor et al. 2001). To the best of our knowledge, little has been done to understand operation-level agility in the service industry. Menor et al. (2001) investigated agility in banks. However, by treating agility as a one-facet concept in their study, they could not fully explain the role of operation-level agility in competitive performance of service firms. Due to the lack of research on operation-level agility in the service industries, the role of IT as an enabler for this significant business capability is also unclear.

In this research we address these knowledge gaps by answering the following research questions:

- What is agility at the operational level in the service industries?
- What is the organizational influence of operation-level agility?
- As essential assets in current business, what are the IT competences that support operation-level agility?
This paper is organized as follows: Section 2 defines the key constructs and develops the hypotheses and research model; Section 3 discusses research methodology followed by the discussion on data analysis. The paper is concluded with implications and discussions.

2 RESEARCH MODEL AND HYPOTHESES

Figure 1 shows the research model of the study.

![Research Model](image_url)

2.1 Operation-level agility and firm performance

The notion of agility has been discussed extensively in the literature. There is a general agreement that when facing a turbulent environment, enterprises must adapt to changes. Otherwise, they lose their competitive advantages. In service industries where competition is severe, a key business competence is to acquire market information and respond to changes in an effective and timely manner (Overby et al. 2006). Prior research has proposed agility at different organizational levels. For example, Sambamurthy et al. (2003) describe enterprise agility as one of the important dynamic capabilities in a turbulent environment. In their study, agility is defined as “the ability to detect opportunities for innovation and seize those competitive market opportunities by assembling requisite assets, knowledge, and relationships with speed and surprise” (p. 245). According to this conceptualization, agility encompasses the capabilities related to interactions with customers, deployment of internal operations, and utilization of its external business partners. Similarly, Overby et al. (2006) define agility as “the ability of firms to sense environmental change and respond readily” (p. 121). This definition emphasizes two capabilities: sensing and responding. Hence, agility includes the ability to detect, anticipate, and sense market opportunities, evolving conditions, and other environmental changes. At the same time, it also includes the ability to seize the opportunity with speed and implement new solutions. Similar to the study done by Sambamurthy et al. (2003), Overby et al. (2006) argue that agility applies to both strategic and operational levels within a firm.

In this study, we focus on agility at an operational level, namely operation-level agility, which is defined as the ability of a firm to achieve speed, accuracy, and cost economy to realize emerging opportunities for operational innovations (Cao & Dowlatshahi 2005, Sambamurthy et al. 2003). It emphasizes the effectiveness and efficiency of a firm’s actions in response to changes in their daily operations. For the purpose of this study, we define operation-level agility as a composite of three interrelated capabilities: market responding capability, process reconfigurability, and supplier management capability. We argue that these capabilities combined would enable a firm to seize opportunities, respond to internal and external changes, and sustain its competitive operational edge.

Market responding capability refers to the ability of firms to sense emerging opportunities and threats and respond to the market to deliver products and services valued by their customers (Grewal &
Operation-level agility is the response to changes in the market. It requires firms to be aware of changes and then seize the opportunities by tailoring their business solutions. Hamel (1994) has argued the market-related capability is one of the core competences in the business. It allows a firm to stay close to its customers, identify emerging needs quickly, and deliver new services in a timely manner (Hamel 1994, Ravichandran & Lertwongsatien 2005, Wang et al. 2004). As an organizational capability, market responding capability integrates multiple functional capabilities such as intelligence collecting, market research, and customer relationship management to tailor services for new tastes of customers. This capability is necessary to help firms sense and realize emerging opportunities for operational innovations.

Process reconfigurability refers to the ability of firms to transform and reconfigure their resources and processes in order to accommodate changes (Pavlou & El Sawy 2006). Transformation and reconfiguration of resources are catalysts for change. While market responding capability helps the firm identify the necessity for change, process reconfigurability allows it to deploy new configurations of functional competences that better match the environment (Pavlou & El Sawy 2006) and ensures that the firm can rapidly redesign and modify existing processes for new market conditions (Sambamurthy et al. 2003). Firms’ ability to integrate and combine existing resources into “novel” combinations to better match their product-market areas helps them respond to changes and deliver new services effectively.

Supplier management capability refers to the ability of a firm to modify or adapt its extended enterprise network (e.g., supply chain) when it needs access to assets, competences, or knowledge not currently resident in its business itself (Dyer & Singh 1998, Sambamurthy et al. 2003). When implementing changes, firms may not possess all the required knowledge, skills, and resources. Nowadays, more and more businesses have decided to outsource some of their business functions and focus on their core competences (Shi et al. 2005). To better leverage the resources of their suppliers and/or clients, businesses form strategic alliance with their business partners to coordinate across firm boundaries and exploit opportunities together (Grover et al. 1996). Therefore, the capabilities to manage suppliers are more important than ever in the current business environment. Any internal changes may require corresponding changes in suppliers in a timely manner. Building missing resources will cause delay and loss of opportunities. To be effective in the acquiring and/or cooperating process, firms need supplier management capability to quickly locate needed resources, negotiate terms, and be able to trust and rely on partners for speed and quality. Therefore, in order to respond quickly to changes, firms must not only reconfigure internal resources and processes, i.e., process reconfigurability, but also acquire high-quality products/services from their partners in time to accomplish the changes. This capability also helps firms respond to changes at the operational level.

In all, we argue that operation-level agility focusing on change implementation in business operations is made up of (a) market responding capability, (b) process reconfigurability, and (c) supplier management capability. This operation-level agility is deemed critical to firms’ performance. Specifically, in fast-cycle industries, operation-level agility leads to more competitive actions in a rapid pace leading to an improved business performance (Meyer 2001, Sambamurthy et al. 2003, Venkatraman & Camillus 1984, Weill & Subramani & Broadbent 2002). In contrast, lack of operation-level agility implies a lack of responsiveness to the environment and the presence of inappropriate, outdated business activities and processes, thus resulting in poor performance. Therefore, especially in service industries where uncertainty and unpredictability are normal, the performance of a firm highly depends on its operation-level agility.

**H1. A higher level of operation-level agility will lead to a higher level of firm performance in service industries.**

### 2.2 IT service competence and operation-level agility

IT is fundamental to the growth of a business. IT has the potential to provide competitive advantages for businesses (Ravichandran & Lertwongsatien 2005). However, IT *per se* may not generate a
sustainable advantage since it can be easily acquired and imitated (Carr 2003). Also, investments in IT may not result in better firm performance since some IT investments can be wasted (Davern & Kauffman 2000). Instead, the implementation of IT competence, the extent to which a firm is knowledgeable about and effectively utilizes information technologies within specific business contexts, can create competitive advantage (Tippins & Sohi 2003).

Recent researchers argue that the relationship between IT competence and business values can be deconstructed through the presence of business competences. For example, Soh and Markus (1995) examine the need for effective deployment of appropriate IT assets to create business value. They argue that the effective use of these IT assets leads to intermediate effects, such as better business competences and processes, which, in turn, influence firm performance. Similarly, Ravichandran and Lertwongsatien (2005) argue that IT capabilities which support core competences of the firm such as market access competence, integrity-related competence, and functional-related competence can contribute to better performance. Tippins and Sohi (2003) confirm that business competences, such as organizational learning capability, mediates the relationship between IT competence and performance. These studies emphasize the importance of understanding the relationship between IT competence and business competences when understanding how IT influence business performance.

In this study, we argue that **IT service competence**, defined as the extent to which a firm can effectively utilize IT to support their businesses and facilitate operation, plays an important role in enhancing business performance by improving operation-level agility. We conceptualize IT service competence as a second-order construct, formed by IT service infrastructure, standardized application platform, and IT service management skills, arguing that IT service competence is central to operation-level agility.

**IT service infrastructure** is the sharable technical and common enterprise-wide platform, such as networking, database services, and standardized operation support, which enables initiatives, such as cycle time improvement and cross functional processes (Bharadwaj 2000). As the foundation of shared IT capabilities upon which the entire business depends, IT service infrastructure is crucial to operation-level agility (Byrd & Turner 2000). It links business units, implements common transaction processing, expedites business operations, allows to quickly access and share business data across the firm, and creates synergies across business units (Bharadwaj 2000, Ravichandran & Lertwongsatien 2005). A non-integrated IT infrastructure can severely restricts an organization's business choices and slow down business process, thus hindering operation-level agility.

**Standardized application platform** refers to enterprise-wide integrated software application platform and standard IT applications (Bharadwaj 2000). Developing standards for IT platform has been considered a priority in both research and professional communities (Markus & Steinfield & Wigand & Minton 2006). By providing uniform technical specifications, interfaces and criteria, it makes it easier to integrate new IT components and improve synergies between work units. It also has been long recognized that modularization is a good software development practice (Byrd & Turner 2000, Ravichandran & Lertwongsatien 2005). It constructs software from separate parts, called modules that separate logical boundaries between components. Such approach allows easy modification for new business processes and integration of new technologies with existing platforms, thereby allowing the IT unit to deliver new capabilities quickly and cost effectively, thus improving operation-level agility.

**IT service management skills** refer to the skill set of IT personnel to manage IT resources to deliver organizational IT services (Tippins & Sohi 2003). It includes knowledge and experiences of IT workers in dealing with daily operation of information systems (IS), handling requests from business users, and monitoring performance of IS to ensure that they meet business needs. Strong IT service management skills can help communication between the IT division and business users, integrate IT and business processes effectively, improve reliability and quality of IT, reduce cost of development and maintenance, and decrease delivery cycle time (Bharadwaj 2000, Byrd & Turner 2000). Therefore, IT service management skills are important part of IT service competence in the context of operation-level agility.
In all, we argue that IT service competence formed by (a) IT service infrastructure, (b) standard application platform, and (c) IT service management skills can support and enhance operation-level agility.

**H2. A higher level of IT service competence will lead to a higher level of firm performance in service industries.**

### 3 RESEARCH METHOD

A large-scale cross-sectional survey to collect firm-level data was conducted in the United States.

#### 3.1 Measurement development

The measurement development process involved three stages: (1) operationalization of research constructs, (2) item development, and (3) validity tests. First, research constructs were operationalized based on the definition of each construct as well as of relevant constructs in the literature. Second, every attempt was made to make use of existing measurements which have good psychometric measurement properties. Modifications of the existing items were also made to suit the context of the study. Table 1 provides a summary of the measurement items used in this study and their sources.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Summary of Measurement Items</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firm Performance</td>
<td>Competitive measures of customer retention, sales growth, profitability, and return on investment</td>
<td>Adopted from Tippins &amp; Sohi (2003)</td>
</tr>
<tr>
<td>Market Responding Capability</td>
<td>Organizational capability to obtain real time information about market changes, respond quickly to market requirements, and deliver products and services on time</td>
<td>Adopted from Grewal &amp; Tansuhaj (2001) and Wang et al. (2004)</td>
</tr>
<tr>
<td>Process-Reconfigurability</td>
<td>Organization capability to quickly reallocate resources, combine existing resources, and timely redesign/reconfigure business processes</td>
<td>Adopted from Pavlou &amp; El Sawy (2006)</td>
</tr>
<tr>
<td>Supplier Management Capability</td>
<td>Organization capability to change volume allocation among suppliers, acquire services and products from potential suppliers, and quickly change suppliers</td>
<td>Adapted from Swafford &amp; Ghosh &amp; Murthy (2006)</td>
</tr>
<tr>
<td>IT Service Infrastructure</td>
<td>Technology infrastructure to electronically link business units and partners, technology infrastructure to expedite business operations, network capacity and speed, and corporate data access</td>
<td>Adopted from Ravichandran &amp; Lertwongsatien (2005) and Weill et al. (2002)</td>
</tr>
<tr>
<td>Standardized Application Platform</td>
<td>Application infrastructure to allow reuse, modularization, integration, and standardization of common application components</td>
<td>Adopted from Bhatt &amp; Grover (2005) and Lee et al. (2008)</td>
</tr>
<tr>
<td>IT Service Management Skills</td>
<td>IT staff’s skills to prioritize and manage IT service requests, possession of well-defined service quality criteria for IT support, and possession of performance standards</td>
<td>Adopted from Tippins &amp; Sohi 2003 (2003)</td>
</tr>
<tr>
<td>Firm Size$^1$ (Control)</td>
<td>The number of full-time employees</td>
<td>Adopted from Tanriverdi (2005)</td>
</tr>
</tbody>
</table>

$^1$ This variable was adopted to control the potential compounding effect of available resources or managerial diseconomies on organizational high-level capabilities and performance (Tanriverdi 2005).

---

**Table 1. Measurement Sources for Research Constructs**
3.2 Research design

We conducted a large-scale cross-sectional survey with firms in service industries in the United States. The United States has experienced a fast growth of service industries and a nation-wide industrial transformation from manufacturing to service. Hence, it is considered as a viable background for examining the research framework of this study.

We applied a series of criteria congruent with the context of the study for the selection of the target samples. First, we focused on service industries which to a significant extent require IT support for their business operations, such as banking/finance, insurance, healthcare/medical, and consulting. Second, we excluded companies with fewer than ten employees from our target sample because such small companies do not provide a background appropriate for investigating capabilities in their operations and IT service.

After the target samples were defined, a cross-sectional survey was conducted using a web-based survey tool. Survey invitations were made to business executives, e.g., president, chief executive officer, chief operating officer, business director, and IT executives, e.g., chief information officer, chief technical officer, and IT director, of the sample firms in the target industries. Around 700 executives in an industrial respondent pool were invited to participate in this survey and a total of 147 complete data samples were achieved after removing small companies, incomplete data, and other inappropriate data, such as non-management positions and non-full-time employees. The final data represents 7 service industry types, such as healthcare/medical (70), banking/finance (32), consulting (21), insurance (17), and others (7). Their firm size varies; less than 250 (55), between 251 and 1000 (26), more than 1000 (66).

4 RESULTS AND ANALYSES

Partial least squares (PLS), a structural equation modeling technique, was used to analyze the data. This technique does not require a large sample size (Chin 1998). In addition, it is appropriate for early stages of theory development (Howell & Higgins 1990). Given that this study is an early attempt to develop a theoretical model that explains how a firm’s competence in IT services enables its operation-level agility leading to firm performance, PLS was considered to be appropriate for this study.

4.1 Measurement model evaluation

The validity of the measurement model was established prior to testing the structural model (Byrne 1998). The convergent validity of the reflective measures is determined in three ways: (1) the item reliability of each item, (2) the composite reliability of the construct, and (3) the average variance extracted (AVE) by the construct.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Number of Items</th>
<th>Item Reliability</th>
<th>Composite Reliability</th>
<th>AVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firm Performance (FPER)</td>
<td>4 items</td>
<td>.654 ~ .863</td>
<td>.881</td>
<td>.651</td>
</tr>
<tr>
<td>Market Responding Capability (MRES)</td>
<td>4 items</td>
<td>.763 ~ .882</td>
<td>.892</td>
<td>.674</td>
</tr>
<tr>
<td>Process-Reconfigurability (RCPR)</td>
<td>4 items</td>
<td>.852 ~ .917</td>
<td>.936</td>
<td>.784</td>
</tr>
<tr>
<td>Supplier Management Capability (SMGC)</td>
<td>3 items</td>
<td>.773 ~ .824</td>
<td>.841</td>
<td>.639</td>
</tr>
<tr>
<td>IT Service Infrastructure (ITSI)</td>
<td>4 items</td>
<td>.862 ~ .914</td>
<td>.938</td>
<td>.792</td>
</tr>
<tr>
<td>Standardized Application Platform (STAP)</td>
<td>4 items</td>
<td>.826 ~ .898</td>
<td>.923</td>
<td>.749</td>
</tr>
<tr>
<td>IT Service Management Skills (ITMS)</td>
<td>4 items</td>
<td>.860 ~ .919</td>
<td>.938</td>
<td>.791</td>
</tr>
</tbody>
</table>

Table 2. Result of Convergent Validity Test
Based on the results reported in Table 2, it was concluded that all the items demonstrated adequate convergent validity. Table 3 shows that the square root of the AVE for each construct was larger than the correlations between itself and the other constructs. This implies that each of the constructs shared greater variance with its own block of measures than with other constructs representing a different block of measures (Chin 1998). Therefore, this result demonstrates that there is good discriminant validity for the items used in this study.

<table>
<thead>
<tr>
<th></th>
<th>FPER</th>
<th>MRES</th>
<th>RCPR</th>
<th>SMGC</th>
<th>ITSI</th>
<th>STAP</th>
<th>ITMS</th>
<th>SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FPER</td>
<td>.807</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRES</td>
<td>.547</td>
<td>.821</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCPR</td>
<td>.483</td>
<td>.674</td>
<td>.885</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMGC</td>
<td>.213</td>
<td>.305</td>
<td>.365</td>
<td>.799</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITSI</td>
<td>.248</td>
<td>.341</td>
<td>.415</td>
<td>.203</td>
<td>.890</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAP</td>
<td>.363</td>
<td>.445</td>
<td>.450</td>
<td>.253</td>
<td>.699</td>
<td>.944</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITMS</td>
<td>.254</td>
<td>.286</td>
<td>.270</td>
<td>.199</td>
<td>.588</td>
<td>.581</td>
<td>.889</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>-.035</td>
<td>-.168</td>
<td>-.189</td>
<td>-.120</td>
<td>-.047</td>
<td>-.004</td>
<td>.037</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 3. Result of Discriminant Validity Test

4.2 Structural model analyses

The estimated path effects and the associated t-values were calculated using the Bootstrapping routine in PLS-Graph (version 03.00 build 1126). Since the operation-level agility and IT service competence were formulated as second-order constructs, the latent scores for each of the first-order constructs were calculated and used as measures for each construct (Chin & Gopal 1995). Figure 2 shows the results of the model analysis.

As shown in Figure 2, all paths were significant. The operation-level agility ($\beta = .554$, $t = 6.981$) was found to be a significant determinant of firm performance at the .01 level. It explained 29.8% of the variances of firm performance. On the other hand, the IT service competence ($\beta = .469$, $t = 6.191$) was found to be a significant determinant of the operation-level agility at the .01 level. It explained 25.4% of the variances of operation-level agility. All second-order loadings (loadings of the 1st order constructs) were highly significant. Therefore, all hypotheses H1 and H2 are supported. Interestingly, firm size, the control variable, was negatively significant ($\beta = -.181$, $t = 2.359$) at the .05 level in determining operation-level agility while it was not a significant determinant of firm performance.
4.3 Implications

The results indicate that a specific set of IT resources forms service organizations’ IT competence to support business needs, hence leading to a higher level of agility in their operations. In turn, this operation-level agility positively influences the competitive performance of the firms. To confirm these causal relationships among IT service competence, operational-level agility, and firm performance, we further conducted a post-hoc analysis of mediator test. Following Baron and Kenny’s (1986) four steps, we stepwise tested (1) the significant effect of IT service competence (IV) on firm performance (DV) without operation-level agility (MV) ($\beta = .344, t = 4.478, p < .01$), (2) the significant effect of the IV on the MV ($\beta = .468, t = 6.587, p < .01$), (3) the significant effect of the MV on the DV ($\beta = .553, t = 10.298, p < .01$), and (4) the insignificant effect of the IV on the DV in the copresence of the MV ($\beta = .099, t = 1.060$). According to Baron and Kenny (1986), the results indicate that operation-level agility is the full mediator between IT service competence and firm performance. Therefore, the post-hoc analysis results also confirm our research model proposed in this study.

In particular, our study shows that a firm’s IT service competence, formed by IT service infrastructure, standardized application platform, and IT service management skills, is a significant driving force for the firm’s agile operations which are represented by its rapid market responses, flexible process reconfiguration, and effective supplier management. The findings are consistent with Bharadwaj’s (2000) perspective that a firm’s IT resources, technology and human IT resources, are the sources of its latent capacity to build and provide the requisite IT services. Furthermore, the findings also provide an empirical evidence of the conceptual premise that IT is a digitized platform of business agility (Overby et al. 2006, Sambamurthy et al. 2003).

The study also shows that IT-supported operation-level agility significantly leads to a better performance, particularly in terms of sales growth, profitability, return on investment, and customer retention. The findings are consistent with the premise of the role of IT-enabled agility in generating higher performance of a firm in a turbulent environment. Considering the context of the study, i.e., the service industries, the findings are deemed important as well as useful to both academics and practitioners. Traditionally, the value of operation-based capabilities has been discussed in manufacturing settings. For example, many techniques of operational leanness and flexibility have been suggested under the settings of manufacturing, such as just-in-time manufacturing (JIT), total quality management (TQM), and lean manufacturing (Burgess 1994, Ettlie 1998, Gerwin 1993, Llorénsa et al. 2005). However, with regards to the recent catastrophe in the financial industries in the United States, it is likely to be more true that operational capabilities, such as operation-level agility, are crucial for firm success under the settings of service industries.

5 CONCLUSION

In this study, we concentrated on the service industries and theoretically proposed a positive relationship among IT service competence, operation-level agility, and firm performance. To capture the combinative values of IT resources and operation-level capabilities which form IT service competence and operation-level agility respectively, the second-order approach was adopted in conceptualizing the core research constructs. Survey data of small- to large-size enterprises in service industries of the United States were used to validate the proposed model. The results indicate that operation-level agility, consisting of market responding capability, process reconfigurability, supplier management capability, is a significant driving force of firm performance. The results also indicate that IT service competence, consisting of IT service infrastructure, standardized application platform, and IT service management skills, serves as a base for the operation-level agility.

This study has several limitations which involve cross-sectional research design and single respondent survey approach. First, this study used a cross-sectional research design. Such a snap-shot approach
may have limitations in terms of studying the causal relationships or time effects between research variables, such as the lead-time of the IT impact (Bharadwaj 2000). Second, single respondent bias has been discussed as a potential source of common method variance (Podsakoff & Organ 1986). To avoid this, various techniques have been suggested in the literature, such as separating survey questionnaire to ask specific expertise with different positions (Lee & Lim & Sambamurthy & Wei 2007). While our sample data are not thought to suffer from this issue\(^2\), multiple-respondent survey may provide more generalizable findings.

Regardless of the aforementioned limitations, this study makes several contributions to the literature. First, this study, both theoretically and empirically, reveals how firms can develop their operation-based agility. The findings of the study indicate the significant role of IT competence in achieving this specific type of agility. Since prior studies in IT-enabled agility have seldom provided empirical evidence, the theory-based models and the empirical findings of the study are both interesting and useful to academics in this research area. Second, this study has a potential contribution to the literature by addressing an unexplored, yet emerging issue of the role of operational capabilities in service industries. Since most of prior studies in operational capabilities have focused on the settings of manufacturing industries, this study may open a new area of discussion among academics and practitioners. Lastly, through this study, we develop new measurements having good psychometric properties. The measurements used in this study can serve as a base for further research in this research area. This study also has some practical contributions by providing guidance for practitioners to strategically invest their IT resources to achieve their agility at operational level.

This study can be extended in several directions. First, this study mainly focuses on operation-level agility in the service industry. Future research can examine another type of agility such as strategic agility. Second, this study investigates IT service competence, including IT service infrastructure, standardized application platform and IT service management skills. Future research can continue to study IT skills of business users that may also have influence on agility. Third, the unit of analysis of this study is the organization. It evaluates overall agility and its impact on the performance of an organization. However, in some large organizations, some divisions may be more agile than others. Future study can explore the topic at department or team levels. Fourth, our study adopts the survey methodology and measures the performance of an organization using historical/present data. Future research can develop a longitudinal study following the effectiveness of business capabilities.

Reference


\(^2\) Harman’s single-factor analysis was conducted to test a potential of common method variance in our sample data. This test involves an exploratory factor analysis (EFA) of all measurements to determine whether the majority of the variance is accounted for by one general factor. The principal component analysis using varimax rotation revealed that a total of seven distinct factors emerged with Eigen value of over 1. The results revealed that each of the seven principal components explained similar amounts of the total variance of 76%, ranging from 5% to 14%. This result indicates that our data does not suffer from common-method bias as indicated by Podsakoff et al. (2003).


UNDERSTANDING AGILITY IN SOFTWARE DEVELOPMENT THROUGH A COMPLEX ADAPTIVE SYSTEMS PERSPECTIVE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0561.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Information Systems Development (ISD), Complexity / Complex adaptive systems, Agile computing, Case Study</td>
</tr>
</tbody>
</table>
UNDERSTANDING AGILITY IN SOFTWARE DEVELOPMENT FROM A COMPLEX ADAPTIVE SYSTEMS PERSPECTIVE

Xiaofeng Wang, Lero, the Irish Software Engineering Research Centre, Limerick, Ireland, xiaofeng.wang@ul.ie
Kieran Conboy, National University of Ireland, Galway, Ireland, kieran.conboy@nuigalway.ie

Abstract
Agile software development methods have emerged in recent years and have become increasingly popular since the start of the century. While much research claims to study agile methods, the meaning of agility itself in software development is yet to be fully understood. Agility is viewed by some as the antithesis of plan, structure discipline and bureaucracy. This study aims to develop a better understanding of agility, using the key concepts of Complex Adaptive Systems as a theoretical lens. The study explores agility from several different angles, including autonomous team, stability and uncertainty, and team learning. A multiple case study research method was employed. The findings of the study emphasize that agility is manifested as stability and discipline, which are just as desirable as flexibility, and context sharing is of the same value and importance as knowledge sharing. In addition, the collective nature of learning is underlined.

Keywords: agility, complex adaptive systems, autonomy, stability, team learning
1 INTRODUCTION

The last ten years or so has seen the emergence of agile software development methods as a response to the inefficiency of existing software development methods in rapidly changing environments (Highsmith 2002), e.g. eXtreme Programming (XP) (Beck 1999) and Scrum (Schwaber & Beedle 2002). A brief reflection on the history of the agile software development movement, however, reveals that agile methods originated as a set of techniques and practices, and the term agile is more a post-rationalization to justify a set of existing “light-weight” methods. Agility in software development has been interpreted in many different ways in practice. Skepticism and criticism of agile methods place agility to the opposite of plan, structure and discipline which are generally considered the core components of more traditional waterfall methods (Rakitin 2001, Stephens & Rosenberg 2003).

To clarify the meaning of agility, Conboy and Fitzgerald (2004) conduct a review of the literature on agility across several disciplines including manufacturing, business and management, and carefully distinguish several intertwined concepts, including flexibility and leaness. Based on the comparison and contrast of these concepts, they provide a broad definition of agility as “the continual readiness of an entity to rapidly or inherently, proactively or reactively, embrace change, through high quality, simplistic, economical components and relationships with its environment” (Conboy & Fitzgerald 2004, p.40). Lyytinen and Rose (2006) explore agility in an information systems development (ISD) context. They claim that ISD agility is concerned with why and how ISD organizations sense and respond swiftly as they develop and maintain information system applications. They outline a theory of ISD agility drawing upon a model of Information Technology (IT) innovation and organizational learning which adopts March’s (1991) concepts of exploration and exploitation. Their empirical study shows that the concept of ISD agility is more multifaceted and contextual than conceived so far in the literature. It relates to being nimble in terms of the velocity to absorb base innovations and innovate with IS products; the velocity to shift from one innovation regime to another (organizational flexibility); the velocity to learn from experiences (trial and error learning); and the velocity to deliver IS solutions. Each one of these demands different competencies and expects managerial shaping of alternative organizational goals and incentives. Their findings suggest that the dynamics and interactions between these four types of agility form different ecological niches. Each one follows a different organizing logic. Managers must view the meaning of agility differently in each niche.

While these studies help to understand agility, and do highlight the lack of theoretical foundation regarding agility in an ISD context, they do not address specifically how agility is manifested in software development environments. Based on this observation, this study investigates the meaning of agility in software development using the lenses of Complex Adaptive Systems (CAS), an important branch of the complexity study which provides insights of how a system can be adaptive to its environment. (Note that in the following sections the full phrase complex adaptive system is used to refer to an instance of a complex system that demonstrates an adaptive nature, while CAS is used to refer to the study and theory of such systems.) The empirical part of the study employs a multiple-case study approach. The remaining part of the paper is organized as follows. Section 2 introduces the key concepts of CAS and builds a conceptual framework based on CAS which guides the empirical investigation; Section 3 describes the research method and the context of the empirical study; then the findings are presented in Section 4 and discussed in Section 5. The paper ends up with a conclusion section where the implications and limitations of the study are reviewed and the future work summarized.

2 A COMPLEX ADAPTIVE SYSTEMS PERSPECTIVE ON AGILITY

A complex adaptive system, roughly defined, consists of a large number of agents, each of which behaves according to some set of rules. These rules require agents to adjust their behaviour to that of other agents. They interact with, and adapt to, each other. CAS seeks to identify common features of
the dynamics of such systems or networks in general (Stacey 2003). There is no single and definitive account of CAS. Anderson (1999), Mitleton-Kelly (2003) and Stacey (2003) provide valuable introductions to CAS in the context of organization and management. Four key concepts of CAS in the centre of these accounts are of particular relevance to this study: inter-connected autonomous agents, self-organization, the edge of chaos and emergence. These key concepts provide a new perspective to investigate different facets of agility as a desirable property for software development teams in constantly changing environments.

The concepts of inter-connected autonomous agents and self-organization suggest that, to be agile, a software development team should be composed of autonomous members who have their own schemata, which generally refer to norms, values, beliefs, and assumptions that are held by individuals (Senge 1990, Schein 1997). Team members are interconnected in such a way that a decision or action by any individual may affect related individuals and the team. A team composed of autonomous but inter-connected members can spontaneously come together to perform a task (or for some other purpose); the team decides what to do, how and when to do it; and no one outside the group directs those activities (Mitleton-Kelly 2003). To do so, a team needs energy imported into and constantly flowing within it, which can be interpreted, partly, as the sharing of information, knowledge or other resources needed to sustain self-organized activities.

The edge of chaos provides organizations “with sufficient stimulation and freedom to experiment and adapt but also with sufficient frameworks and structure to ensure they avoid complete disorderly disintegration” (McMillan 2004, p. 22). Brown and Eisenhardt (1998) contend that, to compete at the edge, organizations must understand what to structure and what not to structure, to foster communication and to capture cross-business synergies. The edge of chaos concept suggests that being agile is neither chaotic nor static. It needs stability but not so much that order prevails and innovation is stifled. It is a delicate balance of both.

The concept of emergence sheds new light on learning, which can be seen as a collective behavior of creating new patterns of thought at the team level based on the interaction of individuals, instead of often seen exclusively as the provision of individual training. Learning means not only training or the acquisition of new skills, but also the gaining of insight and understanding which leads to new knowledge and behavior. When learning leads to new behavior, the team can be said to have adapted and evolved (Mitleton-Kelly 2003). An agile team facilitates team learning and generation of new knowledge. In addition, new knowledge needs to be shared to generate further new learning, knowledge and behavior.

In summary, this study investigates the meaning of agility from three facets: autonomous but sharing team, stability with embraced uncertainty and team learning, as shown in Table 1.

<table>
<thead>
<tr>
<th>Facets of Agility</th>
<th>Underlying CAS Concepts</th>
<th>Relevant Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autonomous but sharing team</td>
<td>Inter-connected autonomous agents</td>
<td>Anderson 1999; Choi et al. 2001;</td>
</tr>
<tr>
<td></td>
<td>Self-organization</td>
<td>Mitleton-Kelly 2003</td>
</tr>
<tr>
<td>Stability with embraced</td>
<td>The edge of chaos</td>
<td>Brown and Eisenhardt 1998;</td>
</tr>
<tr>
<td>uncertainty</td>
<td></td>
<td>Stacey 2003</td>
</tr>
<tr>
<td>Team learning</td>
<td>Emergence</td>
<td>Mitleton-Kelly 2003; Stacey 2003</td>
</tr>
</tbody>
</table>

*Table 1. Agility through the CAS perspective*

### 3 RESEARCH APPROACH

This study adopts an interpretivist stance, emphasizing that agility are situational and can be better understood through the understanding and sense making of people who are involved in software development. In particular, this study employs a qualitative approach, treating agility as a qualitative...
property of a software development team that can be better studied through words and the meanings people ascribe to them rather than numbers or frequencies. The specific research method used in this study is case study, which is an appropriate approach when a research phenomenon is investigated in its real-live context (Yin 2003). A multiple-case design is employed. Given the research focus of the study, the level of inquiry is at the team level, so it seems appropriate to take a software development team as a case. The unit of analysis is the software development team. Three software development teams - XPTeam A, XPTeam B and WaterfallTeam - from two different companies were chosen as the cases. XPTeam A is a representative case; XPTeam B is a confirming case of the first one; and WaterfallTeam is a contrasting case, following the strategy suggested by Yin (2003). The profiles of the three cases are shown in Table 2. XPTeam A is a software development team in SecureSoft, a small software house specialized in network security and management systems development. XPTeam B and WaterfallTeam are software development teams in WorldTech, a major IT company providing both IT projects and services.

<table>
<thead>
<tr>
<th></th>
<th>XPTeam A</th>
<th>XPTeam B</th>
<th>WaterfallTeam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team size</td>
<td>4</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Team composition</td>
<td>3 developers, 1 project manager</td>
<td>6 developers, 1 test manager, 1 project manager</td>
<td>4 developers, 1 project manager</td>
</tr>
<tr>
<td>Development method</td>
<td>XP</td>
<td>XP</td>
<td>Waterfall style mixed with some agile elements</td>
</tr>
<tr>
<td>Years of method use</td>
<td>4.5 - 5 years</td>
<td>11 months to 1.5 years</td>
<td>More than 5 years</td>
</tr>
<tr>
<td>Location</td>
<td>Co-located in an open office space</td>
<td>Co-located in an semi-open office space</td>
<td>Collocated in an semi-open office space</td>
</tr>
<tr>
<td>Software developed</td>
<td>Application for external customer</td>
<td>Web application for internal use</td>
<td>Backend application for internal use</td>
</tr>
</tbody>
</table>

Table 2. The profiles of the three cases

Two rounds of data collection are conducted. The interval between the two rounds is six months. The main data collection method used is semi-structured face-to-face interviews. The questions are all open-ended. The members of each team are interviewed. Each interview lasts between 30 minutes to two hours. In all the cases, most interviewees are interviewed twice. Table 3 lists the people interviewed in each team. Documents regarding the development processes of the case teams are collected when available. Some non-participative observations are conducted as the opportunities occur. Field notes are taken during both rounds of data collection.

<table>
<thead>
<tr>
<th></th>
<th>XPTeam A</th>
<th>XPTeam B</th>
<th>WaterfallTeam</th>
</tr>
</thead>
<tbody>
<tr>
<td>First round interviews</td>
<td>1 group interview (with the 4 team members below), 4 individual interviews</td>
<td>5 individual interviews</td>
<td>1 individual interview</td>
</tr>
<tr>
<td>- Project manager</td>
<td>- Project manager</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Coach</td>
<td>- Team lead</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Developer A</td>
<td>- Tech lead</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Developer B</td>
<td>- Developer A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Test manager</td>
<td>- Test manager</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second round interviews</td>
<td>2 group interviews (with the team members below), 3 individual interviews</td>
<td>6 individual interviews</td>
<td>3 individual interviews</td>
</tr>
<tr>
<td>- Coach</td>
<td>- Project manager</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Developer A</td>
<td>- Project manager</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Developer B</td>
<td>- Team lead</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Developer B</td>
<td>- Tech lead</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Developer C</td>
<td>- Developer A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Test manager</td>
<td>- Developer B</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Two rounds of interviews

The data analysis includes two steps: within-case analysis and cross-case comparison (Eisenhardt 1989). The emphasis is on the cross-case comparison, in which an analysis tactic suggested by
Eisenhardt (1989) is used: the three cases were divided into two groups, XPTeam A and XPTeam B in one group as the cases using agile approach, while WaterfallTeam in the other group as the case that uses waterfall approach. XPTeam A and B are compared firstly for similarities and differences, and then they as a group are contrasted with WaterfallTeam for similarities and differences.

4 MANIFESTATION OF AGILITY IN THE THREE TEAMS

This section presents how agility has been manifested (or shown to be absent) in the three cases.

4.1 Autonomous but sharing team

Team autonomy in XPTeam A and B firstly is shown as competences relevant to software development being distributed among team members. The members of the two teams are involved in all development activities of their projects, and all have to deal with the customers, analyse user requirements and write code together. There are no traditional roles such as system analyst, designer or programmer. Each team member is able to assume all the roles, since comprehensive competences are required to work with user stories, the implementation of which is self-contained and encapsulates different development activities:

“The problem is not to have three persons for analysis, or two persons for design, but a user story inside has to resolve analysis, developing, and, etc., everything.” (Project manager/XPTeam A)

For example, when XPTeam B started the project, there were big gaps among team members in terms of Java related knowledge and skills. With the project going on, the developers with less Java experience learnt quickly from those more experienced, and the team members reached fairly the same level of competence. As a result, there is no dependency on a particular individual, since each team member gets exposure to different areas of a project. Distributed competence is shown in the case of WaterfallTeam too, although the team uses waterfall approach. Like the other two teams, there are no specific roles like analyst, designer or coder in the team. The developers are not specialized on specific tasks. Everybody has chances to do different things.

Team autonomy is also manifested as a disciplined team in XPTeam A and B, which is seemingly contradictory to the idea of autonomy. However, both teams reckon the importance of disciplines. As a member of XPTeam B describes, disciplines are necessary components of an agile process, and they come from the process the team uses:

“There is a set of rules really, and you may not adopt them, you probably adopt most of them, and those rules kind of direct you really, it’s like you need to formalize it so you can be more flexible.” (Test manager/XPTeam B)

Team autonomy does not mean the team members are working on their own; instead, there is constant sharing among them. XPTeam A considers sharing an important aspect of team working. They believe that, as a team, they have to face every moment in any case without barriers. Sharing is also seen as a contributor to a team’s agility by WaterfallTeam who works with the waterfall approach. The difference is that sharing in the two teams using agile processes goes beyond simply knowledge sharing. It extends to context sharing and the sharing of achieved results. What is shared among the team members is not only the technical knowledge related to different areas of a project, which helps to distribute competences among them. It is also the knowledge about who knows what, which is particularly important for a bigger team like XPTeam B, and helps the team members self-organize to implement tasks:

“I think the ten o’clock stand-up meeting is definitely good, because you know what everybody else on the project is working on, and you might say ‘I’m working on this and I’m not sure how to’… and someone says ‘oh yeah actually I did it yesterday’.” (Developer B/XPTeam B)
In XP Team A and B, the developers are attentive to what happens around them, with the help of the open space the teams are working in:

“When you are doing something, you have to listen what the pair, or the single one if you are in pair, what he's doing, what they are saying, you have one ear in this way and the other (in the other way).” (Developer A/XP Team A)

Collective ownership of results is another kind of sharing. The two teams using the agile processes both endorse the collective ownership of code, as suggested by XP. A developer of XP Team B, however, warns that collective ownership can become collective irresponsibility sometimes, which means no one claims to be responsible if there is some problem with a piece of code. In the case of XP Team A, in addition to collective ownership of code, the team also owns collectively other forms of working results, such as designs, solutions, etc., which helps the team to have a sense of common achievement.

4.2 Stability with embraced uncertainty

Stability for software development is a desired property by all three teams, which is seen as an indispensable component in responding to change:

“There has to be some limitation of what you are doing, you cannot be so flexible that things are chopping and changing every single day.” (Test manager/XP Team B)

Stability first of all is demonstrated as a short-term certainty in all the three teams. The short-term certainty means a team has a very clear idea of what they have to do in a short time frame, such as one-week or two-week iterations in the cases of XP Team A and B. Waterfall Team also realizes the importance of the short-term certainty to deal with constant changes from the management:

“Well I guess in terms of uncertainty, you don't know really tomorrow you are going to work on the same project, so on a phased approach you can complete one phase and then this is done. And say after tomorrow, let's say the next phase is cancelled because of the management decision, then you still have a product that works.” (Developer A/Waterfall Team)

In the cases of XP Team A and B, stability is also shown as a sense of frequent achievement and satisfaction. The two XP teams realize that, with their agile process, the team members can be motivated more easily than with the waterfall method, since the developers can see the result of their work at the end of each iteration, rather than working for six months without anyone has ever seen or used the code produced as what can happen in traditional processes. There is evidence to suggest that the developers of Waterfall Team also recognize the importance of motivating people, and believe that a satisfied and motivated team is a source preventing a project from falling apart:

“If someone is not happy with what he's doing, he's not going to do his job well. If he doesn't like it, he doesn't like to co-operate, if he's not happy with people, he wasn't going too far... So the main thing is with people, keep them happy... because if people are unhappy, the project falls apart.” (Developer A/Waterfall Team)

In addition, a team focused on working is also a sign of stability. A focused team has several meanings in the two XP teams: one meaning is to focus on work in a short but appropriate amount of time. It can be an iteration, as in XP Team A and B. Another meaning of being focused is to focus on current work, not wasting time to do future-proof work, which has been emphasized particularly in XP Team B. The third meaning is to focus on development activities and not to mix them with personal desires of learning new things. For example, XP Team A is very attentive of keeping the team focused on development activities by reserving daily studying time to satisfy the developers’ desires to learn.

Last but not least, stability shows as team working at a sustainable pace, with ease and without anxiety, is another aspect of the stability for development. A developer of XP Team A associates this working state with agility directly:
“I think agility is a state of mind... you don't have to feel anxiety, you have to be relaxed when you approach a problem, and XP or Scrum is just a method to obtain this kind of relaxation... If you are happy on what you are doing, if you are not stressed, I think you can say you are agile.” (Developer B/XPTeam A)

Stability co-exists with uncertainty which is unavoidable in the teams using agile methods. Uncertainty needs to be embraced. Embraced uncertainty is manifested firstly as the probability to change directions in the cases. All the teams believe that the iterative nature of their processes gives them more possibility to change directions when needed, including WaterfallTeam, since they use iterative phases within the waterfall process. But the probability to change should be complemented by having a whole picture of the project, which has been emphasized in the two XP teams. XPTeam B observes that having a whole picture of the project occurs not only to the developers, but also to their onsite customer.

4.3 Team learning

XPTeam A understands that learning means doing things differently. If a team wants to be adaptive and evolve, they have to learn. In the two XP teams, learning happens as team learning rather than individual learning, which means a team as a whole acquires new knowledge and competences, and the results of learning are shared among team members. Compared with WaterfallTeam, team learning happens continuously and mutually, through using agile practices in the two XP teams. It happens in daily development activities. It is a continuous experience for the team members. Meantime, since learning happens through interactions among the developers, it is generally bi-directional. A developer of XPTeam B comments:

“I think it (XP) is a very good way of learning as well, because with pair programming which is part of it, you are learning from somebody different every day, and likewise you're able to teach somebody else for you've been doing the day before ... it gives a sense of shared, the project is shared... There's more, definitely more knowledge been shared.” (Developer C/XPTeam B)

Besides, learning is not a daunting experience due to the fact that the teams using the agile processes generally work on small pieces of tasks. The developers learn gradually through implementing them, sometimes with the help of others. The team lead of XPTeam B observes that:

“Because it is down to granular level, it's easier to put better workload over people and also easier for people to get involved, it’s also easier for people who don't have skill learn gradually on the smaller story rather than having to develop something big on their own, so I think it's easier to get a higher level skill without being overly complicated... They are not huge chunk of piece to take on.” (Team lead/XPTeam B)

Due to these attributes, team learning is seen more efficient than individual learning:

“The learning, when we do pair programming it's more efficient. In one year I learn a lot of things that I didn't think (I could do) when I was in the university.” (Developer B/XPTeam A)

Table 4 summaries the findings.

5 DISCUSSION

As shown in this study, agility in the context of software development is highly multifaceted and ambiguous. In this section the different facets of agility demonstrated in the cases are discussed by drawing on relevant agile literature.

5.1 An autonomous but sharing team

Despite the suggestion by advocates of agile that software development processes should be organized to improve and distribute both technical and social competences continuously (Cockburn & Highsmith
2001), few empirical studies in agile research have supported this stance. Only Auvinen et al. (2006) highlight an increased competency in a team where several agile practices are piloted. Similarly, no empirical research in the reviewed literature focused on discipline in agile processes despite the emphasis many agilists place on its importance (e.g. Beck & Boehm 2003).

<table>
<thead>
<tr>
<th>Agility through CAS</th>
<th>Manifested in software development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autonomous but sharing team</td>
<td>Distributed competences</td>
</tr>
<tr>
<td></td>
<td>Disciplined team</td>
</tr>
<tr>
<td></td>
<td>Knowledge sharing</td>
</tr>
<tr>
<td></td>
<td>Context sharing</td>
</tr>
<tr>
<td></td>
<td>Collective ownership of results</td>
</tr>
<tr>
<td>Stability with embraced uncertainty</td>
<td>Short-term certainty</td>
</tr>
<tr>
<td></td>
<td>Team being satisfied, motivated and focused</td>
</tr>
<tr>
<td></td>
<td>Working at a sustainable pace</td>
</tr>
<tr>
<td></td>
<td>Probability to change directions</td>
</tr>
<tr>
<td></td>
<td>Having a whole picture of the project</td>
</tr>
<tr>
<td>Team learning</td>
<td>Learning continuously</td>
</tr>
<tr>
<td></td>
<td>Mutual learning</td>
</tr>
<tr>
<td></td>
<td>Learning gradually</td>
</tr>
</tbody>
</table>

*Table 4. Manifestation of agility in software development*

This study suggests that a team composed of autonomous but interacting developers has a tendency to be agile. Each of them is able to solve various development issues and to interact with customers. Competences are not concentrated on few people so that there is no bottleneck in the development process. Team members are confident and courageous in the interactions with customers and with each other. They are also mature and willing to try new things. An autonomous team, however, does not mean team members can be completely amethodical and ill-disciplined. On the opposite, it is composed of disciplined, self-responsible and committed individuals. Discipline is an essential component of an autonomous team, and is drawn from the interactions among peer team members.

Sharing is a common theme investigated in several agile studies, though most are focused on knowledge sharing (Fredrick 2003, Melnik & Maurer 2004, Poole & Huisman 2001, Schatz & Abdelshafi 2005). Context sharing has also been observed, but is somewhat understated in agile literature. Melnik and Maurer (2004) believe that the so-called “background knowledge” about a project is important to achieve effective communication. It is important for all team members to have a common frame of reference - a common basis of understanding. Poole and Huisman (2001) observe that, in the organisation they studied, there was a measurable increase in the visibility of what everyone was doing on the team subsequent to the adoption of the agile practices. In fact, this improvement in visibility is considered one of the greatest successes the company has achieved. In terms of results sharing, Fredrick (2003) reports the experience of collective ownership of codes. When it is realized, even the most complex business problems can be easily figured out. In contrast, it was found that individual ownership of code made people defensive - people took it personally when someone suggested their code did not work. Schatz and Abdelshafi (2005) also document the collective ownership in their experience report where developers took ownership of the features they created and took pride in showing their work to the stakeholders during sprint reviews. Rising and Janoff (2000) notice that in a team they have studied, at every meeting, as small tasks were completed and the team could see progress toward the goal, everyone was more satisfied with their work and project progress.

The findings of this study confirm that sharing in an agile team not only means knowledge sharing. Context sharing is equally important. To effectively self-manage, a team needs to share the understanding of their working context. Context sharing is a precondition to provide effective feedback, interpret them in a sensible way, and take appropriate actions. Sharing also means results
sharing, such as collective ownership of code and solutions, which reduces the risk of knowledge loss and increases the sense of being a true team.

Another type of sharing, namely problems sharing, is reported by Rising and Janoff (2000) but does not emerge in this study. In the team they have studied, when one team member raises an obstacle in the Scrum meeting, the entire team’s resources come together to bear on that problem, and the entire team immediately owns any one individual’s problems.

5.2 Stability with embraced uncertainty

Several agile studies have noticed team satisfaction and motivation in agile processes (e.g. Rising & Janoff 2000, Poole & Huisman 2001, Drobka et al. 2004). For example, Drobka et al. (2004) conduct a survey of a team using XP and find that it creates a surge in morale since XP provides constant feedback to the developers and at the end of each day the team has a working product. Team members gain a sense of accomplishment from their daily work, because they immediately see the positive impact their efforts have on the project. When morale is high, people are excited about their work, leading to a more effective, efficient development team. Short-term certainty has also been noticed in agile studies, though not so extensively. Murru et al. (2003) claim that XP enhances programmers’ sense of project control. They find that programmers with the experience of Rational Unified Process (RUP) felt that XP’s planning game gave them a stronger feeling of control than traditional planning did. They knew where their project was going and whether it was delayed. Furthermore, programmers were more aware of keeping the project’s strategic goals in focus. This knowledge improved the programmers’ motivation.

The role played by uncertainty is acknowledged by agile advocates (Highsmith & Cockburn 2001, Williams and Cockburn 2003). Williams and Cockburn (2003) believe that uncertainty is inevitable in all software development. Many changes occur during the time that the team is developing the product. It is highly unlikely that any set of predefined steps will lead to a desirable, predictable outcome. It is necessitated short “inspect-and-adapt” cycles and frequent, short feedback loops. Agile software development is about change and feedback. Highsmith and Cockburn (2001) claim that agile organizations and managers understand that to demand certainty in the face of uncertainty is dysfunctional, and agile practices encourage change rather than discourage it. In turbulent business situations, the change tolerance of a development process must be geared to the change rate of a specific environment, not some internal view of how much change is acceptable. Despite these claims of agile proponents, however, few empirical studies of agile processes have focused on uncertainty and how it is embraced, with the exception of Elssamady and Schalliol (2002) who suggest that, when using the XP practices, especially the simple design, one should look ahead and do things incrementally, in order to have a big picture.

This study emphasizes stability as a desired property of development teams that have to deal with continuous changes due to close relationships with customers and evolving requirements. A team needs stability, needs to find a proper heartbeat for their development process so that it would not be dissolved into turbulence. Stability gives developers a sense of security and control over what they are working on. It can be drawn from a short-term certainty provided by a time-boxed development process. Stability for development also means a team is working at a sustainable pace, focused and motivated, working with ease and satisfaction. Certainty and security is only for a short term, however. Uncertainty is inevitable in software development. It comes from both the environment a team is embedded in and the development process itself. Managing uncertainty does not mean to predict what is going to happen and do future proof work today. It is to ensure the probability to change the direction a team goes towards but meantime not to get short-sighted. Team members need to have a whole picture of the project in mind.
5.3 Team learning

Learning is a common theme explored in much agile research (e.g. Dingsoyr & Hanssen 2002, Drobka et al. 2004, Hunt & Thomas 2003, Meso & Jain 2006), but the focus is mainly on individual rather than team learning. In a survey conducted by Drobka et al. (2004), it was found that XP can reduce the learning curve for new team members. Fifty-five percent of the developers believed that using XP shortened their initial project-learning curve. Hunt and Thomas (2003) emphasize that learning in an agile process is a continuous process, and it means learning about more than just the technology involved. It covers how the team works together (or how it doesn’t) and team members themselves, which leads to behavior and mental model change.

Learning means doing things differently. One important consequence of learning for an individual and a team is to change either their behavior or mental model. It is a prerequisite for organizational evolution and co-evolution (Mitleton-Kelly 2003). This study suggests that team learning is different than individual learning, though closely related and dependent on it. From the CAS perspective, team learning is emergent from the interactions of team members. Team learning is a collective result, which means a team as a whole acquires new knowledge and competences, and results of individual learning are shared among team members. In an agile team, team learning happens constantly, mutually and gradually.

6 CONCLUSION

This study investigates how agility is manifested in agile software development through studying the software development processes of three teams, two using XP, one using waterfall approach. Taking the key concepts of CAS as theoretical lenses, the study explores the true meaning of being agile from several different angles, including autonomous team, stability and uncertainty, and team learning. Compared with the existing agile literature, the findings emphasize that stability and discipline are as desirable as flexibility, and context sharing is of the same value and importance as knowledge sharing in agile processes. In addition, the collective nature of learning is underlined. The main theoretical contribution of the study is the understanding of agility in software development which is both theory-informed and empirically grounded. Drawn on CAS, the study lifts the understanding beyond the advocational literature found in agile field (Baskerville & Pries-Heje 2004). The discovered agile properties enrich the understanding of agility in software development. The practical implication of the study is that the findings indicate the desired effects of using agile methods. The agile properties provide a software development team with observable agile indicators from different facets of software development.

One limitation of the study comes from using CAS as the theoretical basis of the study. CAS has originated in natural sciences. There is a deeper concern whether CAS is appropriate to the study of human organizations. A combination of CAS theory with appropriate social theories might be a promising avenue for future research. Some limitations are associated with the case study research approach. One concern is the uniqueness of corporate, team and project characteristics of each case which makes valid comparison and theoretical generalization of the case study results difficult (Kitchenham et al. 2002). Specific to this study, an affecting factor is the diversity of the team profiles (as shown in Table 2). To increase the validity of the study, the contextual information of the cases has been taken into account in the data analysis. Another limitation is that only one agile method, XP, has been involved in our case studies. Future work would be to verify if the findings apply to teams using other agile methods, such as Scrum, Lean system development, etc.

References

A Case Study of Risk Management in Agile Systems Development

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0642.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Agile computing, Risk, Management practices, IT Project Management</td>
</tr>
</tbody>
</table>
A CASE STUDY OF RISK MANAGEMENT IN AGILE SYSTEMS DEVELOPMENT

Coyle, Sharon, Centre for Innovation and Structural Change, J.E. Cairnes Graduate School of Business & Economics, National University of Ireland, Galway (NUIG), Ireland
sharon.coyle@nuigalway.ie

Conboy, Kieran, Centre for Innovation and Structural Change, J.E. Cairnes Graduate School of Business & Economics, National University of Ireland, Galway (NUIG), Ireland
kieran.conboy@nuigalway.ie

Abstract

A principal objective of agile development methods is to reduce risk thereby resulting in more successful and effective information systems. However, while there is an extensive body of academic literature on risk management, very little research has attempted to rigorously apply this as a lens to study risk management in agile development projects. The purpose of this research was to ascertain the extent to which risk management practices are incorporated into agile development projects. A case study was conducted involving a change management consultancy firm whose main business involved the deployment of agile approaches. The study found that while there is a perception that risk management practices are not carried out in agile environments, the extent to which risk management is conducted on these projects is in some cases just as substantial to that carried out on traditional projects. In fact, in this case it appeared that the teams were more rigorous in their risk management than is often the case.
1 INTRODUCTION

In ISD, the rapidly growing use of agile methods shows the urgency of organisations to adapt to change at a more speedy and efficient pace where practitioners have “developed methodologies and practices to embrace, rather than reject higher rates of change” (Williams and Cockburn, 2003). Many agile methods have emerged in recent years to address this issue e.g. eXtreme Programming (XP) (Beck, 1999), the Dynamic Systems Development Method (DSDM) (Stapleton, 1997), Scrum (Schwaber and Beedle, 2002), Crystal (Cockburn, 2001), Agile modelling (Ambler, 2002), Feature Driven Design (FDD) (Coad, de Luca et al., 1999), and Lean Software Development (LSD) (Poppendieck, 2001), along with variants of each e.g. XP-Lite (Aveling, 2004). Agile methods are known for their use of iterative development, active user involvement and their acknowledgement of the need to incorporate changing system requirements and “focus on generating early releases of working products using mostly collaborative techniques” (Reifer 2002). This is a stark contrast to the traditional model for systems development which promotes “elicitation and freezing of requirements in advance” (Fitzgerald 2000) with no overlap between project phases of analysis, design and implementation (Walters, Broady et al. 1994).

A principal objective of agile methods is to reduce well-known risks associated with common ISD project failures by for example, accepting that requirements will change. The first “truly agile method” (Abrahamsson, Warsta et al. 2003) is known as the Dynamic Systems Development Method (DSDM) which promotes the idea that “instead of fixing the amount of functionality in a product and then adjusting time and resources to reach that functionality, it is preferred to fix time and resources and then adjust the amount of functionality accordingly” (Abrahamsson, Warsta et al. 2003). This highlights the element of flexibility in DSDM with regard to adjusting system functionality where system requirements are open to change. However, no matter what the nature of change, there will always be associated risks involved be it either a systems development change or a business operational change.

1.1 Motivation For Research

The idea of being proactive to change is becoming increasingly important across all industries due to an increase in project failures. There are infinite examples across the literature of project failures and media reports of “major engineering and development projects running late or exceeding their budget” (Coppendale 1995). Risk management has been described as the “activity of identifying and controlling undesired project outcomes proactively” (Smith and Merritt 2002). One of the main reasons highlighted for the increasing failure in software projects is that “managers are not taking prudent measures to assess and manage the risks involved” in their projects (Keil, Cule et al. 1998). According to Smith and Merritt (2002), a lack of proactiveness with regard to risk management is one of the main reasons for project delays, increased project running costs and/or eventual project failures.

While there is extensive literature on risk management, research in relation to risk management in agile ISD projects is non-existent. This is surprising considering how quickly agile methods are being adopted in ISD. Many books on agile methods “have remarkably little to say about how a development team determines the risks it faces, prioritises them or takes action to negate their effects” (Smith and Pichler 2005). Essentially, agile methods must “tailor conventional risk management approaches meant for years-long projects into a risk driven agile iteration lasting only seven to thirty days” (Smith and Pichler 2005). How agile projects go about doing this remains unknown.

The primary objective of this research was thus to develop a better understanding of risk management practices in agile ISD projects and the level of formality with which these practices are executed. Specifically, this research focuses on three main elements of risk management, namely risk identification, estimation and evaluation.
2 THEORETICAL BACKGROUND

In an ISD context, Barry Boehm highlighted the concept of managing risks and giving them priority as far back as 1988. Ten years later, Hall (1998) described Boehm as being “the father of software risk management.” Boehm proposed a move away from the ‘staged’ SDLC (Systems Development Lifecycle) to a more iterative or incremental process and this proposed concept in software development is an attempt to lower project risks (Powell and Klein 1996). Boehm’s aim was to eliminate any software difficulties or risks mainly by deriving “risk-driven documents” and “incorporating prototyping as a risk-reduction option” (Boehm 1988). It resulted in what was called the ‘Spiral Model’ that essentially created “a risk-driven approach to the software process rather than a primarily document-driven or code-driven process” (Boehm 1988).

Many authors highlight distinct approaches and frameworks for dealing with risk management. The approaches used can be formal or informal. In an informal environment “much risk identification is undertaken by managers and associated staff without any aiding techniques: they rely solely on experience, intellect and imagination to generate a comprehensive risk list” (Powell and Klein 1996). Formal approaches include methods such as brainstorming or Delphi techniques, which can involve numerous stakeholders. In the literature, there is no universally accepted approach to risk management. This however, is a reasonable and understandable finding in that the process of risk management will differ across various industries and projects. As Raz, Shenhar et al. (2002) state “one cannot expect that a single, universal risk management process and its supporting set of tools and techniques would be applicable to all types of projects” and “just as there are different types of projects, we should expect to see different kinds of risk management practice.”

Risk management fundamentals however remain the same and are consistent across disciplines. The literature shows similar emphasis on the most important activities in risk management namely, those identified by Rowe (1977) and Charette (1990) - the early practitioners of risk management - who outline the three main elements of risk assessment as (i) Risk Identification, (ii) Risk Estimation and (iii) Risk Evaluation.

2.1 Risk Identification

Risk identification is the reduction of descriptive uncertainty (Rowe 1977) which involves “surveying the range of potential threats” (Charette 1990). This element of risk assessment involves detecting issues which could jeopardise or threaten the success of a project (Coppendale 1995; Grey 1995). Chapman (1998) states that “the risk identification and assessment stages have the largest impact on the accuracy of any risk assessment.” As a result, risk identification is the most important stage of the risk management process. Of particular importance to ISD is the early identification of risks where “identifying and dealing with risks early in development lessens long-term costs and helps prevent software disasters” (Boehm 1991). Many software disasters could have been avoided by early identification of any high-risk aspects (Boehm 1991). It is imperative to understand that “the identification of risks carries on throughout the life of a project” (Grey 1995). In other words, risk identification is an ongoing, continuous process that requires regular screening and monitoring.

An important aspect of risk identification is categorising the risks organisations encounter. According to Coppendale (1995), “depending on the size and the complexity of the project there might be between five and fifteen” categories of risk. Categories attempt to group certain types of risk under a particular heading and in doing so “can help you find global risks that can be solved together” (Williams, Walker et al. 1997). Two dominant ways of categorising sources of risk have been identified by Frame (2002). The first way of categorising risk is by function. Schwalbe (2000) identifies these functional categories as Market, Financial and Technological. For example, market risk may be a failure of user-acceptability of a product. The second approach is to categorise risk as internal or external. For the purposes of this research, it was the author’s opinion that an analysis of
the literature in relation to internal and external sources of risk would provide a broader examination of the risk identification process.

There are many sources of risk, some of which include senior management, the client or customer, the project team, organisation of the project itself and finally laws and standards which directly impact the project (Mantel, Meredith et al. 2001). These sources can be placed into their respective categories as being either internal or external risk as follows:

<table>
<thead>
<tr>
<th>Sources of Project Risk</th>
<th>Internal</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senior Management</td>
<td>Acts of Nature</td>
<td></td>
</tr>
<tr>
<td>Project Team &amp; Management</td>
<td>The Client</td>
<td></td>
</tr>
<tr>
<td>Organisation of the Project</td>
<td>Laws and Standards</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Categories and Sources of Project Risk

The two most dominant sources of internal risk identified across the literature are Senior/Project Management and Project Team. The dominating external source of risk is the client. All of these were collectively identified by Mantel, Meredith et al. (2001) and are represented in the above table.

Every source of risk can have numerous risk factors. A risk factor is “a condition that forms a serious threat to the completion of an IT project” (Keil, Cule et al. 1998). Some internal risk factors include project conflict and resource boundaries, which can be linked to sources of project team and senior management risk respectively. There are an infinite number of risk factors (even if one was to analyse just one industry) and a study of all these was beyond the scope of this research. The following table however, shows some of the most dominant risk factors identified by Wiegers (1998), who categories these factors by sector:

<table>
<thead>
<tr>
<th>Project Sector</th>
<th>Risk Factor</th>
<th>% Of Projects at Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIS</td>
<td>Creeping User Requirements</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Excessive Schedule Pressure</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Low Quality</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Cost Overruns</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Inadequate Configuration Control</td>
<td>50</td>
</tr>
<tr>
<td>Commercial</td>
<td>Inadequate User Documentation</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Low User Satisfaction</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Excessive time to market</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Harmful competitive actions</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Litigation expense</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 2: Most common risk factors for various project types (Wiegers 1998)

On analysing the table, some direct link between sources and risk factors is evident. For example, the sources of ‘management’ risk or ‘organisation of the project itself’ could be linked to the risk factor of ‘inadequate configuration control’ due to a flaw in the project’s arrangement and organisation. However, the most interesting correlation is that of the ten risk factors listed in Table 2, at least six of these can be linked to the client as a source of project risk.

Finally, a dominating feature in recent literature deserves recognition where there is strong support among authors that an actual ‘source of risk’ can provide a ‘source of opportunity.’ Chapman and Ward (1997) state “it is only once risk is seen as a good thing people begin to look for opportunities.” Very few people would acknowledge ‘opportunity’ as being a facet of risk as naturally there are negative connotations associated with risk. Hillson (2001) however, states “the decision to encompass both opportunities and threats within a single definition of risk is a clear statement of intent, recognising that both are equally important influences over project success, and both need managing
proactively.” In a general sense the above ideas represent something we all know and understand about risk and the nature of taking gambles – people and organisations usually undertake risks with the aim of benefiting from potential opportunities (Schwalbe 2000). Taking on any form of risk can be a daunting task but as DeMarco & Lister (2003) state, “if a project has no risks, don’t do it.”

2.2 Risk Estimation

At this stage it is hoped that the project team have identified all potential risks and they can now move on to estimating those risks. Risk estimation is the reduction of measurement uncertainty (Rowe 1977) where “the values of the variables describing the system are determined, the various consequences of an event occurring are identified” and finally, “the magnitude of the risk is determined” (Charette 1990). According to Charette (1989) “the basic difficulty in estimating software engineering risks is that most prospective project risk estimates being made are unique” because there are so few comparable projects which were conducted in the past. This aspect of estimation is extremely applicable in ISD environments where there are many generic risk occurrences but very few instances of projects operating under similar circumstances. Therefore, future estimates will undoubtedly be different.

Risk estimation attempts to estimate “the chance (or probability) of potential loss” as well as “the exposure to potential loss i.e. the consequences or magnitude of the identified risks” (Charette 1989). The chance of potential loss is essentially the process of attaching a probability of occurrence to any identified risk. As Hall (1998) states “estimation is the appraisal of risk probability and consequence.” Probability is categorized as being greater than zero and less than one hundred while consequence is decided relative to cost, schedule and technical goals (Hall 1998). If an event is certain to occur it has a probability of exactly one (McManus 2004) or one hundred percent. Calculating the probability of occurrence of a risk factor means calculating “the likelihood that this risk will actually happen” (McManus 2004). According to McManus (2004) probability data should be used to compute the risk. When no actually data on probabilities exist, estimates by individuals most familiar with the project, its risk factors and overall problems are a good substitute (Mantel, Meredith et al. 2001).

According to Frame (2002) and Hall (1998) addressing the consequences of risk factors can be carried out in two ways – either by qualitative or quantitative risk analysis. Qualitative data involves subjective units such as ‘high’, ‘low’ or ‘critical’ while quantitative data uses numerical units such as ‘workdays’, ‘staff months’ or ‘monetary units’ (Smith and Merritt 2002). Informally, the literature reveals that qualitative analysis is based on subjective risk estimation while quantitative analysis is based on objective risk estimation. For example, Mantel, Meredith et al. (2001) outline how approaches that are qualitative in nature include “subjective assessments based on experience or intuition” while quantitative approaches are “bases on mathematical and statistical techniques” which are objective in nature.

2.3 Risk Evaluation

Risk evaluation is an ambiguous term. Many authors use it to describe the process of evaluating the entire risk management approach while others use it to encapsulate the ideas of risk response. For example, Marchewka (2003) states how “evaluation should consider the entire risk management process.” Charette and Rowe however, outline how risk evaluation is the final stage of the risk assessment process where “responses to the risks are anticipated” (Charette 1990). The latter perspective is the stance to which this research has taken. According to Rowe (1977), risk evaluation involves “risk aversive action, which can result in risk reduction or risk acceptance.” Risk aversive action is essentially any mission that is undertaken to control a risk (Rowe 1977). At a glance, the purpose of risk evaluation is to amalgamate the results of the risk estimation phase (Chapman and Ward 1997) and then decide the best action(s) to take. This means that risk evaluation requires input from numerous stakeholders including senior management “in proactive reduction of risk” signifying
the importance of risk action either before a project commences or long before a risk becomes uncontrollable (Thomsett 1993).

According to Charette (1989) risk evaluation consists of the following three steps:

• Establish the ‘acceptable’ level of risk
• Understand how the various risks interact (which may provide a single referral point for the project)
• Determine the action that needs to be taken with the project in relation to these risks

The acceptable level of a risk will depend on “individual propensity to take risks” (Rowe 1977), which lies in the hands of the project manager and as a result is difficult to specify. By avoiding any risk action however, we are eliminating any proactive risk management. This means that if the risk occurs, reactive approaches will be applied such as contingency planning (Frame 2002). Secondly, interacting the various risks to provide a single referral point to the project is a worthwhile exercise as “risks tend to be interrelated with a lot of overlap existing between categories” (McManus 2004). This can make the risk evaluation process more efficient as risk responses can be “grouped according to their intended effect on the risk being treated” (Hillson 2001). Finally, actions that need to be taken will be in the form of risk mitigation, risk avoidance, risk acceptance and risk transfer which are the response options available (Chapman and Ward 1997).

2.4 Incorporating Risk Management into Agile ISD

According to Remenyi (1999) traditional IS developments initially set out to avoid or reduce the changes to a system which is regarded as a huge fallacy and now “the only way forward to a higher degree of success is to identify potential changes as early as possible and to lay down suitable plans for coping with them.” This involves carrying out proactive risk management. In ISD some of the most dominant risk reduction techniques include prototyping, benchmarking or surveying (Boehm 1991) where managers tend to carry out risk action alternatives on an ad hoc basis (Powell and Klein 1996). According to Powell and Klein (1996), “prototyping and incremental development is often used to reduce project risks, for example by developing knowledge, by breaking the project into digestible bits, by reducing time between specification and delivery and by reducing the impact of change requests.” These concepts have triggered a new wave of development approaches, the most notable of these being agile methods.

The growing prevalence of agile methods has been increasingly documented with the use of such methods as XP, SCRUM, Crystal and DSDM. Organisations are using these methods primarily to assist them in responding to change more rapidly and as risk management is a proactive change mechanism, it makes sense that these two practices should be associated with one another in agile ISD literature. There are several agile methods and a detailed analysis of all these methods was beyond the scope of this research. However, in order to put agile methods into context for the reader, the author will give a brief overview of one of these methods, namely, the Dynamic Systems Development Method (DSDM). The author has chosen this as an example as it is considered to be the first “truly agile software development method” (Abrahamsson, Warsta et al. 2003). It is also a worthy example to use because the method not only focuses on systems development from a coding or technical perspective but also places emphasis on higher-level business perspectives.

2.5 Dynamic Systems Development Method (DSDM)

According to Abrahamsson, Warsta et al. (2003) the main idea behind DSDM “is that instead of fixing the amount of functionality in a product and then adjusting time and resources to reach that functionality, it is preferred to fix time and resources and then adjust the amount of functionality accordingly.” The DSDM Consortium advocates that because each organisation is different none of its practices are detailed and the method itself provides complete support over the entire life-cycle phase (Abrahamsson, Warsta et al. 2003). DSDM phases are made up of a functional prototype iteration, design prototype iteration followed by actual implementation with each iteration stage identifying,
agreeing, creating and reviewing the prototypes (Beynon-Davies and Williams 2003). However before any iteration begins, a project will always start with a feasibility and business study. This highlights that the method is not just focused on technical deliverables but it remembers to place importance on business aspects where the business should always be the driver of technological developments.

DSDM has an active community and is developed and used mainly by its Consortium members and although the method is fully available, access to white papers which outline the specific use of the method is limited to such members (Abrahamsson, Salo et al. 2002). Therefore, the extent to which DSDM incorporates risk management in practice is unknown. Empirical evidence of the method is enclosed in such white papers, which are not publicly available.

For the purpose of this research, a member of the DSDM Consortium who participated in this study, kindly provided the researcher access to information regarding DSDM’s fundamental principles for conducting risk management. Without access to these white papers the researcher would not have been able to document the following information in relation to risk management in DSDM whereby its principles are primarily built around the following concepts:

<table>
<thead>
<tr>
<th>Suitability Risk List</th>
<th>Risk Identification</th>
<th>Risk Log update</th>
<th>Monitor</th>
<th>Alert</th>
<th>Assess</th>
</tr>
</thead>
</table>

All risk management practices conducted in DSDM projects are driven by the processes outlined above. The suitability risk list determines at the outset how compatible a particular project is for the use of DSDM and outlines its associated risks. The suitability list is analysed as part of the first formal risk management activity. A risk log is then created, maintained and updated throughout the life of the project (DSDM Consortium 2002-2006). It is well known that agile methods like DSDM were mainly introduced to help combat common ISD project failures, which revolved around time or cost overruns, or a failure to meet business needs. This is also clear from what the DSDM Consortium (2002-2006) has stated whereby “systems that meet the needs of the business are delivered through the incremental and iterative approach with its continuous feedback from users” while “cost and time overruns are avoided by the use of timeboxes.”

As a result many would propose that methods like DSDM would not need to incorporate a high degree of risk management into their processes because the very reason for their existence is to reduce common risks associated with these well-known project failures. However, it appears from this literature research from the DSDM Consortium, that DSDM itself gives no less consideration to risk management than what has been documented for traditional projects. The extent, understanding and application to which these procedures are applied however, required greater investigation that took the form of a case study analysis detailed in the following sections.

### 3 RESEARCH METHODOLOGY

This research utilised a single case study to analyse the present levels of risk identification, estimation and evaluation in the well-known agile method, DSDM. The case study focused on an Irish change management consultancy firm that utilise DSDM in their commitment to delivering client value. All participants in the study also had extensive experience using traditional systems development methodologies and could therefore make accurate comparisons between the practice of risk management on both agile and traditional systems development projects.

A qualitative research approach was chosen for this study because of its greater exploratory nature and therefore its applicability to this research domain because it focuses “on gaining familiarity with the subject area” and gains “insights for more rigorous investigation at a later stage” (Collis and Hussey 2003). This was essential in light of this study because there is very little research conducted in the area of risk management in agile ISD. The researcher’s aim was to fully understand, determine and describe the existing situation of risk management in agile ISD, which goes beyond purely scientific restricting research. Therefore an interpretive stance was taken.
3.1 Data Collection

There were a variety of options available for collecting research information in this study however due to the fact that there was little pre-existing theory on the phenomena being studied (Bonoma 1985) and as the researcher needed to focus on a qualitative methodology, the data collection technique needed to “emphasise meanings and experiences related to the phenomena” under investigation (Collis and Hussey 2003). As such, for the purpose of conducting primary research the author decided to use semi-structured, personal one-to-one interviews, which were carried out with the company founder and managing director as well as four of its key consultants and analysts. Each interview had a duration of approximately one hour. The researcher felt that in choosing this approach it ensured maximum quality and reliability of data as all interviews were tape-recorded and therefore allowed for subsequent accurate analysis of the data. The semi-structured interviews ensured that the researcher maintained control over each interview, without discouraging the discussion of any valuable, additional information.

3.2 Company Background

The company involved in this case study was an Irish change management consulting firm with operations throughout Ireland. It focuses on delivering technologies to clients through the utilisation of DSDM. The managing director has over twenty five years experience of project management and has participated in many change-management initiatives varying from traditional techniques to agile development techniques. The dominating feature of this company was its great success and confidence in utilising and deploying the ideas of agile development to deliver client value. In 2006, the company were also one of the 1666 members of the global Agile Alliance.

4 FINDINGS AND ANALYSIS

The company involved in this study is an active member of the DSDM Consortium and conformed to the principles of risk management outlined by the Consortium, which was discussed previously in section two. Each interviewee expressed how the practice of managing risk can never be curtailed whether the project is agile or traditional based. In fact, one of the consultants interviewed emphasised how in their experience there is a greater level of exposure of the importance of managing risk on agile projects due to high-level stakeholder involvement and collaboration. Strong emphasis on stakeholder collaboration improves the degree of managing risk in agile methods such as DSDM. The managing director explicitly stated:

“Techniques of agile methods themselves are about reducing the risks of project failures and no matter what the nature of the agile method, a risk management framework will always be in place”

This interviewee also expressed how in their experience, due to the turbulent nature of business environments that “risk management has become progressively better because of project environments and the large amount of project failures.” Therefore, risk management procedures are gradually being considered with greater importance in ISD, regardless of the method being agile-based or otherwise. Such statements are important to consider for the remaining findings relating to risk identification, estimation and evaluation as all participants in this research have transitioned from traditional roles to agile but the extent to which their experience of risk management has progressed is relative to the progression of risk management in project environments over time. An experienced workshop facilitator and consultant with the company noted how they bring DSDM risk management practices into traditional project environments, which reinforce any risk methodologies already in place by the client. As stated by this interviewee, because risk management is incorporated into DSDM “it highlights the need for it even when the client does not have a strict risk management methodology.”
4.1 Risk Identification in DSDM

It became evident from all interviews that risk identification is an integral part of the risk management process for the successful completion of a DSDM project. The interviewees generally spoke about three main aspects of their risk identification process, which revolved around an early suitability filter; ongoing workshops and risk log updates. Their comments showed strong association with the formal practices outlined in section two. One interviewee explained that before commencing a project, the suitability filter aims to identify how appropriate DSDM is for a proposed project. The very existence of this suitability filter is a strong indication that DSDM clearly recognises the potential for associated risks when using the method on incompatible projects. Furthermore the managing director noted how even in a scoping workshop they “try and very quickly identify risks from all stakeholder perspectives” and later they would conduct more “formal risk assessment, which may be part of another workshop.” This conveys how risks are considered with importance even at the very early stages of DSDM projects.

As indicated in section two, continuous risk identification is an important aspect throughout a project’s life cycle. Interview participants predominantly referred to the risk log as their main medium for continuous monitoring and updating of risks. Again, one interviewee working as an analyst explained how “risks would be reviewed on a weekly or fortnightly basis in conjunction with status reports that are carried out. What’s outstanding on the risk log would be included in the status report.” The researcher was given access to one of the organisations risk logs which contained a risk number, category, description, colour code (for high, medium or low risks), assignment of risk, date, proposed action and follow-up control.

In section two, client risk involving scope creep was documented as being the main source of project risk across literature. This however, was not the main finding in this research where all interviewees strongly declared how “resourcing” would be their most dominant risk across projects. In particular, obtaining and retaining key personnel in a given project is the single biggest risk across all projects encountered in this case study where a key consultant stated how “resourcing availability is the classic risk we encounter because there are so many competing projects in operation across the organisation.” This risk is associated with the project team risk documented in section two. With regard to overcoming client risk, this consultant stated how:

“In a DSDM project you actually manage scope creep. There are specific DSDM techniques such as prioritisation of requirements and timeboxing, which helps you to manage it. Within each timebox you set the scope upfront and agree that you’re not going to go outside the breadth of what you have at that point.”

In terms of risk categorisation, the literature revealed two main ways of categorising risks – either by function or as being internal or external (section two). The findings in this study revealed that in DSDM risks are typically categorised according to function and “according to the project team stream that owns the risk because effectively this is how the risk gets allocated and it determines who has ownership of the risk.”

Finally, in section two the author discussed a relatively new feature in literature, which promoted aspects of opportunity in risk identification whereby a source of risk can often be a source of opportunity. When all respondents were questioned about this they expressed how they would never associate risk with opportunity in this way. However, when discussing this with the managing director a very interesting concept was put forward whereby:

“Risk management is really another way of identifying a business requirement. You can take a risk approach to develop more business requirements because mitigation of a risk is something that you actually have to do to develop the product so it becomes part of requirements.”

The researcher found this to be a very interesting and valid viewpoint in that if the resolution of an identified risk results in an additional needed requirement, a project might easily find itself embracing
an opportunity by fulfilling this requirement. However, based on this interviewee’s experience these requirements are more often a necessary feature required to make the project work instead of an innovative idea bringing about opportunity.

4.2 Risk Estimation in DSDM

The DSDM Consortium (1999-2003) outlines how the ‘level of risk’ is equal to the likelihood of its occurrence multiplied by the severity of impact, which supports literature in relation to risk estimation. The process was also evident in this case study where risks would be prioritised on this basis. An experienced workshop facilitator stated how their “main mechanism for prioritising risks would use a probability by impact approach.” In contrast however, the managing director stated that they would prioritise a risk relative to the prioritised requirement to which the risk effects. The researcher concluded that while some respondents referred to the main concepts of risk estimation, the nature to which they believe risks are prioritised was extremely dependent on each interviewee’s project role and experience. For example, the most junior team member stated how they would “prioritise risks and identify which are actual show-stoppers but estimating risk probabilities or allocating percentages to the impact on a project is something I haven’t had experience of.” Finally, the approach adopted for risk estimation in this case study showed that both quantitative and qualitative estimation techniques are used and are project dependent and vary from client to client.

4.3 Risk Evaluation in DSDM

A general consensus emerged in this study in that it is not considered feasible to carry out action for all risks because as a leading consultant stated “it just wouldn’t be practical and in some cases this may be more expensive to the project than what it would have to endure if the risk actually occurred.” However, all respondents agreed on one important point when discussing responsive actions and this involved the importance of responding to the highest prioritised risks in the same manner as they would respond to the prioritised ‘must-have’ requirements in specific timeboxes. The researcher noted something important about this level of clarity in that the principles set down by the Consortium for requirements prioritisation and timeboxing have been made applicable to the approaches taken for managing and responding to risks. As one interviewee stated:

“Procedures or actions would not be carried out for all identified risks but only those which have the appropriate priority in the same way as a ‘should-have’ requirement would not be an immediate necessity as would a ‘must-have’ requirement. When actions are taken however, they may include risk reduction, transfer or contingency planning.”

5 CONCLUSIONS & FURTHER RESEARCH

This research was based on the concept of managing risk where the researcher sought to discover the extent of risk management practices in agile ISD projects utilising DSDM. This analysis involved a decomposition of the primary elements of risk management namely risk identification, estimation and evaluation. An analysis of the literature revealed little evidence on the extent of use of risk management practices in agile methods. However, it is widely recognised that agile methods themselves were introduced to combat well-known risks associated with ISD project failures such as scope creep, cost overruns and schedule pressures. Their use of incremental development and active user involvement is an attempt to combat such risks.

Nevertheless, the findings of this research shows that the extent to which risk management is conducted in DSDM is in no way inferior to that carried out on traditional projects. In fact, DSDM shows a strong presence of risk management activities and in some instances shows more rigour for managing risk that what has been documented for traditional projects. The DSDM Consortium have specified the need for formal risk management procedures involving the use of an early suitability filter for example, which specifies project fit with DSDM as well as ongoing use of a formal risk log.
In addition, the research participants involved in this study showed a clear appreciation for the importance of managing risk on projects outlining how the insistence which DSDM places on stakeholder collaboration has a huge impact on risk management as it promotes the formal management of business risks, improves overall awareness of risk management among the project team, increases the number of risks identified throughout the project and ensures logical ownership of risks.

While the research produced many interesting findings there is much scope for further research as follows:

- Given the diversity with which methods are adopted across organizations, a large-scale quantitative study may identify more generalisable themes regarding the adoption and deployment of risk management practices across a large number of projects.
- The case adopted in this research focused solely on one agile method, namely DSDM. Given the diversity that exists across the agile method family, there is a need to examine the extent to which risk management is facilitated by other methods.
- A final potential area for further research may be a comparison between the existing state of risk management in agile project environments and traditional project environments. This form of research may produce interesting findings in relation to risk management maturity over time.

References
In-house development as an alternative for ERP adoption by SMEs: A critical case study

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0168.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Enterprise resource planning (ERP) (packaged systems), Technology adoption, Risk, Closed source software</td>
</tr>
</tbody>
</table>
IN-HOUSE DEVELOPMENT AS AN ALTERNATIVE FOR ERP ADOPTION BY SMES: A CRITICAL CASE STUDY

Poba-Nzaou, Placide, Université du Québec à Trois-Rivières, P.O. Box 500, Trois-Rivières, QC, Canada G9A 5H7, placide.poba-nzaou@uqtr.ca

Raymond, Louis, Université du Québec à Trois-Rivières, P.O. Box 500, Trois-Rivières, QC, Canada G9A 5H7, louis.raymond@uqtr.ca

Abstract

ERP systems are increasingly accessible to small and medium-sized enterprises (SMEs). If the potential benefits of these systems are significant, the same applies to the risk associated with their implementation. The majority of ERP studies relate to software packages supplied by large vendors such as SAP and Oracle and by smaller vendors; but until now, few have studied the adoption of ERP systems developed “in-house”. Furthermore, few studies have explicitly focused on minimizing the risk of these systems at the adoption or pre-implementation stage. Presenting a critical case study which analyzes the adoption of an in-house ERP by a SME in the agri-food industry, this article proposes and tests a process framework of ERP systems adoption, based upon a literature review and a conceptual framework centered on risk minimization. The study shows that 1) in-house ERP seems to represent a credible alternative for ERP adoption by SMEs, 2) to minimize risk at the adoption stage, a SME can proceed in a rather intuitive and unstructured manner, based however upon certain principles, policies and practices. The successful ERP implementation in this case indicates that it is not always necessary to resort to formalized project management in order to minimize implementation risk.

Keywords: ERP, in-house development, adoption, implementation risk, SME.

1 INTRODUCTION

Despite the fact that ERP adoption is one of the most important phenomena “in the corporate use of information technology in the 1990s” (Davenport 1998, p. 122), their adoption by small and medium-sized enterprises (SMEs) is a relatively new phenomenon (Everdingen, Hillegersberg and Waarts 2000). Recently, Light and Sawyer (2007) have found some weaknesses in ERP research with regard its theoretical foundation, also noting the scarcity of research focused on non-commercially licensed ERP software such as open source ERP and “in-house” or custom-developed ERP. In this last alternative, the organization develops its own ERP system, made-to-measure to satisfy its specific needs, with or without the assistance of a third party. The availability of new software development tools and methods (Wirfs-Brock 2008) as well as the wide adoption of open standards by developer communities make the in-house development of ERP systems feasible (Olsen and Saetre 2007). Certain characteristics of this alternative make it attractive for SMEs: greater adaptability, no vendor lock-in, property of the source code, cost comparable to the other alternatives (Olsen and Saetre op. cit.). Furthermore, most of the studies on ERP adoption tend to examine large enterprises and their findings cannot easily be extended to SMEs because of the fundamental differences between the two (Dandridge 1979).

The growing importance and risk of ERP projects, in SMEs in particular, imply that researchers must “focus on ways to improve ERP implementation” (Robey, Ross and Boudreau 2002 p. 19). With regard to risk minimization, Kliem (2000) emphasises the efficiency of risk management when it is
introduced at the earliest possible opportunity in the system life cycle. In response to the appeals stated above, and building on past research, the present study attempts to answer the following research question: *What can be done to minimize the risk of ERP system implementation in a small firm, from the adoption stage onwards?* To obtain initial insights on this question, a positivist case study was made of a small firm that has developed and implemented an in-house ERP system. The objective of the study is three-fold: a) contextualize and describe the various stages in the SME’s in-house ERP adoption process, b) identify the firm’s management practices that reduce the ERP development and implementation risk, and c) understand the determinants and effects of these practices. This research builds upon Boudreau and Robey’s (1999) theoretical framework in using four theoretical lenses to discuss the case, that is, organizational change theory (Robey et al. 2002), diffusion of innovation theory (Rogers 2003), complexity theory (Cohen 1999), and the resource-based view (Beard and Sumner 2004).

2 RESEARCH FRAMEWORK

While the notion of risk is deemed to be important for IS researchers and practitioners, there is as of yet no consensus on how to define, measure and manage IS risk; and most risk management models do not have a theoretical foundation (Alter and Sherer 2004). In the present study, based upon Barki, Rivard and Talbot’s (2001) model founded on contingency theory, it is proposed that the success of an ERP implementation will be influenced by the degree of alignment or “fit” between the SME’s risk exposure level and its risk management profile.

In reviewing the literature on ERP system adoption, we found that most of the models are without any theoretical foundation with the exception of Verville and Halingten’s (2003) model. The framework developed for the present study is thus based on this model because of its theoretical and empirical foundation and on Esteves and Pastor’s (1999) because it is somewhat complementary to the first.

A number of IT implementation studies have used Tornatsky and Fleischer's (1990) technology-organization-environment (TOE) framework to characterize the implementation context, emphasizing three groups of contextual factors: 1) characteristics of the environmental context such as external pressures from the firm's business partners, 2) characteristics of the organizational context such as the firm's structure, and resources, including managerial and entrepreneurial (in the case of SMEs) factors, given the key role played by certain individuals in the implementation process, and 3) characteristics of the technological context, including the information technologies already implemented by the firm.

Given the lack of consensus on the nature of IS risk factors and on the principal components of such factors (Alter and Scherer 2004), a review of prior studies led to the identification of six categories of risk exposure, namely organizational, business, technological, entrepreneurial or managerial, contractual and financial risk (Austin and Nolan 1999, O'Leary 2000, Caldeira and Ward 2003, Bernard, Rivard and Aubert 2002, Ariss, Raghunathan and Kunnathar 2000).

Figure 1 presents the initial research framework derived from a review of the scientific and empirical literature. The framework was initially validated by 13 academic and professional ERP experts through a two-round Delphi procedure. According to this framework, the ERP adoption process in a SME is influenced by a general context and a specific context (Ross and Vitale 2002, Kumar, Maheshwari, and Kumar 2003). The process itself can be broken down into seven stages (Verville and Halingten 2003). The dynamics of change that allow the organization to move from one stage to the next is explained by “motors” of organizational change and by the form of the change (Boudreau and Robey 1999). The choices available as a result of the process are the six alternatives proposed in the framework (Alshaw, Themistocleous and Almadami 2005), one of which is in-house ERP development. Lastly, the fit between exposure to risk and risk management profile (Barki et al. 2001) has an impact on the quality of the new system, on user satisfaction, on individual managers and employees, and on the organization (Gable, Sedera and Chan 2003).
CASE STUDY RESEARCH METHOD

The field of knowledge on ERP, while continuing to grow, had achieved a certain level of maturity (Botta-Genoulaz, Millet and Grabot 2005). Chen and Hirschheim (2004) note that when a field becomes more mature, one expects that research efforts will tend more to “theory-testing” than to “theory-building”. The present test of a research framework on the adoption of in-house ERP thus seems opportune. As the phenomenon under study is hardly separable from its context (Ross and Vitale 2000) and the adoption of in-house ERP is relatively recent and rare, a single “critical” case design was taken (Gallivan and Keil 2003). As is shown in Table 1, the firm selected for the case study, Bio-Epsilon, a small (31 employees) Canadian firm in the agri-food industry, adopted its ERP system at the end of a 6-month process.

<table>
<thead>
<tr>
<th>Date</th>
<th>Events</th>
<th>Phases</th>
</tr>
</thead>
<tbody>
<tr>
<td>October 2005</td>
<td>Decision to acquire an ERP to support the firm’s operations in the new building</td>
<td>Decision to adopt Planning</td>
</tr>
<tr>
<td>November 2005 to March 2006</td>
<td>Collection of information on ERP systems from five main sources</td>
<td>Search for information</td>
</tr>
<tr>
<td>March 2006</td>
<td>Evaluation of two ERP solutions from a de facto short list</td>
<td>Evaluation</td>
</tr>
<tr>
<td>July 2006</td>
<td>Start of ERP implementation</td>
<td>Implementation of the ERP</td>
</tr>
<tr>
<td></td>
<td>First functionalities of the ERP are put into service (end of implementation planned for December 2006)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Adoption of an in-house ERP solution at Bio-Epsilon.

This case is based on actual facts and existing firms. A transposition has been made to preserve the anonymity of these firms and the confidentiality of the data.
For purposes of internal validity in case study research, evidence was obtained from three different sources: semi-structured interviews, written documents, and a questionnaire (Dubé and Paré 2003). The main source of the data collected consisted of four interviews conducted at Bio-Epsilon’s premises in August 2006 by one of the researchers. The four “key informants” met were the owner-manager, the marketing manager, the quality control manager, and the sales manager (Patton, 1990). Each interview lasted approximately one hour and a half. The selection of these informants was based on a “snowball” sampling procedure. Saturation was achieved at the third interview (Lincoln and Guba 1985).

Similarly to Sarker and Lee (2001), an informal interview was also held with the consultant who accompanied Bio-Epsilon along the adoption process. The documents consulted were essentially promotional material on the firm’s activities and products, information on its markets, project documentation and documentation from the consulted ERP suppliers. The questionnaire was filled out immediately after the interview by both, the owner-manager and the marketing manager with the assistance of the researcher.

Data processing and analysis were based on the interactive model proposed by Miles and Huberman (2003). Prior to data collection, a list of four pre-determined categories and eleven sub-categories was established from the literature review and the initial research framework (Paré 2004). Given the flexibility inherent to the case study method (Dubé and Paré 2003), this list was modified during and after data collection and “room was made for modifications dictated by the data itself” (Lapointe and Rivard 2005, p. 470). Coding was done initially by the researcher who did the interviews, then done anew by a second researcher, following which discrepancies were resolved. Use was also made of “pattern matching” (Yin 2003) as well as the following process data analysis strategies: narrative, visual mapping, quantification, and temporal bracketing (Langley 1999).

4 RESULTS

Bio-Epsilon is a Canadian SME specialised in the commercialisation and distribution of food products originating in biological agriculture. Created in 1996 by a conglomerate on the initiative of two entrepreneurs, Bio-Epsilon was a subsidiary of Global-Delta until June 2005. That same year, for financial and fiscal reasons, the chief executive and until then minority shareholder bought back the greater part of the shares from the parent firm and thus became majority shareholder. Today, 96% of the equity is held by the management team and 4% by the ex-parent firm.

4.1 Global and specific contexts of ERP adoption at Bio-Epsilon

Bio-Epsilon has two types of customers. On one hand, food retailers specialised in the selling of food products originating in biological agriculture to which it sells finished products constitute approximately 80% of its turnover. On the other hand, it sells semi-finished products to a number of craft or industrial producers in the biological agri-food sector. From these semi-finished products, these last customers make biological food products that compete with those of Bio-Epsilon. This relationship based on simultaneous cooperation and competition (Shen 1997) is accepted by Bio-Epsilon’s owner-manager. The firm’s business environment is characterised by strong pressures with regard to the price of products sold and to ever-increasing requirements with regard to the quality of products and services. The customers’ demands manifest themselves in terms of time of delivery and availability of products.

---

2 The interview guide and the questionnaire are available from the authors upon request.
Bio-Epsilon is a firm whose commercial dependency upon its customers is rather weak, but it is rather strong upon the Global-Delta firm, which is its main supplier. The ex-parent firm provides 77\% of its supplies. Today, Bio-Epsilon processes over 30,000 order lines per year and forecasts an annual growth rate of 15\%. With such growth objectives and also with objectives of increasing the quality of services rendered to its customers, the firm decided to invest in the construction of a new building adapted to its operational needs. The acquisition of a new ERP system is part of this organisational project. This investment is accompanied by quantified objectives in terms of customer satisfaction. More precisely, these objectives should translate into fewer complaints lodged by customers and in the delivery of products within no more than 24 hours after reception of a customer’s order.

Bio-Epsilon is an organisation whose size is small and whose organisational structure is simple and decentralised. Certain important decisions are taken jointly by the owner-manager and the marketing director or on the basis of a consensus. Bio-Epsilon has 31 salaried employees, including five managers, and had a turnover of 14 million dollars in 2006. Due to its strong growth, this number of personnel has doubled in the last three years. To increase the flexibility of its operations, Bio-Epsilon out-sources the transportation and delivery of its products to three firms.

To do business, Bio-Epsilon mainly relies on a “niche” type strategy. The firm distinguishes itself from its competitors by its responsiveness and the diversity and quality of its products, whose brand is recognised, and by the quality of services rendered to customers. It also distinguishes itself by its marketing, justifying its employment of a full-time graphic artist dedicated to designing the products’ packaging and leaflets and to the firm’s marketing communication.

The owner-manager and marketing director both have university degrees and more than twenty years’ professional experience in the commercialisation of food products. The owner-manager has a positive attitude toward IT and IT management procedures. His management style is of the participative-management type.

As soon as the new building construction project started to take form, Mr. Nelson, Bio-Epsilon’s owner-manager, mandated an outside consultant to assist the firm in formalising the new work organisation and the new workstations. This project lasted for approximately 12 months.

Previously, the firm’s applications portfolio consisted of a “best of breed” ERP (Light, Holland and Wills 2001), composed of the following elements: marketing and sales, and financial management modules provided by a large software supplier, Bestofapps. It also included a specific application developed in the 4D technology environment to complement the functionalities of the Bestofapps software with regard to warehousing management. Payroll processing is out-sourced. There are no personnel dedicated to IT, but most employees use IT tools intensely. As a result, Bio-Epsilon’s level of IT use and IT management sophistication is high. Most management processes are automated through a relatively rich but poorly-integrated applications portfolio. Thus data are captured more than once in the various administrative units, generating numerous errors despite multiple redundant verifications. The information system had thus become an obstacle to the attainment of the firm’s quality objectives. One example illustrates this:

The marketing director indicates that when promotional materials are to be posted to customers, only 50\% of these appear automatically due to database inconsistency. “For example, when we want to send promotional materials, half of our customers do not appear because they are not all entered in the same way in the database. This looks trivial but the consequences are enormous. It’s of great concern because we have to manually process more than 1000 records each month.”

---

3 4D is a software publisher (integrated development environment, databases, application servers and web servers) whose products are meant to accelerate the development and deployment cycles for professional applications (www.4D.fr).
The triggering factor in the idea to adopt an ERP system lies in moving to a new building especially constructed to house Bio-Epsilon’s head office and warehouse. Initial motivations to acquire such a software package were operational in nature (Ross and Vitale 2000). Management wants to be able to localise product batches in the new three-storey warehouse in a more precise manner.

In addition to the essentially operational motivations previously cited, objectives of a strategic and technical nature also appeared. Bio-Epsilon’s management hoped that a new ERP system would also allow it to improve the quality of services rendered to customers. It also wanted to avail itself of a unique technical platform to manage all of the firm’s activities.

Eight organisational stakeholders participated in the process of identifying the needs relative to the new ERP system, as mentioned by Mr. Nelson: “7 or 8 out of 31 employees, that’s not too bad, that means a quarter, knowing that the firm’s small size allowed everyone to express themselves”. User participation was organised at two levels. First, in each department, one or two representative users were charged with gathering the users’ needs and assure the liaison between the project and the latter. Then, information and consultation meetings with the representative users were held.

During the adoption process, Bio-Epsilon followed one guiding principle: adapting the system to the organization. The principles represent the highest level of abstraction for an architectural approach composed of three levels, that is, principles, policies and practices (Colbert 2004).

### 4.2 The ERP adoption process at Bio-Epsilon

As presented in Figure 2, the ERP adoption process at Bio-Epsilon can be decomposed into six phases: the adoption decision, planning, search for information, evaluation of the proposed solutions, choice of the best solution, and negotiation.

![Figure 2. Dynamic of ERP adoption at Bio-Epsilon.](image-url)
In the first phase, the decision was jointly taken by the owner-manager and the marketing director, given the limitations of the present (legacy) information system to support Bio-Epsilon’s operations, limitations that would be accentuated when the firm moves to its new building. These two individuals were guided by one principle, that is, “the new system must be adapted to the organisation”, and by five policies, that is “no operating system other than Mac OS”, “work with an ERP supplier that can assure the system’s continuity”, “work with partners whose size is small”, “work with people we know”, and “work with a responsive supplier”.

In the second phase neither the budget nor the schedule of the next phases of the ERP adoption project were formally planned. However, management fixed an objective of deciding on the choice of an ERP system no later than March 2006 in order to dispose of sufficient time to implement it. This milestone allowed Bio-Epsilon to put the new system into service during the summer of 2006 such that it would be up-and-running by the time the firm moved into the new building, planned for November 2nd, 2006. While the amount of investment had not been formally estimated, Mr. Nelson had estimated the time required by the ERP adoption process. The budget for the acquisition of the new system was to be in the 75 000$ to 90 000$ (CAD) range.

In the third phase, the search for information was made “on the fly” in the words of the owner-manager. In total, Bio-Epsilon collected information on ERPs from five main sources:

- Contacts with the ex-parent firm, Global-Delta. This organisation had been using AlphaMIS’s ERP for the last six years. AlphaMIS is one of the three largest ERP suppliers in the world. For Mr. Nelson, while the implementation was done “without pain”, a lot of work and investment was needed as it took a number of years before the system performed adequately. He took from this experience that the implementation of an ERP system provided by a large supplier such as AlphaMIS would be a complex task for Bio-Epsilon. Nonetheless, he was attracted by the idea of adopting the same ERP as Global-Delta to benefit from the compatibility between systems supplied by the same provider, and from an online access to Global-Delta’s inventory and other information. For this reason, he then contacted AlphaMIS.

- Contacts with a large supplier. Mr. Nelson sent an email to AlphaMIS. A few days later, he was contacted by a sales representative who presented him with AlphaMIS product offer to SMEs. It came out from this discussion that Bio-Epsilon could acquire an ERP system for approximately 75 000 dollars (CAD). In concluding, the sales representative promised to send a proposal to Bio-Epsilon.

- Contacts with the informal network. Mr. Nelson’s discussions with other owner-managers of SMEs that have implemented an ERP in their firm seem to have confirmed his worries. Two points come out of these discussions, one being that the choices made when initially setting the system’s parameters are difficult to change afterwards, and the other being that the cost of implementing an ERP is quite high. In relation to this last point, certain accounts of other owner-managers made him remember his own bad experience with Bestofapps with regard to the quality of service provided.

- Contacts with SMESoft, a firm that had implemented the Bestofapps software package at Bio-Epsilon. On one hand, given the difficulties encountered in the search for information on Mac OS-compatible ERP systems, and on the other hand given the good quality of the relationship with this partner, Mr. Nelson has decided to seek advice from SMESoft. Besides, it is this firm that had always been charged with maintaining Bio-Epsilon’s applications portfolio, computer network and workstations. It had been SMESoft that had implemented the first Macintosh computer at Bio-Epsilon, eleven years before. SMESoft’s chief executive indicated that they could develop a solution to respond to Bio-Epsilon’s requirements because, as software developers, they knew how to do so.

The ERP proposed by SMESoft, a small firm, was the only solution found by both the CEO and the marketing manager to be truly attractive and Mac OS-compatible. While he was interested at the outset to work anew with a service provider and software developer that knew his firm « on the tip of its fingers » and with which he had had close collaboration in the last eleven years, Mr. Nelson
thought that it would be rather imprudent to make such an important investment simply on this basis: “So I was keen on that solution, except that I said to myself, for such an important investment, it would be irresponsible to start out only with a ‘feeling’.”

- Call to an experienced outside consultant. Being uncertain of his choice, Mr. Nelson called upon an outside consultant within his informal network to give him advice on the adequacy of the ERP system proposed by SMESoft, given Bio-Epsilon’s needs. As the consultant was employed by OpenSoft, an integrator of open source solutions, he seized this occasion to propose OpenSoft ERP, an open source software package, to Bio-Epsilon. Consequently, the mission of the consultant was transformed into assistance in the choice between two ERP solutions: SpecificSoft ERP (in-house development) and OpenSoft ERP (open source).

Another consultant from OpenSoft was given the task of preparing the integrator’s proposal to Bio-Epsilon. So, Bio-Epsilon met itself with a de facto shortlist made up of two alternatives: an in-house ERP or an open source ERP. Afterward, both systems were evaluated by the external consultant and Bio-Epsilon’s eight user representatives. Detailed specifications were prepared by the consultant and sent to each supplier along with an invitation to proceed to a demonstration of their product. The functional scope of the ERP system as defined in the specifications covered the following functions: customer relationship management, sales administration, purchasing, quality and control, warehouse management and accounting.

During the demonstration, Bio-Epsilon representatives asked questions principally related to firm’s core activities. The result of this initial demonstration favoured OpenSoft which differentiated itself from its competitor by the quality of its presentation, particularly its scenarios, and the extent of its ERP system’s functional coverage. Having noticed that a number of important questions had remained unanswered by both candidates at the end of this demonstration, it was decided to send them renewed specifications in order to have another demonstration. After the second demonstration, Bio-Epsilon’s management and user representatives as well as the consultant prepared a list of about twenty selection criteria to enlighten the decision. Both suppliers were then evaluated on the basis of these criteria and on the impression made during the demonstration. Following several discussions, a consensus emerged in favour of SpecificSoft, the in-house ERP development solution.

Prior to initiating the implementation of the new ERP system, the owner-manager engaged in negotiations with SpecificSoft to obtain a fixed price that included not only the system’s development and configuration but also the training of all system users. The contract signed between both parties granted the property of the source code to Bio-Epsilon.

The system was put in service during the summer, a more tranquil period for Bio-Epsilon in terms of customer orders. The new system functioned in parallel with the legacy system for one month after going live. Overall, users show themselves very satisfied with the ERP system adopted. They are most satisfied with the system’s user-friendliness and comprehensibility, and with the support provided by the supplier. At the organizational level, it is the integration of business processes that is perceived as being the most important impact of the new ERP system. At the individual level, it is the productivity gains generated by the system via the elimination of redundant tasks that are most valued.

5 DISCUSSION

During the ERP adoption process, Bio-Epsilon seems to have been attracted by the “shared vision” (Slevin and Pinto 1987) of its functioning in the new building that would be enabled by the new system. In fact, photos and plans were sited on the walls of most of the headquarter offices. At Bio-Epsilon, the consideration of the risk of implementation from the adoption stage was based on a reactive, informal, intuitive and incremental approach (Bili and Raymond 1993). Bio-Epsilon’s risk management profile can be described as an intuitive, fairly informal and apparently unstructured approach to risk, based on specific principles, policies and practices. An assertion of the owner-
manager illustrates this fact: “I was guided by events ‘on the fly’. I gathered information right and left. It was difficult to obtain information on MacOS-compatible ERP systems because 95% of companies use PCs.”

Such as presented in Table 2, the principles, policies and practices employed by Bio-Epsilon had an effect at different stages in the adoption process, on the envisaged alternatives and on the firm’s exposure to ERP implementation risk.

The ERP systems supplied by large vendors such as AlphaMIS were definitively eliminated from consideration, as they were perceived to be in conflict with the “ERP system is adapted to the organization” principle and the “work with a small-sized supplier” policy. The competition between the two remaining potential solutions brought to the foreground the most adequate choice for Bio-Epsilon. Furthermore, the answers to “field questions” put between both demonstrations as well as the work realized by the competitors between the two demonstrations allowed Bio-Epsilon to measure the responsiveness of every supplier. SpecificSoft’s and OpenSoft’s ERP solution were both in accordance with the previous principle and policy as well as with the “no operating system other than MacOS” policy. On the other hand, only SpecificSoft’s solution turned out to be more in accordance with the “work with a responsive supplier” and “work with people we know” policies. As an organizational metaphor can be a powerful yet subtle way of conveying meaning to results (Patton 1990), one can say that SpecificSoft was more capable than OpenSoft of overcoming the “virtual barrier” protecting the organization against ERP implementation risk, that is a barrier whose architecture was made-up of the principles, policies and practices employed by Bio-Epsilon during its ERP adoption process. Data analysis indicates that the architecture of principles policies and practices employed by Bio-Epsilon to minimize risk forms a whole, a “configuration” (Fiss 2007) whose internal consistency led to a successful implementation of the ERP system. Consequently, rather than the “profile deviation” alignment perspective posited initially, following Barki et al. (2001), it is the “gestalt” alignment perspective (Viliyath and Srinivasan 1995) that seems more appropriate to the case studied. In this perspective, equifinality is assumed in that different internally consistent
configurations of risk exposure and risk management profile may be equally effective, that is, may minimize ERP implementation risk.

According to Rogers (2003), five characteristics of innovations may affect their rate of adoption: relative advantage, compatibility, complexity, trialability, observability. Evidence from this case confirms the influence of three characteristics of the “in-house” ERP solution that turned out to be most important: greater compatibility with the firm’s processes and business model, less complexity; better trialability and greater evolutionary capability. The requirement for an evolutionary capability of the ERP system originates in Bio-Epsilon’s previous ERP experience and in the firm’s growth perspectives. Finally, diffusion of innovation theory explains neither the firm’s search for compatibility with regard to the size of organization (small size), nor with regard to a dominant trait of its strategic orientation (responsiveness) as indicated by the case data. This interpretation suggests that diffusion of innovation theory needs to be extended when applied to the adoption of a complex innovation such as an ERP system, in the context of SMEs.

According to Cole (2002), social processes constitute examples of complex phenomena. As ERP adoption is a social process in which the technology represents only one of the dimensions (Jones and Hughes 2001), it is thus a complex process. Consequently, the ERP adoption process may exhibit observable patterns of behaviour (Choi, Dooley and Rungtusanathan 2001). Data from the case indicate a pattern composed of five phases. With regard to Bio-Epsilon’s evolution from phase to phase along the adoption process, it is the shared vision of the functioning of the organization in the new building that acted as “an attractive vision of the future” (Elenkov, Judge and Wright 2005, p. 668) and urged it to look for a new ERP system. This vision was also a referred to all during the adoption process. Hence the “strange-attractor” motor stemming from complexity theory (Nutt and Backoff 1997) appears to be adequate in explaining the dynamic that moved Bio-Epsilon from one phase to another during the adoption process. The strange attractor is a configuration towards which tends the behaviour of a complex system in the long term (Capra 2005). Similarly, the “evolutionist” motor (Boudreau and Robey 1999) appears to be relevant in explaining the emergence of SpecificSoft and its ERP system as the suitable choice for Bio-Epsilon, as the result of competition during the adoption process.

As one can say that the “attractive future” of the organization envisioned by Bio-Epsilon’s CEO is made-up of a combination of experience, physical infrastructure, human capital resources, organizational resources, and business results, it appear that the resource-based view as applied to ERP (Beard and Sumner 2004) is a complementary theoretical lens for explaining what happened in this case. In summary, following Caldeira and Ward (2003), the outcome of the adoption process at Bio-Epsilon is valuable because adopting the ERP system applied to the core competitive processes of the business. It is also rare because it is based on a “custom” ERP system integrated in an enterprise vision. It is imperfectly imitable because of a personal relationship/partnership with the IT supplier for mutual long-term benefits. It is non-substitutable because of Bio-Epsilon’s continuous incremental innovation in partnership with IT suppliers to enhance its products and the services.

6 CONCLUSION

In an environment characterised by globalisation and based on knowledge, many SMEs are subjected to increased pressures with regard to competitiveness, innovation, flexibility, quality, and information processing capability. In attempting to respond to these challenges, a number of these firms have adopted an ERP system, and some of them are doing so by developing their system in-house. As a complex evolutionary phenomenon, ERP adoption is deemed by common wisdom to involve substantial risk. This study has demonstrated that it is nonetheless possible for SMEs to manage this risk, albeit through a single case. In attempting to describe and understand the dynamics of the ERP adoption process within change management and risk management perspectives, it is hoped that this study has provided a significant conceptual and practical contribution.
References


CONSUMER ACCEPTANCE OF BIOMETRICS FOR IDENTITY VERIFICATION IN FINANCIAL TRANSACTIONS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0279.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Adoption, Identity, Trust, Structural Equation Modeling</td>
</tr>
</tbody>
</table>
CONSUMER ACCEPTANCE OF BIOMETRICS FOR IDENTITY VERIFICATION IN FINANCIAL TRANSACTIONS

Breward, Michael, McMaster University, DeGroote School of Business, 1280 Main Street West, Hamilton, ON, Canada, N3C 4L4, breward@mcmaster.ca

Head, Milena, McMaster University, DeGroote School of Business, 1280 Main Street West, Hamilton, ON, Canada, N3C 4L4, headm@mcmaster.ca

Hassanein, Khaled, McMaster University, DeGroote School of Business, 1280 Main Street West, Hamilton, ON, Canada, N3C 4L4, hassank@mcmaster.ca

Abstract

Biometrics can unequivocally verify a person’s identity and, thus, have significant potential to curb the crime of identity theft, which costs individuals, organizations, and society billions of dollars annually. While financial institutions may be highly motivated to examine the feasibility of adopting biometrics for identity verification, research regarding consumer acceptance of this technology is limited. This paper proposes and empirically validates a research model for consumer acceptance of biometric identity verification for financial transactions. Additionally, conditions of voluntariness and user control are explored for their effects on user attitudes within this context. Implications for theory, practice, and future research are outlined.

Keywords: biometric verification, privacy, security, attitude, control, voluntariness, adoption, trust
1 INTRODUCTION

Biometrics refer to a host of physiological or behavioural characteristics that are used to uniquely identify or verify the identity of individuals. Despite the fact that the biometrics industry is still in its infancy, it has experienced tremendous growth in the recent past (IBG 2007). This growth has been spurred by governments in the interests of national security. This attention on the part of governments around the world has resulted in increased government funding for research and development in this area. This has inevitably lead to biometric devices becoming not only more reliable, but also less expensive. As a result, organizations that are strictly constrained by traditional cost-benefit analyses now see the viability and potential of biometrics across a variety of applications.

One possible application of biometrics is to help slow the growth of identity theft (IDT). In the United States alone, the losses attributed to identity theft were estimated to be over $50 billion dollars during 2006 (Javelin 2006); and this figure represents only the direct costs of the crime without considering the plethora of associated indirect costs. For the individual, there are the very real and often significant costs to repair their finances and credit rating, obtain new identification documents, etc. In addition, even though the average time required for an individual to repair the damage after their identity has been stolen is estimated at 40 hours (Javelin 2006), in the more extreme cases it can extend into hundreds of hours (Petouhoff 2006). There are also the associated costs of all the resources, in both the private and public sectors, required to prevent, detect, deal with, and prosecute identity theft perpetrators.

Typically, it is financial institutions (primarily banks) that absorb many of the financial losses incurred by their customers as a result of identity theft. Therefore, in an effort to curb the growth of this type of crime, financial institutions are highly motivated to examine the feasibility of adopting biometric identity verification methods. However, like most for-profit organizations, banks are extremely cognizant of the importance of maintaining customer satisfaction and, therefore, must be mindful of the degree to which customers will accept the introduction of any new technology, especially if it is unknown and/or misunderstood. Therefore, any feasibility study for using these technologies must investigate consumer attitudes towards, and their potential acceptance of, biometric verification as a method of establishing identity for financial transactions.

Although extensive research exists with respect to the technical aspects of biometric technologies (Lease 2005), research regarding consumer acceptance of biometrics is sparse. This may simply be a reflection of the relative immaturity of the biometrics market (Lease 2005). The purpose of this research is to explore the impact of several antecedents upon consumer attitudes towards using biometrics, for identity verification purposes, within the banking industry. This paper starts with an overview of biometrics and relevant literature to support a proposed research model. The model, which examines the factors that influence bank customers’ attitude toward using biometrics for financial transactions at ATMs, is validated through an empirical study. The results of this study are presented with implications for researchers and practitioners.

2 BIOMETRICS

The word biometrics is derived from the Greek words bios (life) and metrikos (measure). Biometrics is the science of measuring either physiological or behavioural characteristics (Zorkadis and Donos 2004) for the purpose of determining or verifying identity (Bolle & Connell & Pankanti & Ratha & Senior 2004, IBG 2007). Verification answers the question “Am I who I claim to be?” In this context, the system verifies the identity of that person and makes a yes/no decision based upon a one-to-one comparison between the newly scanned biometric data and a previously stored version (Jain & Ross & Prabhakar 2004). Identification, on the other hand, answers the question “Who am I?” In this case, the newly acquired biometric information from an individual is compared to all available biometric data files in the database using a one-to-many comparison process (Prabhakar & Pankati & Jain 2003).
In terms of computational requirements, verification is much easier than identification; in addition, the former is much more accurate than the latter.

Biometric measurements fall into two categories: physiological and behavioural (Zorkadis and Donos 2004). Physiological traits that have been used include fingerprints, hand vein pattern, fingernail bed, iris, retina, body odor, skin reflection, ear shape, teeth, DNA, hand geometry, palm, and face (Bolle et al. 2004, Jain et al. 2004).Behavioural measures include voice, gait, signature, keystroke dynamics, and lip motion (Bolle et al. 2004).

This paper focuses on fingerprint verification in financial transactions for several reasons. First, although fingerprint verification is less reliable than iris or retinal scans, it is more reliable than voice recognition. In addition, relative to iris and retinal scans, fingerprint verification is considerably less expensive. Third, while there may be some stigma attached to fingerprints given their association with criminals, it is also better understood by the general public relative to other biometrics. Finally, it is also a more mature technology when compared with other forms of biometrics.

3 THEORETICAL DEVELOPMENT AND RESEARCH MODEL

Ajzen (1991) defines attitude toward behaviour as “the degree to which a person has a favourable or unfavourable evaluation or appraisal of the behaviour in question” (Ajzen 1991, p.188). Within the context of technology adoption, the behaviour of interest is use of the system, which in this case is biometric identity verification to access one’s bank accounts at an automated teller machine (ATM). Consequently, attitude towards behaviour is a potential user’s affective net assessment of the costs and benefits associated with using the new technology (Morris and Venkatesh 2000). In the context of biometrics, privacy and security are primary concerns (or costs) while perceived usefulness is a primary benefit. Trust in the institution is also a relevant factor that could impact the attitude of biometric use in this context.

3.1 Privacy and Security Concerns

Flavián and Guinalíu (2006) see the concepts of privacy and security as highly interrelated. “Privacy is linked to a set of legal requirements and good practices with regard to the handling of personal data...Security refers to the technical guarantees that ensure that the legal requirements and good practices with regard to privacy will be effectively met. For example, the company may promise that the data will not be given to third parties without the consumer’s consent. Yet hackers might get hold of the data and hand them over to malefactors. His invasion or privacy can only be prevented by the use of suitable security measures” (Flavián and Guinalíu 2006, p.604). It is these types of nuances between the two concepts that make it unclear in the minds of consumer as to where privacy ends and security begins. This lack of distinction on the part of consumers may be due to their perception that it is not relevant to them and perhaps too technical (D’Hertefelt 2004). This confusion has been further exacerbated by legislators and companies (Flavián and Guinalíu 2006).

Throughout e-commerce adoption literature, the related aspects of privacy and security have garnered considerable attention (Park and Kim 2003, Flavián and Guinalíu 2006). It has been empirically established that, despite the tremendous growth in e-commerce over the recent past, it would be that much greater if privacy and/or security issues were addressed more adequately in the eyes of the consumer (Park and Kim 2003). It is estimated that, in 2000, online retail sales were reduced by $18 billion due to privacy concerns and that 92% of households with internet access do not trust online companies to keep their personal information private (FTC 2000). Over two thirds of Americans are concerned about cyber crime and hackers (McCrohan 2003). In another study, 52% of respondents simply abandoned an online purchase over privacy concerns (Ranganathan and Grandon 2002). If consumers are reluctant to provide information to e-vendors, it is reasonable to surmise that consumers will regard the request for biometric information from their financial institution with even more concern, as biometrics are extremely personal by definition.
Given the preceding discussion, perhaps the semantic distinction between privacy and security is immaterial to consumers. Therefore, perhaps treating privacy and security concerns together is a more appropriate approach. Similarly, this interdependence of related yet different elements with no overriding single concept, also suggests that the construct is formative. To test whether this was the case, a survey was conducted that asked 402 people to answer an open-ended question as to their concerns with respect to using biometrics for identity verification at ATMs. Of the 253 respondents that identified privacy and/or security concerns, almost 60% cited multiple considerations that drew from security and privacy dimensions. Based upon the results of this survey, privacy and security concerns were operationalised as a single, formative construct and used items based upon those given by Pavlou & Liang & Xue (2007). As the findings that privacy and security concerns do impact consumer use of the internet for banking and purchases, combined with the highly personal nature of biometric information, the following hypothesis is proposed:

**H1:** Individuals with a higher degree of privacy and security concerns will demonstrate a less positive attitude towards adopting biometric verification technology for accessing their bank account(s).

### 3.2 Perceived Usefulness

An individual’s perception with respect to the usefulness of technology has been demonstrated to be a key determinant in shaping an individual’s attitude towards the behaviour of adopting said technology (Davis 1989). Perceived usefulness has also been shown to influence consumer intention to adopt the Internet as a shopping channel (Gefen & Karahanna & Straub 2003). Using biometric identity verification in financial transactions holds several potential benefits to consumers including convenience, speed, and improved security. Hence, perceived usefulness should have a positive influence on consumers’ attitude towards using this technology. Therefore, the following hypothesis is proposed:

**H2:** Individuals with a higher degree of perceived usefulness will demonstrate a more positive attitude towards adopting biometric verification technology for accessing their bank account(s).

Within the context of biometrics, consumers must weigh the perceived usefulness of biometric verification against concerns they may have with respect to disclosing this type of highly personal information. Featherman and Pavlou (2003) demonstrated that consumer beliefs about risk reduce not only their intention to use an e-service, but also their perceived usefulness of this service. As surrendering any personal information exposes oneself to both potential privacy and security risks, and given the highly personal nature of biometric information, the following hypothesis is proposed:

**H3:** Individuals with a higher degree of privacy and security concerns will demonstrate a lower degree of perceived usefulness towards biometric verification technology for accessing their bank account(s).

### 3.3 Trust

Consumers will be more willing to interact with and purchase products from a vendor if they can trust that the vendor’s word can be relied upon that their vulnerabilities will not be exploited (Geyskens & Steenkamp & Scheer & Kumar 1996). For example, there is a wealth of empirical research suggesting that a heightened perception of trust positively influences consumer attitudes and intentions to purchase from a website (McKnight & Choudhury & Kacmar 2002, Gefen et al. 2003, Hassanein and Head 2007).

Considering the reluctance of consumers to share personal information with e-vendors, one would presume even more dissonance could be evoked by their financial institution’s request for biometric information as a means of identity verification. Just as trust is a salient belief when dealing with e-vendors in general, and when sharing personal information in particular, it is reasonable to suggest that trust could also play a significant role in whether or not consumers would be willing to surrender their biometric information in order to access their bank accounts. For example, trust has been shown to
increases one’s intention to share personal information with an e-vendor (McKnight et al. 2002). Thus the following hypotheses are presented:

**H4a:** Individuals with a higher degree of trust in their bank will demonstrate a more positive attitude towards adopting biometric verification technology for accessing their bank account(s).

**H4b:** Individuals with a higher degree of trust in their bank will demonstrate a lower degree of perceived privacy and security concerns with respect to using biometric identity verification technology for accessing their bank account(s).

The above discussion and hypothesis considers the dimension of trust towards the consumer’s bank, or what McKnight et al. (2002) refer to as institutional trust. Various studies have demonstrated that an antecedent to institutional trust is predisposition to trust (Dinev & Bellotto & Hart & Russo & Serra & Colautti 2006, Kim & Ferrin & Rao 2008). As the phrases imply, institutional trust is contextual and/or specific to an entity, while predisposition to trust is a person’s tendency to be generally trusting irrespective of specific contexts and/or entities. People with a higher tendency to trust in general typically demonstrate a higher degree of trust in a specific entity; conversely, people that have a lower tendency to generally trust are likely to demonstrate a lower degree of trust in a specific entity (Dinev et al. 2006, Kim et al. 2008). Also, extending the reasoning above that institutional trust may reduce perceived privacy and security concerns, it is reasonable to suppose that predisposition to trust may have the same effect. Therefore, the following hypotheses are proposed.

**H5a:** Individuals that have a greater predisposition to trust will demonstrate a higher degree of trust towards their bank.

**H5b:** Individuals that have a greater predisposition to trust will demonstrate a lower degree of perceived privacy and security concerns with respect to using biometric identity verification technology for accessing their bank account(s).

### 3.4 Personal Innovativeness in Information Technology (PIIT)

Based upon the marketing literature, Agarwal and Prasad (1998) define Personal Innovativeness in the domain of Information Technology (PIIT) “as the willingness of an individual to try out any new information technology” (p. 206). They suggest that people with higher PIIT are more likely to experiment with IT innovations, demonstrate a higher degree of comfort, confidence, and expertise in their evaluation process, and are usually first to adopt them (Featherman & Valacich & Wells 2006). In addition, early adopters of new technology tend to be more comfortable with increased levels of uncertainty and, by extension, tend to be risk takers (Rogers 1995, Agarwal and Prasad 1998). Thus, individuals with higher PIIT are more likely to take risks; and as the use of biometrics for identity verification may be perceived as risky, in the minds of consumers, due to the potential breach of privacy and/or security, the following hypothesis is proposed.

**H6:** Individuals that have a higher degree of PIIT will demonstrate a lower degree of perceived privacy and security concerns with respect to using biometric identity verification technology for accessing their bank account(s).

The proposed research model and hypotheses 1 through 6 are depicted in Figure 1 below.
3.5 The Roles of Control and Voluntariness

Control is seen as “central to the concept of privacy.” (Van Dyke & Midha & Nemati 2007 p. 68) This position is also held by Fried (1984), who states (p. 209): “Privacy is not simply an absence of information about us in the minds of others, rather it is the control we have over information about ourselves”. Information privacy may be defined as the right that individually identifiable information not be disseminated to other individuals or organizations without consent; however, in cases where personal information is held by another party, the individual should be able to exert substantial control over both the data and its use (Clarke 1999). Most consumers define privacy as “ownership and control of personal information.” (Acquisti and Grossklags 2005, p. 28) Singh (2006) sees it as a lynchpin connecting a variety of interrelated constructs. “[The] control of personal information connects security, trust, privacy, and identity.” (Singh 2006, p. 74) While there are a variety of divergent views of where control, and/or empowerment, lies in relation to privacy, research suggests that it is extremely salient in the minds of consumers when deciding whether or not to part with their information (Olivero and Lunt 2004). Within the context of biometrics, one would expect to see the same results, especially given their highly personal and sensitive nature. Therefore, the following hypotheses are proposed.

H7: Consumers with a higher degree of control over their biometric data will have a more positive attitude towards the use of biometric identity verification technology for accessing their bank account(s).

Voluntariness may be defined as “the extent to which the adoption of an innovation is perceived to be under an individual’s volitional control” (Plouffe & Hulland & Vandenberg, p.68). Voluntariness has been demonstrated to influence technology adoption intention in work environments (Venkatesh & Morris & Davis & Davis 2003) and with respect to smart card-based electronic payment systems (Plouffe et al. 2001). Extending this concept to the current context, we hypothesize that:

H8: When the use of biometric identity verification for accessing one’s bank account(s) is voluntary, consumers’ attitude towards its use will be more positive than where it is involuntary.

4 METHODOLOGY

4.1 Experimental Design and Procedure

Most biometric verification applications are not yet widely used and/or available such that they have not been experienced first hand by consumers. Therefore, scenario-based research is appropriate for this investigation (Sheng & Nah & Siau 2006). Additionally, the scenario method allows the researcher to ascertain which of the contexts examined will have a higher degree of acceptance to consumers (Bria & Gessler & Queseth & Stridh & Unbehaun & Wu & Zander 2001, Sheng et al. 2006). This has implications for both researchers and practitioners. In the case of the former, it will add to theoretical understanding and highlight those areas that should be examined further; and for the latter, it will allow for better decision making in terms of the allocation of scarce resources (Bria et al. 2001, Sheng et al. 2006).

Control and voluntariness were the two factors (each with two levels) manipulated in the scenarios presented to subjects in this study. The concept of the consumer having less control was operationalised as the bank maintaining a consumer’s complete biometric; while more control was operationalised as the bank maintaining half the biometric identifier while the consumer retained the other half of their biometric identifier on a smart card. Voluntariness was operationalised as the consumer having the option of using biometric identity verification or a debit card to gain access to their ATM, while involuntariness was operationalised as the bank using only biometric identity verification for ATM use. In order to minimize potential bias as respondents moved from one scenario to the next while answering the same questions, a between subjects design was used in which respondents were given only one of the four scenarios (Keppel 1991). In addition, respondents were
asked the questions regarding predisposition to trust, PIIT and trust in the bank prior to being given their scenario in order to establish the level of these traits independent of the scenario, thereby reducing any potential bias the latter might have on these pre-existing conditions.

4.2 Subjects

Subjects were recruited from the MBA programs of two major Canadian universities and included both fulltime and part-time students. To participate in the study, subjects had to 1) be over 18; 2) not work for a bank; 3) have a bank account; 4) live in Canada; and 5) use an ATM at least once a month. Participants were randomly assigned to the four scenario groups. Approximately 700 students were contacted which ultimately yielded 275 usable surveys. Of the 275 respondents, 160 were male and 115 were female. In terms of age, 65 were 18-24 years old, 159 were 25-34 years old, 43 were 35-44 years old, and 8 were 45-54 years old. The response across the scenarios was relatively even. Scenario 1, 2, 3 and 4 had 77, 59, 72 and 67 usable surveys respectively.

4.3 Instrument Scales and Validation

With the exception of Privacy and Security Concerns, which is a formative construct, all the other constructs in the proposed model of Figure 1 are reflective. In the case of formative constructs, indicators are considered to cause the latent variable and, as such, the direction of causality is from the indicator to the construct. In the case of reflective constructs, the reverse is true (Kim et al. 2008). This has ramifications with respect to evaluating construct quality and validity. In the case of reflective constructs, the use of principal components analysis and the examination of factor loadings, cross-loadings, average variance extracted (AVE), Cronbach’s α-values, etc., are required to evaluate convergent and discriminant validities. Such is not the case with formative constructs as their indicators do not need to correlate and are assumed not to co-vary. In fact, formative constructs must be examined for multicollinearity since, if the items are too highly correlated, then they are essentially measuring the same thing (Bollen 1989).

To evaluate the external validity of the Privacy and Security Concerns formative construct, both linear regression and a PLS model were run using a two construct Multiple Indicators, Multiple Causes (MIMIC) model (Diamantopoulos and Winklhofer 2001, Jarvis & MacKenzie & Podsakoff 2003). A linear regression was run using ten indicators for Privacy and Security Concerns (obtained from a variety of literature sources and through experts in the field) as independent variables and the mean of the Attitude items as the dependent variable. Using the suggested cutoff value of 3.3 for variance inflation factors (VIF) (Diamantopoulos and Siguaw 2006), four items of the initial ten items were eliminated. The remaining six items had VIF values ranging from 2.07 to 3.24.

A two construct MIMIC model was developed in PLS using Privacy and Security Concerns (with the six items that remained as a result of the linear regression) as the exogenous construct and Attitude as the endogenous construct. This construct demonstrated good composite reliability (0.967) as well as a significant path coefficient (β = -0.711, t-stat = 20.689) on Attitude.

Looking at the reflective constructs, PIIT, Predisposition to Trust, Trust, Perceived Usefulness and Attitude came from existing literature (Table 1), where they had been repeatedly shown to exhibit strong content validity. Convergent validity was assessed using a principal factor analysis with Varimax rotation, as shown in Table 2. Hair & Anderson & Tatham & Black (1995) suggested that an item is significant if its factor loading is greater than 0.50. Using this threshold, one item (Usefulness 4) was dropped due to high cross-loadings with other constructs. Construct reliability was also assessed using Cronbach’s α-values. As shown in Table 2, α-values ranged from 0.818 (for Trust) to 0.956 (for perceived usefulness). Nunnally (1978) recommends that the Cronbach α of a scale should be greater than 0.7 for items to be used together as a construct. Therefore, all our reflective constructs passed the test of construct reliability. Furthermore, discriminant validity is demonstrated in Table 3 as the square-roots of AVE (the diagonal values) are greater than the cross-correlations, as per Fornell and Larker’s (1981) criteria.
Table 1. Content Validity of Reflective Constructs

<table>
<thead>
<tr>
<th>Construct</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIIT</td>
<td>Agarwal and Prasad (1998) – 4 Item Scale</td>
</tr>
<tr>
<td>Predisposition to Trust</td>
<td>Cheung and Lee (2001) – 4 Item Scale</td>
</tr>
<tr>
<td>Trust</td>
<td>Gefen et al. (2003) – 5 Item Scale</td>
</tr>
<tr>
<td>Perceived Usefulness</td>
<td>Moore and Benbasat (1991) – 4 Item Scale</td>
</tr>
<tr>
<td>Attitude</td>
<td>Morris and Venkatesh (2000) – 4 Item Scale</td>
</tr>
</tbody>
</table>

Table 2. Convergent validity assessment (Principal Components Analysis and Cronbach α)

<table>
<thead>
<tr>
<th>Attitude</th>
<th>DispTrust</th>
<th>PIIT</th>
<th>Usefulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.937</td>
<td>0.895</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Attitude 0.937, DispTrust 0.895, PIIT 0.830, Usefulness 0.956

Table 3. Correlation matrix and discriminant validity assessment

<table>
<thead>
<tr>
<th>Attitude</th>
<th>DispTrust</th>
<th>PIIT</th>
<th>Trust</th>
<th>Usefulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.941</td>
<td>0.841</td>
<td>0.813</td>
<td>0.862</td>
<td>0.958</td>
</tr>
</tbody>
</table>

5 RESULTS

The model proposed in Figure 1 was tested using SmartPLS (v 2.0.M3). The results are shown in Figure 2 and summarized in Table 2. Approximately 66% of the variance in the attitude towards adopting biometric verification technology for accessing bank accounts was accounted for by the variables in the model ($R^2=0.661$). All path coefficients of the causal links in our hypothesized model are significant. These findings support hypotheses 1 through 6 at a minimum $p<0.05$ level.
Figure 2. The proposed structural model

Table 4. Hypothesis validation for structural Model

Based upon the results of an ANOVA test comparing the means of Attitude towards using biometrics for identity verification at ATMs, control does have a significant impact, \( (F = 8.328, p = .004) \) which supports H7. In the situation where an individual’s biometric information resides exclusively with the bank, the Attitude mean is 4.054 versus a mean of 4.560 when part of the information resides with the bank and part is maintained by the customer in the form of a smart card. The same cannot be said with respect to voluntariness. The results of that ANOVA test comparing the means of Attitude when the customer has a choice as to whether or not they will use biometrics as a method of identity verification versus it being made mandatory demonstrated that choice had no effect \( (F = 1.339, p = 0.248) \). The mean Attitude score with biometrics being involuntary was 4.182 versus 4.387 with it being voluntary. Therefore, H8 was not supported.

In a further analysis, an ANOVA test was run with scenario as the fixed factor and Attitude mean as the dependent variable. Interestingly, the scenario in which Attitude was the highest (mean = 4.674) was scenario 2, in which the consumer shared control of their biometric information, but participation was mandatory. Scenario 4 (shared control and voluntary) had the second highest Attitude score (mean = 4.459), scenario 3 (bank control and voluntary) was third (mean = 4.319), and scenario 1 (bank control and involuntary) was a distant fourth (mean = 3.805). In addition, the post hoc Bonferroni test shows that the attitude means between Scenario 1 and Scenario 2 are significantly different, as are the attitude means between Scenario 1 and 4.

6 CONCLUSIONS

From an academic perspective, this research makes important contributions by developing and validating a research model for consumer acceptance of biometric identity verification in financial transactions. This acceptance was also tested under various conditions (i.e. control and voluntariness) using a scenario based approach. The empirical results of this research represent an important first step in understanding consumers’ attitudes towards using biometrics as a means of identity verification in
financial transactions. Some interesting and salient findings from this study, which can spur future research and provide initial direction to practitioners, include:

- The significant negative influence of privacy and security concerns on both attitude and usefulness suggest that the threat of biometric information being compromised, either inadvertently (i.e. a security breach) or intentionally (i.e. being shared with other entities) is a top-of-mind issue for consumers. Banks wishing to employ biometric verification should target their marketing campaigns at educating consumers with respect to the superiority of biometric technologies relative to other forms of identity verification, as well as addressing any concerns about what is actually stored (e.g. encrypted version of fingerprint as opposed to actual fingerprint image).

- It appears that consumers understand the value of using biometrics for identity verification at their banks. However future research should explore if the positive effect of usefulness upon attitude outweighs the negative effect of privacy and security concerns, under varying scenarios and contexts.

- Institutional trust was found to influence attitude both directly and indirectly. While the direct impact of institutional trust on attitude was significant (p<.05), it appears to have a more indirect influence through privacy and security concerns. This mediation effect should be examined further in future research, under varying scenarios and contexts.

- As expected, an individual’s predisposition to trust and personal innovativeness in the domain of information technology had a significantly negative effect on privacy and security concerns. Hence, privacy and security concerns of biometric use for bank account verification are somewhat mitigated for individuals that are more trusting in nature and/or more daring to try new technologies. This could have implications for which segment of customers banks would choose to market their biometric verification programs to first as early adopters.

- Consumer’s attitude towards using biometric identity verification at their banks was significantly higher when they felt they had more control over their biometric data (by sharing the biometric data so that banks only kept a partial piece of their biometric). The same was not true for voluntariness. However, when voluntariness was combined with control to generate four scenarios, there was an impact. In particular, when enrolment in a biometric verification program was voluntary, the effect of control was negligible. While further investigation is needed to fully understand this phenomenon, one explanation could be that when the program is voluntary, the aspect of consumer control becomes less relevant. Alternatively, the fact that the entire program is voluntary gives the consumer de facto control in the sense that they don’t have to use it in the first place. The interaction of control and voluntariness could provide for some interesting future research that could provide practitioners with valuable insights of deploying biometric verification technologies.

As with any study, there are limitations. In this study, fingerprints were used as the method of identity verification. Given their association with the criminal element, this biometric may make consumers feel they are being equated with criminals which might have impacted their responses to our survey. That being said, there is the potential that any biometric could result in similar misgivings (e.g. people may worry about the long term medical implications of iris and retinal scans). The sample for this study consisted of graduate business students which limits generalizability. Also, this study was conducted in Canada where biometrics use in consumer applications is virtually non-existent and there are no national identity cards. It would be worthwhile doing an international study that varied these two aspects as some countries do have very strict national identity card programs and/or use biometrics for various government and private sector applications.
References


A EUROPEAN STUDY OF E-BUSINESS MATURITY AND ICT-BENEFITS: IS THERE A CONDITIONAL RELATIONSHIP?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0641.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-commerce (B2B / B2C / B2G / G2C), Empirical study, Adoption, Perceptions</td>
</tr>
</tbody>
</table>
A EUROPEAN STUDY OF E-BUSINESS MATURITY AND ICT-BENEFITS: IS THERE A CONDITIONAL RELATIONSHIP?

Batenburg, Ronald, Institute of Information and Computing Science, Utrecht University, Padualaan 14, Utrecht, The Netherlands, r.s.batenburg@cs.uu.nl
Constantiou, Ioanna, Department of Informatics, Copenhagen Business School, Howitzvej 60, # 5.12, DK 2000 Frederiksberg, Denmark, ic.inf@cbs.dk

Abstract

The present study explores the relationship between the e-business maturity and the perceived benefits from (Information and Communication Technologies) ICT at the firm level. We aim to debunk this relationship in terms of its strength and stability, and to explore the conditions which may influence it. Taking an economic approach, we hypothesise that the relation between e-business maturity and perceived benefits from ICT adoption will be influenced the costs of intra-organisational adaptations due to ICT. Using data from an European survey on ICT adoption and e-business maturity of the firms (N=7,072) we show that the correlation between the e-business maturity and the perceived benefits from ICT adoption is indeed positive, significant and stable over countries, industries, firm size and age. Further, the findings confirm the hypothesis that intra-organisational adaptations due to ICT moderate the positive correlation between a firms’ e-business maturity and perceived benefits from ICT.

Keywords: e-Business maturity, ICT adoption, perceived benefits, European survey.

1 INTRODUCTION

Recent international studies reveal that European countries are in the highest ranks of e-readiness and ICT growth rates (Economist Intelligence Unit, 2006; Economist Intelligence Unit, 2008; EITO, 2007). Still, the adoption of ICT within Europe strongly varies between countries and sectors. Over the last decade, the European and national policies particularly focus on stimulating ICT adoption by the small and medium sized enterprises, the SMEs (Castaings and Tarantola, 2008; Castaings et al., 2007; European Commission, 2002; Nardo et al., 2004; Renner et al., 2008). This sector structurally lags behind the ICT diffusion rates of large firms. Although from a scale perspective this might not seem a problem *per se*, the more fundamental problem observed is the low awareness, or even mistrust, of SME firms about the potential benefits of IT (Lucchetti and Sterlacchini, 2004). The problem, an over-critical and sceptical attitude towards the potential benefits of ICT, is not only a matter by SME firms. It actually draws back to the many uncertainties when investing in ICT, both at the costs and benefits side. Although the ‘productivity paradox’ (Solow) and ‘IT doesn’t matter’ (Carr) discussions have passed, still the number of studies that report failure of the ICT adoption and implementations remain large (cf. Ward and Peppard, 2003). Quite ironically, the national governments that firmly stimulate the uptake of ICT, represent organizations that particularly seem to fail in getting value from ICT. The Dutch government for example, struggle with several public examples of ICT disasters and mis-planning (Court of Audit, 2008).

So from both a policy and scientific perspective, a key challenge within area of IS adoption and uptake is to improve the understanding of how managers actually *perceive* the benefits from ICT adoption, and how this relates to their actual investments in ICT, e.g. e-business applications. It seems obvious that the firms which have invested much in ICT may perceive more benefits from these investments than the others, but this relationship is not trivial. There are firms which might have ‘under-invested’
in ICT, while there are other firms which might have over-invested in ICT. In the context of IS adoption research, the current market observations raise some questions about the nature of the relationship between the investment (i.e. ICT maturity) and the perceived benefits for the different types of firms.

The present study builds upon the increasing knowledge about organisational adoption of ICT (for a thorough review see Venkatesh et al., 2003), and economic analysis of ICT in the organisational context (Zhu and Kraemer, 2005). Motivated by these two perspectives, we conduct a comparative analysis of European firms’ perceptions of benefits and impacts of ICT and their e-business maturity. Empirical data from the e-Business W@tch project, a cross-national survey sponsored by the European Commission among 7,000 firms from 10 different EU-countries, are used. The main research objective is addressed by two research questions:

- How do the perceived benefits of ICT adoption relate the e-business maturity of the different types of European firms?
- How does intra-organisational impact of ICT affect the relation between the perceived benefits of ICT and the e-business maturity?

In order to answer these research questions we develop hypothesis and measurements that enable us to describe and explain the European landscape of e-business maturity. The study contributes to the IS adoption research by shedding some light on the relationship between perceived benefit of ICT and e-business maturity, while exploring potential differences in the relationship for firms from different countries, industries as well as firms of different size and age. The study results indicate a positive relationship between the e-business maturity and perceived benefits from ICT adoption which is moderated by the intra-organisational adaptations due to ICT. The positive relationship is stable across firms from different countries and industries and despite their different size and age.

The paper is structured as follows. The next section presents a brief literature review of the adoption theories and models and the economic elements used in the proposed research. Then the data and measurements of the study are described. The following section presents results from the bivariate and the multivariate analysis. The paper concludes underlining the main findings, limitations and proposing some future research directions.

2 THEORETICAL CONTEXT

2.1 Some basic theoretical insights on the adoption of ICT

One of the most widely used theories in the Information Systems, which attempts to explain the reasons behind innovation adoption, is diffusion of innovation (DoI) (Rogers, 1995; 2003). DoI (1995) has been used to study the adoption of many computer-related technologies. The model contains four main elements: innovation, communication channels, time, and the social system. Rogers’ model is prominent in addressing the role of change agents, individuals who influence clients’ innovation-decisions in a direction deemed desirable by the change agency (2003).

Since e-business systems involve high costs for the organisation, the adoption decision is an authority innovation-decision type. This means that the decision to adopt or reject the system is made by a relatively few individuals in the organisation who possess status, power, or technical expertise (Rogers, 1995). The top management in the organisation is mostly involved in the selection of e-business systems and in charge of providing the resources for its implementation. They delegate other decision-making tasks to other members of the organisation. Besides, they can influence the way the rest of the organisation views the system and its changes.

According to Rogers (2003), and other rational diffusion theorists (Agarwal and Prasad, 1997; Moore and Benbasat, 1991), there are certain characteristics of innovations which affect their rate of adoption. These characteristics can be applied to e-business. First, the relative advantage of e-business solutions relates to cost savings in time and effort by improving business processes. Second,
compatibility of e-business solutions may relate to synchronization of data with suppliers and customers and increases the adoption probabilities. Third, e-business solutions may be characterised by complexity which decreases the adoption probability because of the high cost of implementation and customisation. Fourth, e-business solutions have a very limited trialability since they can only be seen in other adopter organisations, and this in turn decreases the adoption probability. Finally, observability is low for e-business solutions and this decreases the adoption probability. Most observations of e-business come from indirect paths such as media and hearsay. These sources usually give evaluations of e-business which are used in the decision making process.

Diffusion of innovation theory has been criticized for not taking into account the particularities of complex information technologies (Lyytinen and Damsgaard, 2001). The different approaches in the study of ICT diffusion (such as Baskerville and Pries-Heje, 2001; Cooper and Zmud, 1990; Edquist, 1997) seem to take a narrow perspective while emphasising specific topics of interest with no single theory being able to explain the particularities of the different technologies (Jones and Myers, 2001). Besides, researchers used theories from social psychology to investigate organisational adoption of the innovation. The most commonly used theories are Technology Acceptance Model (TAM) (Davis, 1989), the Decomposed Theory of Planned Behaviour (Taylor and Todd, 1995) and the Unified Theory of the Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003). These models aim to explore the user attitudes or intentions, and thus technology adoption and use through various social, personal and technical constructs. However, they do not investigate the value elements affecting the adoption decision process or the economic implications from ICT adoption to the organisation.

The present study focuses on e-business applications and it uses theoretical insights from industrial economics to get further insight into ICT diffusion and the perceived value of ICT (Wilkins and Swatman, 2006). In particular, it builds upon the notion of relative advantage as postulated by Rogers and it investigates the firms’ perceptions of the benefits from ICT adoption and the relationship with e-business maturity.

2.2 An economic approach to ICT adoption and its perceived benefits

Most of the research on the adoption of an inter-organisational system (IOS) focuses on the expected operational benefits (e.g. Zhu and Kraemer, 2002). The firm using the Internet infrastructure to connect with other partners is expected to increase its operation efficiency, which in turn leads to more efficient information processing and a decrease in the transaction costs, search costs and other direct and indirect costs involved. These costs savings have been depicted as some of the expected benefits from the adoption of EDI systems as well (e.g. Zhu and Kraemer, 2002). Recently this theoretical discussion has been extended to include open standards IOS. In such case, additional benefits for the firm include the ability to find new partners and to exchange information, reaching new customers and expanding to new markets as well as the improvement of the value chain co-ordination (Zhu et al., 2006b).

The adoption of ICT products also involves network-related value (Kauffman et al., 2000). Research on industrial economics has highlighted the importance of increasing returns in the adoption of a new system (Arthur, 1989; Davis, 1992). In particular, there are positive network effects when the user’s value from the use of the system increases as the number of users at the same or compatible systems’ users increases, since the firm can exchange information with them (Katz and Shapiro, 1994). Network effects may accelerate or impede the diffusion process. For example a user of a legacy system, enjoying the benefits of network effects by exchanging data with other firms which adopted the old system, would expect to be able to maintain those benefits when using a new system. The adoption of a new incompatible technology may limit network effects for users of the old technology (Farrell and Saloner, 1986). Further, the compatibility between systems shifts the focus of attention from the overall services package, which includes the network size, to specific component cost and
performance characteristics (Economides, 1989; Katz and Shapiro, 1994; Matutes and Regibeau, 1988).

As firms gradually invest in e-business applications and they climb the e-business maturity ladder, they may increasingly experience the advantages and benefits from ICT adoption. Eventually, it has become a basic notion that it takes time to actually ‘see’ the benefits of ICT – especially with major ICT-investments such as ERP (cf. Poston and Grabski 2001). We expect that the same holds for e-business as a next wave in terms of ICT maturity (Ward and Peppard 2003). According to this notion the first hypothesis we propose for the present study is:

$H_1$: There is a positive relationship between firm’s perceived benefits from ICT and the level of e-business maturity.

ICT adoption may change business activities and it may create new opportunities or new needs for the firm. This in turn generates significant costs to the firm. Apart from the financial costs involved in the investment decision for the new system (e.g. Zhu et al., 2006a), the adoption costs are related to the managerial complexity and the risks involved for the firm changing processes when adopting the new system. We can refer to numerous studies that report about misinvestments because of underestimation of the organisational consequences of ICT (cf. for ERP: Hong and Kim, 2002; Rajagopal, 2002).

In a classic study, Brynjolfsson and Hitt (1995) show that firms only benefit from IT investments if they align these investments with their organisational structure. Specifically, they demonstrate that the firms which invested in computers and they decentralised organisational structure at the same time, outperformed other firms in terms of their market value. In line with these findings an extensive stream of literature on business/IT alignment emerged in the nineties, highlighting the importance of mutually adjusting the organisational/business domain and the IT domain (Henderson and Venkatraman, 1993; Luftman and Brier, 1999; Luftman and Kempaiah, 2007; Papp, 1998; Papp et al., 1995). Following the basic argument of Brynjolfsson and Hitt (1995), we expect that the relation between the perceived ICT benefits and the e-business maturity will be at least ‘blurred’ if firms experience intra-organisational changes due to the adoption of ICT. Intra-organisational adaptation may put a hold on the potential benefits of ICT. The firms that grow in ICT or e-business maturity without intra-organisational changes are likely to be those who had their processes and structures already in place. These firms may benefit relatively more and/or earlier from ICT, compared to those that are forced to execute significant intra-organisational adaptation. Hence the second hypothesis of the present study is:

$H_2$: The positive relationship between the perceived benefits and e-business maturity is negatively influenced by the perceptions of the intra-organisational impact of ICT

3 RESEARCH APPROACH: DATA AND MEASUREMENTS

3.1 The dataset

The data used for this study were collected by Empirica that coordinates the e-Business W@tch project of which a number of European wide surveys is an important part (Selhofer et al., 2008). The data collection was based on telephone (CATI) interviews with IT managers employed at firms from 15 different sectors. The e-Business W@tch survey project started in 2002 in the 15 EU member states at that time. New EU member states were included in later waves along with additional questions. Besides, the survey expanded to include additional industrial sectors.

Field work of the E-Business W@tch project was executed by national market research organisations. The standard procedure for each country was to sample enterprises that “were active at the national territory of the country”, “have their primary business activity in one of the sectors specified by NACE categories” and “can be defined as a business organisation of one or more establishments comprised as
one legal unit” (European Commission, 2008). Each national sample was drawn from “acknowledged business directories and databases” and stratified by economic activity (one or more of the NACE-defined economic sectors) and employee size. Size was of specific importance for stratification because the e-Business W@tch survey particularly aimed to estimate the accurate level of usage and adoption of IT for countries and industries. The company size was strongly correlated with the IT investments because of scale and scope of activities. In practice, three employee size strata were defined for each country and industry: to include a share of at least 10% of large companies (250 employees or more), 30% of medium sized enterprises (50-249 employees) and 25% of small enterprises (10-49 employees). Besides, the so-called “micro enterprises” with less than 10 employees were also included (by 35%).

This study used the 2006 e-business w@tch dataset. As country and sector are important variables for the analysis. Thus, it was ensured that the sectors surveyed in the 2006 wave were equally represented in each country. Besides, firms with 10 employees were selected. This resulted in a pre-selection (see Table 1) of 7,237 firms, from 10 different countries and 10 different sectors (i.e. Food = Food and beverages, Foot = Footwear, Pulp = Pulp and Paper, ICT = ICT Manufacturing, Elec = Consumer electronics, Ship = Shipbuilding and repair, Cons = Construction, Tour = Tourism, Tele = Telecommunications, Hosp = Hospital activities).

<table>
<thead>
<tr>
<th></th>
<th>Food</th>
<th>Foot</th>
<th>Pulp</th>
<th>ICT</th>
<th>Elect</th>
<th>Ship</th>
<th>Cons</th>
<th>Tour</th>
<th>Tele</th>
<th>Hosp</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>78</td>
<td>26</td>
<td>132</td>
<td>190</td>
<td>20</td>
<td>8</td>
<td>75</td>
<td>70</td>
<td>72</td>
<td>80</td>
<td>751</td>
</tr>
<tr>
<td>Germany</td>
<td>53</td>
<td>68</td>
<td>163</td>
<td>169</td>
<td>66</td>
<td>15</td>
<td>51</td>
<td>54</td>
<td>60</td>
<td>101</td>
<td>800</td>
</tr>
<tr>
<td>Italy</td>
<td>50</td>
<td>200</td>
<td>85</td>
<td>182</td>
<td>30</td>
<td>21</td>
<td>50</td>
<td>48</td>
<td>50</td>
<td>40</td>
<td>756</td>
</tr>
<tr>
<td>Poland</td>
<td>50</td>
<td>135</td>
<td>75</td>
<td>76</td>
<td>141</td>
<td>3</td>
<td>50</td>
<td>50</td>
<td>75</td>
<td>97</td>
<td>752</td>
</tr>
<tr>
<td>Spain</td>
<td>49</td>
<td>181</td>
<td>117</td>
<td>132</td>
<td>17</td>
<td>23</td>
<td>49</td>
<td>46</td>
<td>103</td>
<td>37</td>
<td>754</td>
</tr>
<tr>
<td>UK</td>
<td>59</td>
<td>20</td>
<td>140</td>
<td>167</td>
<td>59</td>
<td>8</td>
<td>59</td>
<td>57</td>
<td>147</td>
<td>34</td>
<td>750</td>
</tr>
<tr>
<td>Czech Rep.</td>
<td>74</td>
<td>70</td>
<td>105</td>
<td>99</td>
<td>130</td>
<td>2</td>
<td>75</td>
<td>75</td>
<td>70</td>
<td>50</td>
<td>750</td>
</tr>
<tr>
<td>Finland</td>
<td>149</td>
<td>18</td>
<td>66</td>
<td>104</td>
<td>9</td>
<td>4</td>
<td>141</td>
<td>134</td>
<td>95</td>
<td>32</td>
<td>752</td>
</tr>
<tr>
<td>Hungary</td>
<td>153</td>
<td>40</td>
<td>50</td>
<td>95</td>
<td>19</td>
<td>2</td>
<td>152</td>
<td>141</td>
<td>60</td>
<td>60</td>
<td>772</td>
</tr>
<tr>
<td>Netherlands</td>
<td>60</td>
<td>11</td>
<td>31</td>
<td>63</td>
<td>16</td>
<td>12</td>
<td>52</td>
<td>50</td>
<td>97</td>
<td>8</td>
<td>400</td>
</tr>
<tr>
<td>Total</td>
<td>775</td>
<td>769</td>
<td>964</td>
<td>1,277</td>
<td>507</td>
<td>98</td>
<td>754</td>
<td>725</td>
<td>829</td>
<td>539</td>
<td>7,237</td>
</tr>
</tbody>
</table>

Table 1. Country and industry composition of the 2006 e-Business W@tch dataset used for this study.

According to Table 1, the spectrum of countries and industries seems to be sufficiently diverse to cover different national level of economic development, geographical spread, including a relevant spread of manufacturing versus service-oriented industries.

3.2 The measurements

3.2.1 E-business maturity

From the e-Business W@tch survey we selected 11 questions which addressed the use of ICT (web) applications to support inter-organisational (B2B) activities. The questions explored the inter-firm collaboration in general, or specific processes such as invoicing. The response scale was categorical including “Yes”, “No” or “Don’t know”. The latter category appeared in a small number of the cases (less than 2%), which were removed from the dataset. The 11 dichotomous variables were analysed on their inter-correlations, but as the different e-business applications were not logically interrelated, (e.g. firms might invest in one type of e-business without investing in another type) we did not aim at exploring the scalability or a latent factor model for this set of 11 dummy variables.
“Does your company use online technologies or specific IT solutions”

<table>
<thead>
<tr>
<th>Indicator</th>
<th>% “yes”</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>Factor loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. to collaborate with business partners to forecast product or service demand</td>
<td>16.6</td>
<td>.510</td>
<td>.219</td>
<td>.263</td>
<td>.228</td>
<td>.179</td>
<td>.186</td>
<td>.151</td>
<td>.223</td>
<td>.206</td>
<td>.182</td>
<td>.587</td>
</tr>
<tr>
<td>2. to collaborate with business partners in the design of new products or services</td>
<td>14.7</td>
<td>1</td>
<td>.180</td>
<td>.228</td>
<td>.236</td>
<td>.177</td>
<td>.176</td>
<td>.138</td>
<td>.214</td>
<td>.192</td>
<td>.158</td>
<td>.558</td>
</tr>
<tr>
<td>3. to send e-invoices to customers in the public sector</td>
<td>16.6</td>
<td>1</td>
<td>.602</td>
<td>.426</td>
<td>.164</td>
<td>.173</td>
<td>.111</td>
<td>.163</td>
<td>.155</td>
<td>.207</td>
<td>.630</td>
<td></td>
</tr>
<tr>
<td>4. to send e-invoices to customers in the private sector</td>
<td>17.6</td>
<td>1</td>
<td>.498</td>
<td>.163</td>
<td>.197</td>
<td>.139</td>
<td>.198</td>
<td>.183</td>
<td>.212</td>
<td>.687</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. to receive e-invoices from suppliers</td>
<td>20.4</td>
<td>1</td>
<td>.205</td>
<td>.200</td>
<td>.151</td>
<td>.198</td>
<td>.193</td>
<td>.150</td>
<td>.639</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. to place orders for goods or services online</td>
<td>60.9</td>
<td>1</td>
<td>.140</td>
<td>.218</td>
<td>.165</td>
<td>.157</td>
<td>.118</td>
<td>.422</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. to support the selection of suppliers or procurement processes</td>
<td>12.5</td>
<td>1</td>
<td>.097</td>
<td>.286</td>
<td>.190</td>
<td>.112</td>
<td>.370</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. to order goods or book services online</td>
<td>29.2</td>
<td>1</td>
<td>.215</td>
<td>.166</td>
<td>.092</td>
<td>.497</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. to support marketing or sales processes</td>
<td>14.0</td>
<td>1</td>
<td>.220</td>
<td>.145</td>
<td>.497</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. to link to the ICT system of suppliers</td>
<td>11.1</td>
<td>1</td>
<td>.383</td>
<td>.498</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. to link to the ICT system of customers</td>
<td>11.3</td>
<td>1</td>
<td>.449</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Measurement of the E-business maturity concept: frequencies, inter-correlations and factor loadings of 11 indicators (N=7,237)

We added the scores of the 11 indicators to create one construct, which we labelled ‘e-business maturity’. Factor loadings from the PCA were used to allocate weights for the 11 indicators. According to Table 2, most of the e-business application were used by a minority of the firms (‘yes’ varies between 11% and 29%), except from the use of online ordering. This means that the constructed factor was skewed to the left-hand side of the distribution as nearly half of the firms answered ‘yes’ to none or only one of the 11 indicators (skewness=1.38, kurtosis=1.49).

3.2.2 Perceived ICT benefits

The e-Business w@tch questionnaire included questions asking the respondents’ perceptions of ICT benefits for their business on six different performance areas: (1) revenue growth, (2) business processes efficiency, (3) internal work organisation, (4) procurement costs of supply goods, (5) products and services quality, (6) customer service and (7) productivity. We considered these items as
indicators of one latent concept, and we explored whether they could be aggregated to a single e-business maturity scale. Principal Component Analysis (PCA) was first applied to test construct validity. The Kaiser-Meyer-Olkin test and the Barlett’s Test of Sphericity were used to determine whether the sample size was sufficient relative to the number of items, and the correlation matrix was not an identity matrix. Both tests were passed (KMO statistic is 0.894, Bartlett’s Test is significant, p <.001). As a result of the PCA, the one factor solution holds an eigenvalue of 3.75, while all factor loadings were above 0.50 and all inter-correlations were significant at the 1%-level (see Table 3 below). In addition, the Cronbach’s Alpha over the 11-item set, by 0.86, supported scale reliability (Nunnally and Burnstien, 1994).

<table>
<thead>
<tr>
<th>“All in all (…) Please tell us for each of the following areas whether ICT has a positive, negative, or no influence”</th>
<th>% “positive”</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Factor loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Revenue growth</td>
<td>43.2</td>
<td>.505</td>
<td>.375</td>
<td>.383</td>
<td>.433</td>
<td>.464</td>
<td>.516</td>
<td>.718</td>
</tr>
<tr>
<td>2. Business process efficiency</td>
<td>58.1</td>
<td>1</td>
<td>.563</td>
<td>.388</td>
<td>.409</td>
<td>.491</td>
<td>.566</td>
<td>.774</td>
</tr>
<tr>
<td>3. internal work organisation</td>
<td>55.7</td>
<td>1</td>
<td>.367</td>
<td>.394</td>
<td>.458</td>
<td>.527</td>
<td>.721</td>
<td></td>
</tr>
<tr>
<td>4. procurement costs of supply goods</td>
<td>40.2</td>
<td>1</td>
<td>.422</td>
<td>.388</td>
<td>.418</td>
<td>.641</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. product/service quality</td>
<td>38.8</td>
<td>1</td>
<td>.515</td>
<td>.489</td>
<td>.712</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. customer service</td>
<td>53.7</td>
<td>1</td>
<td>.515</td>
<td>.752</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. productivity</td>
<td>53.4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 3. Measurement of the ICT benefits concept: frequencies, inter-correlations and factor loadings of 7 indicators (N=7,237)*

3.2.3 Intra-organisational impact of ICT

The respondents were asked about the importance of ICT on four organisational aspects of their company: (1) the organisational structure, (2) task and job descriptions, (3) education and training of employees, and (4) outsourcing decisions. Categorical responses of “yes” and “no” were collected and analysed; while responses in the category “don’t know” were excluded. Again PCA, correlation and scale analysis were executed to explore whether these dichotomous items could be aggregated to one factor. Based the results (eigenvalue of the 1-factor model is 2.42, Cronbach’s Alpha is 0.78) a sum score factor was subsequently constructed (see Table 4).

<table>
<thead>
<tr>
<th>“Would you say that ICT had an important influence on the following organisational aspects in your company …”</th>
<th>% “yes”</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Factor loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The organisational structure</td>
<td>35.2</td>
<td>.592</td>
<td>.424</td>
<td>.411</td>
<td>.787</td>
</tr>
<tr>
<td>2. Task and job descriptions</td>
<td>35.2</td>
<td>1</td>
<td>.535</td>
<td>.440</td>
<td>.838</td>
</tr>
<tr>
<td>3. Education and training of employees</td>
<td>36.5</td>
<td>1</td>
<td>.418</td>
<td>.764</td>
<td></td>
</tr>
<tr>
<td>4. Outsourcing decisions</td>
<td>26.1</td>
<td>1</td>
<td></td>
<td></td>
<td>.715</td>
</tr>
</tbody>
</table>

*Table 4. Measurement of the intra-organisational impact of ICT concept: frequencies, inter-correlations and factor loadings of 4 indicators (N=7,237)*

3.2.4 Control variables

Control variables were treated as background characteristics of a firm which might potentially influenced either e-business maturity, or perceived ICT-benefits, or both. The control variables were:
- Industry type including 9 categories or branches (see Table 1);
- Country of firm’s location including the 9 countries represented in the dataset (see Table 1);
- Firm’s size, measured by the number of employees

The size indicated the ‘scale’ of a firm, which might be positively correlated with e-business maturity. The firm’s age might be correlated to e-business maturity, because the ‘older’ organisations might be restricted by earlier investments in pre-Internet technologies, and they probably employed older and less e-skilled employees. With regard to the relationship between industry and country on one hand, and e-business adoption on the other e-Business W@tch reports (European Commission, 2008) revealed Scandinavian countries and the UK were front-runners in e-business application. This might be related to the cultural differences and the public policies developed to stimulate technology and innovation. However, the effects of culture on the ICT adoption should be investigated by cross-national analyses (e.g. Miller et al., 2006; Van Everdingen and Waarts, 2003). Besides, industries such as the ICT manufacturing and telecommunications were known as early adopters of e-business because they involved products and services which were inherent to their own business domains (European Commission, 2008).

4 ANALYSIS: RESULTS AND FINDINGS

We first explore the relationship between e-business maturity and perceived ICT-benefits to test the first hypothesis (H1). Table 6 shows the (Pearson) correlations, controlling for country and industry by split sampling, and controlling for size and age by calculating partial correlations.

<table>
<thead>
<tr>
<th></th>
<th>Pearson correlation</th>
<th>Partial correlation, controlled for size</th>
<th>Partial correlation, controlled for age</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample</td>
<td>.40</td>
<td>.38</td>
<td>.40</td>
<td>7,237</td>
</tr>
<tr>
<td>France</td>
<td>.40</td>
<td>.35</td>
<td>.39</td>
<td>751</td>
</tr>
<tr>
<td>Germany</td>
<td>.44</td>
<td>.42</td>
<td>.44</td>
<td>800</td>
</tr>
<tr>
<td>Italy</td>
<td>.32</td>
<td>.28</td>
<td>.33</td>
<td>756</td>
</tr>
<tr>
<td>Poland</td>
<td>.43</td>
<td>.42</td>
<td>.43</td>
<td>752</td>
</tr>
<tr>
<td>Spain</td>
<td>.42</td>
<td>.41</td>
<td>.44</td>
<td>754</td>
</tr>
<tr>
<td>UK</td>
<td>.44</td>
<td>.43</td>
<td>.44</td>
<td>750</td>
</tr>
<tr>
<td>Czech Rep.</td>
<td>.36</td>
<td>.32</td>
<td>.36</td>
<td>750</td>
</tr>
<tr>
<td>Finland</td>
<td>.41</td>
<td>.35</td>
<td>.41</td>
<td>752</td>
</tr>
<tr>
<td>Hungary</td>
<td>.36</td>
<td>.36</td>
<td>.37</td>
<td>772</td>
</tr>
<tr>
<td>Netherlands</td>
<td>.50</td>
<td>.47</td>
<td>.50</td>
<td>400</td>
</tr>
<tr>
<td>Food and beverages</td>
<td>.39</td>
<td>.33</td>
<td>.39</td>
<td>775</td>
</tr>
<tr>
<td>Footwear</td>
<td>.36</td>
<td>.34</td>
<td>.36</td>
<td>769</td>
</tr>
<tr>
<td>Pulp and paper</td>
<td>.36</td>
<td>.32</td>
<td>.36</td>
<td>964</td>
</tr>
<tr>
<td>ICT Manufacturing</td>
<td>.36</td>
<td>.33</td>
<td>.36</td>
<td>1,277</td>
</tr>
<tr>
<td>Consumer electronics</td>
<td>.35</td>
<td>.33</td>
<td>.35</td>
<td>507</td>
</tr>
<tr>
<td>Shipbuilding and repair</td>
<td>.38</td>
<td>.36</td>
<td>.39</td>
<td>98</td>
</tr>
<tr>
<td>Construction</td>
<td>.37</td>
<td>.35</td>
<td>.37</td>
<td>754</td>
</tr>
<tr>
<td>Tourism</td>
<td>.45</td>
<td>.40</td>
<td>.45</td>
<td>725</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>.38</td>
<td>.36</td>
<td>.38</td>
<td>829</td>
</tr>
<tr>
<td>Hospital activities</td>
<td>.21</td>
<td>.16</td>
<td>.22</td>
<td>539</td>
</tr>
</tbody>
</table>

Table 6. Results from correlation analysis testing hypothesis H¹ (N=7,072)

All correlations are significant (p<.001). Controlling for firm size and age does not change the size of the correlation coefficient, nor does country and industry. These findings support the first hypothesis H¹ indicating that there is a positive correlation between perceived benefits from ICT adoption and e-business maturity of the firm.
Table 7 presents the results from the regression analysis, including the same variables. From the total sample column, we can conclude that hypothesis $H^1$ is confirmed again ($\beta = 0.327, p < 0.001$). To test the second hypothesis ($H^2$), we explore the influence of the intra-organisational changes due to ICT on the relationship between the perceived benefits and the e-business maturity. We primarily aim to test whether this influence is a moderator that affects the strength of the relationship between perceived ICT benefits and e-business maturity (Gonzalez-Benito, 2007). To this end we compare the explanatory power of the regression model for different stages of the moderator – i.e. the level of perceived intra-organisational impact of ICT. The dataset is split into three groups of equal size, distinguishing between organisations with relative low, medium and relative scores on the variable that measures the intra-organisational changes due to ICT. Table 7 presents the results.

<table>
<thead>
<tr>
<th>Subsamples: intra-organisational impact of ICT is$^a$:</th>
<th>Total sample</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-business maturity</td>
<td>Beta</td>
<td>Sig</td>
<td>Beta</td>
<td>Sig</td>
</tr>
<tr>
<td>Total sample</td>
<td>.327 ***</td>
<td>.263 ***</td>
<td>.199 ***</td>
<td>.178 ***</td>
</tr>
<tr>
<td>Industry$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food and beverages</td>
<td>-.131 ***</td>
<td>-.132 ***</td>
<td>-.113 ***</td>
<td>-.115 ***</td>
</tr>
<tr>
<td>Footwear</td>
<td>-.148 ***</td>
<td>-.124 ***</td>
<td>-.134 ***</td>
<td>-.087 ***</td>
</tr>
<tr>
<td>Pulp and Paper</td>
<td>-.079 ***</td>
<td>-.066 ***</td>
<td>-.082 ***</td>
<td>-.038 ***</td>
</tr>
<tr>
<td>Consumer electronics</td>
<td>.004</td>
<td>.006</td>
<td>.024</td>
<td>.005</td>
</tr>
<tr>
<td>Shipbuilding and repair</td>
<td>-.030 ***</td>
<td>-.019</td>
<td>-.028</td>
<td>-.014</td>
</tr>
<tr>
<td>Construction</td>
<td>-.110 ***</td>
<td>-.115 ***</td>
<td>-.088 ***</td>
<td>-.095 ***</td>
</tr>
<tr>
<td>Tourism</td>
<td>-.056 ***</td>
<td>-.072 ***</td>
<td>-.041</td>
<td>-.001</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>.044 ***</td>
<td>.040</td>
<td>.044</td>
<td>.016</td>
</tr>
<tr>
<td>Hospital activities</td>
<td>-.069 ***</td>
<td>-.022</td>
<td>-.118 ***</td>
<td>-.135</td>
</tr>
<tr>
<td>Country$^c$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>-.107 ***</td>
<td>-.099 ***</td>
<td>-.123 ***</td>
<td>-.141 ***</td>
</tr>
<tr>
<td>Germany</td>
<td>-.029</td>
<td>.006</td>
<td>-.104</td>
<td>-.068</td>
</tr>
<tr>
<td>Italy</td>
<td>.036</td>
<td>-.041</td>
<td>-.038</td>
<td>.031</td>
</tr>
<tr>
<td>Poland</td>
<td>-.009</td>
<td>-.038</td>
<td>-.080 **</td>
<td>-.101</td>
</tr>
<tr>
<td>Spain</td>
<td>.024</td>
<td>-.029</td>
<td>-.031</td>
<td>-.037</td>
</tr>
<tr>
<td>UK</td>
<td>.008</td>
<td>.031</td>
<td>-.019</td>
<td>.002</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>.032</td>
<td>.049</td>
<td>-.002</td>
<td>.041</td>
</tr>
<tr>
<td>Finland</td>
<td>.022 **</td>
<td>.093 ***</td>
<td>-.003</td>
<td>-.055</td>
</tr>
<tr>
<td>Hungary</td>
<td>-.081 ***</td>
<td>-.024</td>
<td>-.015</td>
<td>-.016</td>
</tr>
<tr>
<td>Employee size (log transformed)$^d$</td>
<td>.167 ***</td>
<td>.106 ***</td>
<td>.045</td>
<td>.014</td>
</tr>
<tr>
<td>Firm’s age$^e$</td>
<td>-.009</td>
<td>-.026</td>
<td>.006</td>
<td>.006</td>
</tr>
<tr>
<td>N</td>
<td>7,071</td>
<td>3,097</td>
<td>2,232</td>
<td>1,741</td>
</tr>
<tr>
<td>R2</td>
<td>.234</td>
<td>.150</td>
<td>.130</td>
<td>.099</td>
</tr>
</tbody>
</table>

Table 7. Results of OLS regression analysis testing hypothesis $H^1$ and $H^2$ (N=7,071)

$***$ p<.001; $**$ p<.01; $*$ p<.05

$^a$ Low = mentioned none of the four intra-organisational domains (see section 3.2.3); medium = mentioned one or two; high = mentioned three or four

$^b$ ICT manufacturing = category of reference, all other categories coded 0/1

$^c$ The Netherlands = category of reference, all other categories are coded 0/1

$^d$ log transformed


The three subsample analyses shows that with the three subsequent categories of intra-organisational impact of ICT, the strength of the relationship between the e-business maturity and the perceived ICT-benefits decreases (beta drop from +.26 to +.18). These findings support the hypothesis $H^2$. It should
also be noted that the explained variance of the total model decreases along the three groups (from 15% to 10%). Thus, the intra-organisational influence of ICT creates costs (e.g. due to the technology complexity), which moderates or counterbalances the perceived benefits of e-business maturity.

5 DISCUSSION AND CONCLUSION

The present study used survey data from over 7,000 European firms, to demonstrate that the organisations which have invested on ICT and display e-business maturity perceived more benefits from ICT than the firms with lower e-business maturity. This result is remarkably stable over 10 different European countries and 10 different industries. It also remains stable when controlling for age and the size of the firm. However, there might be firms which experience small benefits from ICT while having invested heavily in e-business. Those firms are a minority. Overall we can derive a quite a ‘positive message’ with regard to e-business development from the study’s results.

In this study we also found indications that when the ICT strongly influences intra-organisational domains of a firm the perceived benefits from ICT hold a weak relation with e-business maturity. This implies that firms should be made aware of the necessary organisational changes when adopting e-business applications. Depending on the ‘as it is’ intra-organization situation, the benefits from ICT may not be evident from the start, as they require resources, investments and – in terms of attitude – patience. Hence, the public policies should be focused on facilitating firms in handling both the respective benefits and costs from ICT adoption. In turn, if firms’ perceptions of the benefits related to ICT become more positive, this may further stimulate investments, ICT adoption and increase e-business maturity.

Limitations and further research

The empirical part of this research is based on a secondary analysis of an existing dataset (i.e. data from the organized 2006 E-business w@tch survey). A limitation is therefore that the measurements do not fully match the intended concepts from our hypotheses. This applies, for instance, to the limited number answer categories in measuring ICT adoption (‘yes’, ‘no’) and the general formulation of ‘ICT’ in measuring their perceived benefits. In further research these measurements could be improved by specifying the type of software or applications used, the history, intensity and goal of usage, and so on. Related to this, it is recommended to further validate scales to measure e-business maturity, perceived benefits from ICT and intra-organisational changes. As European firms become more mature in terms of ICT and e-business, this requires new maturity measurements and likewise new explanatory research frameworks (cf. Aryee et al, 2008).

In addition, also our results are subject to limitations, but provide drivers for further investigation. Table 6 showed that European countries only slightly differ in the correlation between e-business maturity and ICT benefits, but Table 7 displays different levels of perceived ICT benefits by country. This is an interesting subject to explore further, e.g. on the fact that e-business measurement is rightly assumed to be generic. It might be the case that respondents from different countries do differ in their level of technical knowledge to answer technology specific questions properly. Extending this study through cross-national comparison implies the analysis of other cultural and socio-economic factors on e-business adoption (Miller et al. 2006), maturity and perceived benefits from ICT (Agourram and Ingham, 2003).

References


DEALING WITH TIGHT COUPLINGS AND MULTIPLE INTERACTIONS IN COMPLEX TECHNOLOGICAL SYSTEMS

| Journal: | 17th European Conference on Information Systems |
| Manuscript ID: | ECIS2009-0218.R1 |
| Submission Type: | Research Paper |
| Keyword: | Case Study, Complexity / Complex adaptive systems, Implementation, Health information systems / medical record systems / care delivery / |
DEALING WITH TIGHT COUPLINGS AND MULTIPLE INTERACTIONS IN COMPLEX TECHNOLOGICAL SYSTEMS

Aanestad, Margunn, University of Oslo, Department of Informatics, PO Box 1080 Blindern, Oslo, Norway, margunn@ifi.uio.no
Jensen, Tina Blegind, Aarhus School of Business, Department of Management, Haslegaardsvej 10, 8210 Aarhus V, Denmark, tibj@asb.dk
Grisot, Miria, University of Oslo, Department of Informatics, PO Box 1080 Blindern, Oslo, Norway, miriag@ifi.uio.no

Abstract

In this paper we discuss the challenges of dealing with interdependencies in complex assemblages of heterogeneous and interconnected information systems (IS), which we conceptualize as organization-wide information infrastructures. We draw on Perrow’s studies of complex technological systems, where interactions, mechanisms, and couplings are emphasized. We base our paper on an empirical case study from a Norwegian hospital, where a seemingly trivial project aimed at the introduction of scanners turned out to be more complex than expected. This we claim is partly due to the interdependencies and tight couplings between information systems, actors, and work practices in the hospital environment. The paper’s main focus is on describing what it entails in practice to deal with these interdependencies during and after implementation. We emphasize the work of sorting out and dealing with various types of interactions and couplings.

Keywords: Interdependencies, information infrastructures, couplings, health care.
1 INTRODUCTION

Organizations today rarely operate with isolated, task-specific information systems (IS). In general, they have interconnected collages of multiple and different systems, intended to support multiple work tasks, interaction and communication (Ciborra 1997; Star 2002). Some systems can be large, monolithic, and generic, such as enterprise systems, whereas others can be smaller and more dedicated. Legacy systems abound and thus the collection of systems span several technological generations. In addition, there is usually some degree of attempted, achieved or at least wished-for integration of these systems (Hasselbring 2000).

Much has been written about the implementation and use of IS in organizations. Studies have looked at changes in organizational practices (Vaast and Walsham 2005), collaboration and communication flows (Avgerou et al. 2004), and power structures (Robey and Boudreau 1999; Silva and Backhouse 2003; Howcroft and Mitev 2000) with the implementation of IS. Furthermore, researchers have studied the meanings that users ascribe to technology and how users’ attitudes and perceptions relating to new technology are socially shaped (Orlikowski and Gash 1994; Davidson 2006). In the existing literature on IS implementation and use, however, we only find limited accounts of how new IS are related to and get intertwined with the existing portfolio of systems in the organization, also known as the installed base already in place and conceptualized as information infrastructures (Ciborra et al. 2000; Star and Ruhleder 1996). In contrast to isolated and task-specific information systems, information infrastructures pose particular challenges for the organization (Hanseth and Lyttinen 2004; Monteiro and Hanseth 1995) as they are characterized as being large, complex, heterogeneous, and emergent. There has been surprisingly little emphasis on the impact of interactions and tight couplings between the various systems that make up the information infrastructure.

In this paper we address the challenges of extending the existing portfolio of systems already in place in the organization, and of reorganizing work routines accordingly. We consider the installation and post-installation period where a new technology is to be integrated into current work practices (Jasperson et al. 2005). We emphasize the role of interdependencies and tight couplings between the new system and the pre-existing information infrastructure. The paper’s main focus is on describing what it entails in practice to deal with these interdependencies during and after the implementation of new technology.

We illustrate these challenges by analyzing a case study from health care, where the issues of managing multiple IS (each of which contains parts of the relevant patient information) are highly relevant. We report on findings from a longitudinal study following the introduction of an Electronic Patient Record (EPR) system at a major Norwegian hospital. Most of the empirical material presented here comes from a seemingly trivial project where scanners were introduced in order to scan paper documents. The project turned out to entail more deep-going organizational changes and to be more time-consuming than expected. This we claim is partly due to the interdependencies and tight couplings between IS, actors, and work practices in the hospital environment. Issues such as these, we believe, constitute one significant reason for why the digitization of health care is complex and has been protracted. Consequently, we wish to examine this complexity that is evident even with seemingly simple projects.

For this we employ the information infrastructure perspective, which is presented in the next section. This is followed by a presentation of concepts used by Charles Perrow that help us conceptualize and discuss issues related to complex technological systems. Next, we present the research approach and the empirical case study, and finally we discuss the findings and their implications for the field of IS implementation studies.
2 ORGANIZATION-WIDE INFORMATION INFRASTRUCTURES

We base our work on a theoretical perspective that has its focus on the complexity of socio-technical assemblages. The information infrastructure perspective, as presented by e.g. Hanseth and Lyytinen (2004) or Monteiro and Hanseth (1995), has been developed to address the large-scale, heterogeneous, interconnected collections of systems. This way of conceptualizing the phenomenon emphasizes the long-term, gradual development of an information infrastructure. Rather than being implemented, it evolves over time through expansion, extension, and replacements. It is a socio-technical and heterogeneous collection of elements and it encompass much more than just the technological infrastructure. Star and Ruhleder (1996) argue that information infrastructures extend beyond human skills and materiality to include social, organizational, as well as moral elements. In contrast to the notion of IS, an information infrastructure is used by many for different purposes, not just one specific task. Thus an organizational information infrastructure is not “discrete, stable, independent and fixed” (Orlikowski and Iacono 2001), but multiple, evolving, and interconnected. Often information infrastructures span multiple inter-organizational contexts where they mediate interactions among several organizations (Gal et al. 2007).

Interventions in interconnected and complex environments can only be planned and rationally executed to a limited degree. Studies of information infrastructures tend to describe unintended effects of actions and the consequential slippage and “drift” of the processes and results (Ciborra et al. 2000). The same kind of phenomena is emphasized in the improvisational model for change (Orlikowski and Hofman 1997), which is claimed to be particularly relevant for open-ended and customizable technologies. Orlikowski and Hofman emphasize that implementing new information systems is not a one-time event, but an ongoing learning and adaptation process. Everything cannot be anticipated ahead of time, but the degree, depth, and complexity of changes is only fully understood when implementation is underway. Consequently, they see the resolution to this as lying in approaches that “enable organizations to systematically absorb, respond to, and even leverage unexpected events, evolving technological capabilities, emerging practices, and unanticipated outcomes. Such a model for managing change would accommodate – indeed, encourage – ongoing and iterative experimentation, use, and learning. Such a model sees change management more as an ongoing improvisation than a staged event.”(Orlikowski and Hofman 1997, p. 12)

In this paper we emphasize the need for such adaptive or learning-oriented approaches when dealing with the interconnectedness of the organization-wide information infrastructures. In order to analyze the nature and consequences of the interdependencies, we have chosen to draw on Charles Perrow’s conceptualization of complex socio-technical systems, where the role of complex interactions and tight couplings is emphasized (Perrow 1984). In laying the ground for the “normal accidents theory”, Perrow conceptualizes complex systems along two axes:

- The first axis concerns the character of interactions between the parts or elements of the system. The different parts of the system (e.g. organizational units) may collaborate through simple, linear relations. The linear interactions are expected and planned, and they stand out as clearly visible and understandable. Alternatively, the parts may collaborate through more interactively complex mechanisms with multiple dependencies and feedback loops. The complex interactions are unfamiliar, unplanned, and unexpected. They are not visible and not easily comprehensible (Perrow 1999 [1984], p. 78).

- The other axis concerns whether the linkages between the parts of the system are loose or tight. If the linkages are tight it means that action at one location has immediate and given effects elsewhere. When for instance the steps of action are tightly prescribed in advance, or when there is no slack in the production that allows for flexibility in handling disruptions, we have a tightly coupled system. On the other hand, a loosely coupled system may be slower to respond to interventions, but also more robust with respect to fault recovery due to its slack and buffer resources.
Perrow exemplifies how systems can differ with respect to their type of interactions and couplings along the two axes. For instance, a post office is characterized as a routine bureaucracy where the interactions between work practices, actors, and information systems are relatively linear and loosely coupled. In such a system, the order of sequences is straightforward. Changes in one part of the system can take place without having any prominent impact on other parts, e.g. mail can pile up without undue alarm. In contrast, a power grid system is also linear but tightly coupled, and disturbances in one network segment will immediately affect the other segments. Perrow suggests that universities are examples of organizations that exhibit complex interaction mechanisms, but are loosely coupled, while a nuclear plant is characterized by complex interactions and is also tightly coupled. Many unexpected interactions are possible and changes in one part of the system may have considerable consequences in other parts. This insight is the basis for Perrow’s major thesis that accidents are unavoidable (hence “normal accidents”) if the systems are complex and tightly coupled.

In this paper we use the information infrastructure perspective to define our research object: we do not study one single stand-alone information system, but rather an extended, socio-technical network that evolves gradually over time. We complement this view with the concepts of interactions and couplings used by Perrow to discuss the complexity in various forms of socio-technical organizations. We see Perrow’s concepts as helpful because they complement the information infrastructure theory with a more fine-grained “lens” to examine the interdependencies of the different parts in information infrastructures. Also Perrow’s conceptualization reminds us that different “configurations” of the socio-technical imply different forms of complexity and may have different effects. In the following we illustrate how the aspects of interaction mechanisms and couplings in the organization-wide information infrastructures had a practical impact on the implementation of a small project.

3 RESEARCH METHOD

Our empirical material emerged from longitudinal research collaboration between researchers at the University of Oslo and the IT department at a major Norwegian hospital in Oslo (Rikshospitalet, Oslo). The hospital is a highly specialized university hospital with regard to complicated treatments. In 2006 it had almost 8000 employees and treated more than 300 000 patients.

As a part of the implementation of an Electronic Patient Record (EPR) system at the hospital, a scanning project was initiated in 2003. This project was studied from November 2003 to October 2005. The study site was one of the four pilot departments – the Women’s Clinic, one of the largest departments at the hospital and a major user of the patient record archive’s services. The first author conducted seven formal interviews with the project management group and with staff from the IT and archive department. These interviews focused on the project’s progress and the interviewees’ perceptions of obstacles and opportunities. The interviews lasted for between 45 minutes and 1.5 hours and were usually taped and partly transcribed.

The study of the scanning project was a “side activity” related to a more structured and general study of the EPR implementation. This meant that no new formal research project was defined, and the access to the field site was more informal. For instance the researcher participated in 6 meetings at one of the pilot departments for the project and in return for access wrote minutes from the meetings. Some of the meetings were project meetings where IT staff was present, while others were internal meetings with only staff from the department. In addition, the researcher assisted the personnel in compiling detailed information about document flows in the department and documented the actual scanning work through time studies, including existing errors and problems etc. These results were fed back to the department and the project management. In total 17 hours of participant observation were conducted, and during this work a number of informal conversations were carried out; these added to the background perception of the process. Based on transcriptions of the interviews and field notes, we built a storyline of the scanning project and analyzed the material on the basis of the conceptualizations that we have presented above.
4 THE STORY LINE OF THE SCANNING PROJECT

4.1 Background of the case study

The transition from a paper-based patient record to a complete EPR system started in 1996 with a few departmental installations of the EPR system. The process was delayed because the hospital moved into new facilities in 2000, but in 2001 the EPR system was installed in all clinical departments.

The functionality of the EPR system, however, did not allow it to fully replace the previous paper-based system. The EPR consisted to a large extent of textual information and a large part of relevant clinical information was still on paper or stored in other specialized systems. Due to this, the legally valid record was still the paper-based record and all entries in the EPR were thus printed on paper and archived with the corresponding paper patient record. The vendor seemed unable to deliver updated and improved versions within the time frame expected by the hospital (contractually, the EPR system should have been fully developed by 1999). In 2003, this situation caused a critical space problem in the archive. Only limited space was allocated to the central archive in the new facilities as it was expected that a fully functional EPR system would be in place. While waiting for the full version of the EPR, the hospital management decided to scan both the existing patient record files in the archive and the individual paper documents in the clinical departments to decrease the paper flow to the archive.

The Scanning Project started in the spring 2003 with a bid for tender on scanners as well as software that imported the scanned documents into the EPR system. The plan was to select, test and finalize a solution over the summer, start the scanning in early autumn, and stop the flow of paper documents to the archive by January 2004. The project was perceived as a minor, straightforward, and problem solving project aimed at solving the archive crisis. The estimates predicted the employment of one support person from the IT department with a 50% position for 6 months. In December 2004, after some delays, the scanners were installed in four pilot departments and the secretaries started testing scanners and software. Most of 2004 and 2005 were spent on teasing out the actual work practices associated with digitizing the work flow within the departments and the hospital. The volume of paper being sent to the archive started to decrease more than a year later than initially expected.

Figure 1. Chronology of the scanning project: The dotted line indicates dependencies on actors beyond the local context (i.e. the Women’s Clinic, in the middle)
We enter the case in December 2003 and look at the scanning project in one of the pilot departments, the Women’s Clinic. We describe three events (events 1, 2, and 3 in figure 1 above) that made the project deviate from its planned course and somehow complicated the process.

4.2 Event 1 (2003-2004): Resolving local procedural and technical constraints

During the first pilot period (December 2003 - January 2004), bugs were identified and fixed and the employees at the Women’s Clinic developed routines for how to arrange the scanning. The first documents targeted for scanning were the incoming referral letters, which are letters sent by external doctors to request a patient’s admission to the hospital. These letters could be several pages long and usually all of the pages were scanned and saved as one file in the appropriate folder in the EPR and tagged with the date of reception. In some instances the referral letter was incomplete. For instance the letter would mention that additional examinations had been performed and that the results would be forwarded to the hospital when they were ready (e.g. when received from the laboratory). In such cases, the forwarded documents arriving later would logically belong to the same referral letter; however, it was not technically possible to append the after-sent documents to the previously scanned referral letter. The scanner software generated images using the TIFF file format, and in order to create one file, the different pages had to be scanned at the same time and joined before exporting them to the EPR system. Thus in the case of incomplete referral letters, an additional file had to be generated for the separate (after-sent) documents. The only way to link these documents in the EPR was to tag them both by date of the referral letter. To do this properly, the person who scanned the documents had to figure out which referral a particular document belonged to and use the same date for storing it. Initially, this work routine was accepted in the pilot departments, but it was later debated again. After some discussion, it was decided that after-sent additional information should be stored in an EPR folder called “External correspondence, miscellaneous”.

The scanning activities were also shaped by other technical characteristics of the existing information infrastructure. The activities required cooperation between three different software systems: the scanning software, the Electronic Patient Record (EPR) system, and the Patient Administrative System (PAS). Before a document could be scanned, a digital record had to be created in the patient’s name in the EPR system. However, even before the record could be created in the EPR system, the patient information had to be entered into the patient administrative system (PAS). Initially, it was assumed that the speed of the process would improve if the same person was responsible for all these steps; however, the PAS application was running on the hospital’s Windows NT network. The new scanning application required Windows 2000 and was installed on separate machines initially, while waiting for the eventual upgrading of the whole hospital network. Consequently, the two applications were running on different machines in different work places, and at least two persons had to be involved in conducting the task.

Thus we see that the redesign of work practices did not happen freely “from scratch” but had to start from the given constraints, in this case constraints posed by the software, the operating systems, and the existing network technology. The redesign furthermore necessitated cooperation with actors beyond the department. The temporary workaround utilizing different workstations and different network segments involved the hospital’s IT department. Furthermore, negotiations with other departments were required to establish common work routines with respect to the scanning. But still these issues did not entail dependence on actors beyond the hospital (i.e. external actors), such as the next example illustrates.

4.3 Event 2 (2004-2005): Dependence on external vendors

Some procedures for improved mail handling were drafted and implemented in the Women’s Clinic during 2004, primarily to ascertain that incoming letters were distributed as efficiently as possible within the department, and that the evaluation was speedy. However, major changes could not be carried out because the EPR system, according to the contract specification, was not yet ‘complete’.

Proceedings ECIS 2009
The EPR system primarily served as a document repository and did not yet offer much work process support. For instance, the system did not have a message functionality that could be used to notify the doctor when a referral letter was awaiting evaluation. The paper letter would be physically placed in the doctor’s mail shelf, and thereby serve to notify the doctor about the task to be handled (i.e., the evaluation of a referral letter). Ideally, the paper letter should be removed from the workflow after scanning; however, that was not possible since the notification functionality was not yet implemented in the EPR system. When asking for this functionality in March 2004, the project team in the Women’s Clinic was told (jokingly) by the IT department representatives that it would probably be a Christmas present. Then again in November 2004, they were told that it was a part of the version being currently tested.

The point we want to emphasize here is that the provision of this functionality was not within the mandate of the project, or even of the hospital. The vendor would make the decision of whether and when to implement this functionality based on weighing costs of development against a potential increase of market share and income generation. Thus not all wishes from the various customers (hospitals) were fulfilled. In practice, this dependency on the vendor meant that the doctors continued to receive paper letters in their mail shelves (even if these letters were also being scanned). When finally implemented in 2005, the notification functionality allowed the removal of the scanned paper documents from the workflow. The problem illustrated here is just one example of how the clinical workflows were complicated with the introduction of scanners and how changes were dependent on external partners. In this case, we see how the use of scanners did not immediately live up to the initial aim of reducing the amount of paper. Rather, it complicated the work even further as the healthcare professionals now had to work with both paper-based and electronic documents.

4.4 Event 3 (2005-): Defining departmental and hospital-wide standards and procedures

The various clinical departments were expected to reorganize their work processes. The hospital can be considered a professional bureaucracy (Mintzberg 1983) and no-one in the administration had detailed and extensive knowledge about the idiosyncrasies of documentation practices and work patterns. This necessitated a decentralized approach to workflow redesign. Still, there existed a need to coordinate these redesign activities at an organizational level, since local changes in work practices could have consequences for other collaborating actors at other levels. The reorganization was initiated in the beginning of the project but it was becoming increasingly formalized by the end of 2004 and in 2005.

The local project team at the Women’s Clinic was requested by the project management to develop procedures within given time frames. The project team drafted procedures for e.g., incoming mail handling, for evaluation of referral letters, and for detailed instructions as to which documents should be scanned and which not. These procedures were formally part of the organizations’ quality assurance system, and thus this request had some legitimacy and became a rather well-working way to get the department to actually design changes. The procedures also served to standardize learning across the multiple departments. The standardization was achieved through a lengthy process where numerous versions of multiple procedures (both departmental and general, hospital-wide) were circulated and discussed. The project team asked for general template procedures in order to accomplish this task. The templates were made and circulated, and the project team worked with them for some time before sharing the drafts and discussing revisions with other departments. The required coordination of the redesign was thus achieved through a process where the local or departmental idiosyncrasies were exposed and adjusted, and where the overall requirements (central as opposed to local/departmental) also were met.

These discussions were not always straightforward. For instance, the general rule was that all incoming mail should be opened at once and processed immediately to avoid delays. In the Women’s Clinic, the mail was distributed to the four sections before it was opened and processed. Initially, the involved persons thought it would be a good idea to centralize mail handling and scanning. Later in
the process, other secretaries became involved, and they voiced objections to this model. The personnel in one section handled sensitive personal information (e.g. related to gender identity disorders, anonymous semen donors etc.) and they registered this information in a version of the EPR which had strictly limited accessibility. They wanted to continue to open their mail themselves in order to maintain the patient privacy. The general rule was also a problematic issue in the delivery section, which received more applications than it could grant. They rejected approximately half of the yearly 5000 applications from women wanting to give birth there, and these referral letters were forwarded to other hospitals. The hospital’s established procedure (to scan incoming referral letters at once) worked well for most departments. But in the delivery section it would not be appropriate, as it would necessitate the unnecessary creation of around 2500 files in the EPR system for women who would not be admitted after all.

Several of such special cases were brought forth in the ongoing discussions in the project meetings, and the job of distinguishing between what should be general, hospital-wide procedures, and what should be specific to each local department took several months. This process needed to allow for local variations in the procedures (e.g. as the case of the delivery section shows), but also to eliminate these variations when they were not necessary and considered counter-productive.

5 ANALYZING THE FINDINGS

We have described some concrete examples of how a seemingly trivial project of implementing scanners turned out to take much longer time than expected and was dependent on a multiplicity of actors and actions. We do not offer this account as an example of negligent estimation or suboptimal project management. Neither is it an example of an unusual IS project. The case is rather trivial, even banal, and certainly not dramatic. The story illustrates the impact of interactions and couplings; issues that are crucial, we claim, but of such a mundane character that they are often overlooked by IS researchers trying to conceptualize IS-related change processes. In the discussion below, we focus on these interactions and couplings as well as the strategies employed to manage them during the scanning project.

5.1 Linear or interactive interactions, tight or loose couplings

In some instances described above, the type of interaction was simple and linear whereas in others, they were highly interactive. Event 2 illustrates how the hospital was dependent on the vendor to develop and deliver the notification functionality of the EPR system. The relation between the hospital and the vendor can be characterized as simple; the vendor was the only one who could provide a solution, and as long as it was not provided, a delay occurred with respect to the particular usage pattern of the EPR system. The difference between the scanner software’s requirements (Windows 2000) and the existing Windows NT network, as illustrated in event 1, was another example of a linear and simple relation. This time, the dependence was on another entity (the IT department) within the hospital organization, not an external actor. Nevertheless, the project team in the Women’s Clinic could only wait for the change to happen. They did not have any power to change the timeline for network upgrading as this was a separate process in the organization. In the meantime, they needed to design a workaround to deal with this issue. The possibility to design workarounds and tolerate delays in this process shows that the couplings were loose, not tight. There were multiple ways for the organization to keep on working, despite the delays introduced by the other actors and processes.

The definition of procedures for the new work practices showed a more interactive pattern as illustrated in event 3. The various departments needed to sketch the new procedures and try them out locally. At the same time, the employees at the Women’s Clinic had to discuss with the other departments in cases where the procedures were dependent on or had consequences for them. The procedures were tightly linked with the documents associated with the internal patient transfer between departments. In addition, the hospital’s quality system imposed a structure of distinguishing between general and local procedures, and we witnessed a degree of interactively sorting out the
procedures that belonged to the hospital versus the departments. Experiences from the pilot departments were collected and compared in the process of sorting this out. With respect to the difficulty of replacing the referral letter, it was discovered that a single artifact (i.e. the paper letter) served two functions simultaneously; it was both an information carrier and a “signal” for action in the process. One of these functions could easily be digitized, while the challenges of digitizing the other were discovered along the way. Such double- or multi-mode functioning is, according to Perrow (who called it “common-mode functioning”), one of the major reasons behind interactively complex systems.

5.2 Sorting out and dealing with interdependencies

Most of the interdependencies accounted for in the case description were not predicted in advance. This is not to say that they were impossible or even difficult to predict. For instance, it could easily have been foreseen that the TIFF file format would exclude appending the test results to a referral letter, if anyone had asked the question. However, even if predicted, it was doubtful whether the product, a commercial off-the-shelf application, would have been changed. In principle, these issues (at least some of them) could have been predicted if the organization had run a process where they used for instance storyboards or use cases. However, the introduction of scanners was not considered a major project that required involvement from IT professionals. Since the process was handled by the workers themselves, storyboards and use cases were unknown techniques and were not used, except for quite simple drawings with boxes and arrows in order to show the work flow. In practice, these dependencies were only realized as the actual and practical work progressed. This is not necessarily because they were complex and intractable; rather, there were just too many and too small dependencies, and it was a task in itself (“beside” the primary project work) to find out what to look for and to realize who needed to be involved.

It is the work of sorting things out and dealing with interdependencies that we wish to draw the reader’s attention to. In the case, we see that the discussions often centered on sorting out what to take first and second, and what to prioritize and what to downplay. This process was linked to the uncertainty in the process because of its links beyond the local context. The team members needed to sort out a number of issues such as: What are the preconditions and consequences (upstream and downstream)? What are the limiting factors and what are the constraints of change? What can we do and not do? Where do we have any impact and what is beyond our scope of influence? The actors also needed to consider how to run a changed procedure in parallel with maintaining the pre-existing paper-based ordering so that the Women’s Clinic could still interact with the non-pilot departments. In response to our inquiry on the progress of providing a feature that the users had requested, one of the project managers from the IT department replied: “No, we haven’t fixed that and we will not do that for the time being. Remember, we’re trying to eat an elephant here. We cannot take it all at once”.

The changes had to occur in a coordinated way so that the ongoing work of patient care would not suffer during the transition process. This interlinked nature of the organizational routines and the dependencies on other actors, both internal and external, delayed the process compared to the initial estimates.

5.3 Infrastructural implementation

We believe it is significant in this respect that the implementation of scanners was not considered a strategic and important project. Infrastructural technologies are expected to just “be there” and support the core organizational activities. In a hospital, the activities are centered on patient treatment and care, and the information infrastructure is seldom considered to be of strategic importance in its own right. Especially this goes for the ongoing incremental extension or replacement of the elements of an information infrastructure, which is often seen as necessary maintenance. This may have the effect that change processes (both implementation and assimilation) associated with them receives low organizational attention, priority, and allocation of resources.
The change process was by necessity decentralized, but the clinical staff lacked expertise and applicable models for how to handle the required changes, as these were not related to their core competency. Thus, when no external resources (such as consultants) were available, the process may be handled in a suboptimal way. Since no resources were allocated, the actors perceived these tasks as additional to their ordinary work burden, and they easily defected the meetings and activities. This character of the process may be a challenge since it delays actions and dissipates the change momentum in the organization.

6 DISCUSSION AND IMPLICATIONS

In this paper we have argued that the existence of interconnections and dependencies is a fundamental trait of the information infrastructures of today’s organizations. While these interdependencies can be of many types and degrees, we have focused on the mundane and trivial types of interdependencies rather than large-scale and dramatic ones. When changes occur in the organization, small interdependencies may become significant enough to have an impact on larger organizational processes.

In our case study, we emphasized how the actors involved in the project resolved the challenges associated with the interdependencies. This was a cumbersome process of sorting out the consequences of them. It involved negotiations with other actors (outside the project), coordination of multiple local changes with one another, and coordination of local with overall changes (i.e. both cross-site and cross-level negotiations).

We claim that IS implementation studies have to a limited degree focused explicitly on how new IS are related to and get intertwined with the existing portfolio of systems in the organization. The emphasis is usually rather on project management and reception in the organization. This orientation is also reflected in how failures and problems are conceptualized. Common “folk explanations” for delays or failures of IS implementation often emphasize incompetent project management, inappropriate design, or lack of user involvement. In addition, IS researchers have emphasized the importance of upper-level management support (Myers 1994), the role of users’ resistance due to various reasons, such as badly designed systems, or organizational politics (Markus 1983; Keen 1981). Failure to learn from previous experience or to transmit learning between projects has been seen a major reason for IS failures (Lyytinen and Robey 1999). Also with Orlikowski and Hofman’s improvised change model, the need for learning-oriented and adaptive approaches is emphasized (Orlikowski and Hofman, 1997).

Our study does not lead us to disagree with these insights. On the contrary, we argue that our emphasis on minor instances of interactions and interdependencies are complementary to these studies. We believe that such details are important in any project, although they may be overlooked or abstracted away in the accounts that are produced. We want to draw attention to this specific issue since we believe attention to this can contribute with conceptualizations that are practically useful in IS projects. These conceptualizations need to describe how we can identify and deal with the issue of interdependencies that emerge because of the infrastructural character of an organization’s systems portfolio. Our empirical case study could be considered as a first step in this direction.

The question now is how actors, who are involved in similar projects where multiple types of interactions and couplings are encountered, may act and react? The answer is not straightforward, and we will not be able to provide managers or other actors involved in IS implementations with clear-cut answers and a list of best practices. However, one way of acting in such a context is suggested by the “bootstrapping” strategy proposed by Hanseth and Aanestad (2003). In discussing how to establish information infrastructures in health care, the strategy addresses the dilemma when the number of adopters is below critical mass, i.e. too low to ensure general benefits from the infrastructure. The focus of the bootstrapping approach is on how one can work in such a situation in order to establish a self-reinforcing growth process. It advises the selection of an appropriate starting point in order to create a platform that allows some action to be taken, given the existing constraints and opportunities.
The next step is to search the context for opportunities and then to expand this platform on the basis of these available opportunities. General principles within this model are to expand existing solutions rather than request new solutions (more users before more functionality), to experiment in shielded areas only, to learn from simple use areas before complex ones, and to identify usage areas with low cost and high gains for the prospective users. These principles help in sorting the immediate steps from the future steps, in choosing what to do first and last.

The benefit of this strategy lies in breaking down a task that seems impossible when considered as a whole. This is of relevance also with respect to handling and dealing with interdependencies during implementation, a task which an IT employee characterized as “eating an elephant”. If it is possible to plan organizational implementation in a way where a starting point has been carefully selected so that the dependencies extending beyond the project’s sphere of influence are minimized, and where the need to negotiate cross-site and cross-level are also minimized, then the greater the chance for successful implementation of this limited change.

If a starting point like this can be found where a limited change can be achieved, then a platform has been built from which to proceed. At this stage, some interdependencies have been uncovered and perhaps even sorted out. Ideally, the project team should have learnt from this process, and be in a better position to predict and prepare for new and “larger” (i.e. more far-reaching or comprehensive) interdependencies in the next rounds. On the contrary, if no such selection has been carried out, and the whole project (“the elephant”) is taken on, the more likely are delays and frustrations in the change process due to the couplings and dependencies of systems and procedures beyond the project site. It may not be possible to achieve this in a perfect or optimal way. For instance, there may be no area where there is not a need to coordinate change across locations. We have no reason to claim that the selection of a starting point or the planning of a sequence of steps in our case was not the most optimal strategy. However, even if the single optimal strategy cannot be defined, this way of thinking may effect a change in mindset. We believe that the project management should be oriented towards vigilance and detection and focused on learning to deal with interdependencies rather than insisting on meeting milestones and deliverables.

7 CONCLUDING REMARKS

The purpose of this paper was to illuminate a specific source of complexity associated with implementing IS that becomes an integrated part of an organizational information infrastructure. We have emphasized the role of complex interactions and tight couplings in such a process and described how a long and cumbersome process of dealing with such interdependencies delayed the project we studied. The challenges associated with the interdependencies had to be resolved in a manner which entailed sorting out the consequences of the couplings that were proposed, introduced or detected. It also involved negotiations with other actors, over whom the project did not have control. Finally, due to the nature of an organizational infrastructure, the solutions had to align multiple local changes with one another and coordinate local with overall (organization-wide) changes. We believe this phenomenon is understudied, yet significant. It constitutes a particularly challenging aspect of leveraging information infrastructures to achieve organizational change. Finally, we have sketched how a learning-oriented approach to deal with these multiple interdependencies and couplings could look like.

References


THE MEDIATING ROLE OF IT KNOWLEDGE INTEGRATION CAPABILITY IN THE RELATIONSHIP BETWEEN TEAM PERFORMANCE AND TEAM CLIMATE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0520.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>IT capability, Knowledge integration, Organizational theory, Distributed organization / teams</td>
</tr>
</tbody>
</table>
THE MEDIATING ROLE OF IT KNOWLEDGE INTEGRATION CAPABILITY IN THE RELATIONSHIP BETWEEN TEAM PERFORMANCE AND TEAM CLIMATE

Abstract

The purpose of our paper is twofold: on one hand we draw on the team climate literature to understand how IT knowledge integration capability can be promoted within team-based structures; on the other hand we rely on resource-based view framework for understanding how IT knowledge integration capability affects teams performance, in terms of effectiveness, and efficiency. We tested our research model on a sample of 410 members and leaders of 69 organizational work teams. Results show the critical role played by team climate for favouring IT knowledge integration capability, which in turn affects team outcomes.

1 INTRODUCTION

Organizational researchers and managers acknowledge the importance of teams to organizations (Hoegl and Gemuenden, 2001). In modern organizations, knowledge workers are becoming more and more important, and cooperation became completely immanent to the labouring activity itself. Productivity, wealth, and the creation of social surplus take the form of cooperative interactivity through linguistic, communicational, and affective networks. In mainstream literature terms, this means the increasing relevance of teams within modern organizations (Appelbaum and Batt, 1994; Argote and McGrath, 1993; Adler et al., 2008).

Researchers have developed several models and constructs in the attempt to analyze the degree to which team members interact for explaining their ability to accomplish complex tasks (i.e. creative, innovative, and knowledge-based task). Indeed, complex tasks are regularly conducted through teams, as teams represent the most immediate social entities through which individuals obtain resources (Faraj and Xiao, 2006; Hoegl et al. 2003). For example, Li and Zhang (2002) found that members’ integration facilitates innovative solutions in the product development domain, which relies on the ability to exchange, coordinate and aggregate individual contributions (Faraj and Sproull, 2000). Thus, the ability of team to integrate resources through the leverage of team processes represents a favourable condition for facing the complexity of their task. In order to cope with such complexity, team members should be able to exchange and integrate knowledge for obtaining positive performance (Garud and Kumaraswamy 2005; Gold et al. 2001; Lee and Choi 2003; Pfeffer and Sutton 1999; Wasko and Faraj 2005).

According to the resource-based view framework, knowledge represents one of key resources to obtain and transform other resources. The way through which team members rely on processes for favouring the creation, codification, and sharing of knowledge, has been recognized to positively affect performance (Gold et al. 2001). Despite the widely recognized importance of team capability to activate processes for integrating knowledge, the effectiveness of knowledge management process capability needs to come along with technological capability that supports and incorporates knowledge management processes (Burgelman 1994). While established rules and directives contribute to firm performance, proper IT support is required to maximize the effect of the rules and directives as the knowledge integration mechanism (Ray et al., 2004). IT as an organizational capability can facilitate
the process of transforming knowledge into action, and the effective exploitation of the rules and directives may depend upon a well-established technological infrastructure (Armour, 2001). Thus, the harmonious combination of knowledge management process capability and IT capability represents the effective configuration of resources and knowledge to determine team performance.

Moreover, according to previous literature pointing out that facilitating conditions have emerged as an important precursor to knowledge integration (Grant, 1996a), we believe that IT knowledge integration capability is affected by the existence of a team-level climate that favours the circulation of information among team members. Indeed, a positive team climate creates the types of team environments in that collaboration, team learning, and thus innovative activities are encouraged (Lumpkin and Dess 1996). In such environments, teams can integrate knowledge to reduce members isolation (Kanter 1983). Thus, a positive team climate may enhance an organizational performance by creating new knowledge (Ireland, Hitt and Sirmon 2003). Through developing a positive climate, teams are able to create a new knowledge and exploit it (Dess et al. 2003; Kazanjian, Drazin and Glynn 2001).

The remainder of this paper is structured as follows. First, we present constructs and hypotheses. Next we discuss the methodology, including data collection, and analyses. We conclude with a discussion of our results and implications for theory and practice.

2 THEORY AND HYPOTHESES

The first objective of this paper is to understand how IT knowledge integration capability can be promoted within team-based structures.

Following some previous researches (i.e. Grant 1996a,b), we name IT knowledge integration capability, a type of combinative capability (Kogut and Zander, 1992), which consists of knowledge management process capability, and technological capability. Knowledge management process capability is defined as the capability to capturing, storing, sharing, and using knowledge (Davenport and Prusak 1998; Bock et al., 2005). Technological capability is defined as the capability to leverage IT in order to obtain tangible and/or intangible benefits (Bharadwaj, 2000).

Extant literature in organizational behaviour domain proposed different types of teams, and analysed different team mechanism and properties (Cohen and Bailey, 1997; Barrick et al. 2007). Team climate plays a pivotal role because it refers to the creation and influence of social contexts (Bock et al., 2005) that influence team members’ behaviour. Climate represents a shared perception of the types of behaviours, practices, and procedures that are supported in a specific context (Schneider et al., 1998). According to Glomb and Liao (2003) team climate influences individual behaviour through social information processing mechanism (Salancik and Pfeffer, 1978), thus leveraging the way individuals think and feel about a certain aspect of their environment (i.e. specific team environment). In this study, we examine the relationship between two critical properties of team climate (autonomy and experimentation) and IT knowledge integration capability. We focus on these two dimensions of climate because they represent the salient aspects of organizational climate which have been outlined by scholars interested in understanding individuals’ knowledge sharing (Bock et al. 2005). Individuals tendency toward knowledge sharing is mostly influenced by those climate facets that can be traced back to a tolerance of well-reasoned failure (Leonard and Sensiper 1998), and a freedom to make decisions (Kirkman et al. 2004).

Teams vary in the degree to which they are characterized by an autonomy climate, or decision-making authority for their actions (Bruhn and Gibson, 2006). Autonomy climate may be defined as “the degree to which the task provides substantial freedom, independence, and discretion in scheduling the work and in determining the procedures to be used in carrying it out” (Hackman and Oldham, 1980, p. 79). Autonomy climate refers to freedom, independence, and discretion in the task of a team (Cordery et al., 1991; Hackman, 1987; Kirkman and Rosen, 1999; Langfred, 2000). Organizations with centralized decision-making decrease (1) collaborative processes within the team as
communication increasingly flows vertically rather than horizontally, and (2) team members’ willingness to fully contribute their knowledge to problem-solving processes. On the contrary, team members in teams with a high degree of autonomy climate are reliant upon themselves for task decisions, which will likely increase the sharing of information, the coordination of task activities horizontally within the team (Hoegl and Parboteeah, 2006), and knowledge creation (Zellmer-Bruhn and Gibson, 2006). Moreover, autonomy climate strongly influences team learning because members in autonomous teams are encouraged to develop new ideas and/or to adapt ideas developed in other parts of the organization to fit their particular local context (Zellmer-Bruhn and Gibson, 2006). IT users need to acquire and/or develop IT and task-related knowledge to be able to use technology effectively in order to gain the planned objectives (Attewell, 1992; Nelson et al., 1995; Sein et al., 1999; Lassila and Branchau, 1999; Rogers, 2003; Sharma and Yetton, 2007). Two dimensions characterize users’ knowledge: an individual one, and an inter-individual one. The inter-individual cognition (Sharma and Yetton, 2007) is based on a transactive memory system (Argote 2005; Liang et al. 1995) and the development of collaborative task knowledge (Kang and Santhanam 2003-04). According to Wegner et al. (1985), transactive memory develops when individuals use the memories of other individuals as external storage, and it consists of (1) knowledge contained in the memories of individual group members, and (2) knowledge relevant communication processes among group members. Sharma and Yetton (2007, p. 221) follows Kang and Santhanam in defining collaborative task knowledge as “an understanding of the interdependent relationships among all users’ work procedures, enabling users to assess the collective consequences of their individual ways of using a collaborative application in the context of a business process, coordinating across individuals and contributing to group performance”. Autonomy climate supports the creation of these two dimensions of the inter-individual cognition through the fostering of information sharing, collaboration, knowledge creation, and team learning. Therefore, autonomy climate may enhance the development of IT knowledge integration capability. Formally,

\[ \text{Hypothesis 1: Increasing autonomy climate at team level will have a positive impact on IT Knowledge Integration Capability.} \]

Teams vary not only in terms of autonomy climate, but also in terms of innovative behavior. Teams may be innovative in the use of knowledge and technology if they incorporate the rule of agreement (i.e. non blocking the ideas of others) as a norm of their climate. The norm of agreement is critical for the creation of an experimental climate in teams, which is defined as a climate that provides room for experimentation and is tolerant of “competent” mistakes (Vera and Crossan 2004). Low levels of room for experimentation and low tolerance for mistakes represent climate that pursue efficiency over effectiveness and exploitation over exploration (Crossan and Hurst 2003). In contrast, high levels of possibility to experiment and tolerance for error are not associated with blind risk taking and lack of discipline, but represent a climate that promotes action as opposed to reflection as a way to understand and deal with reality (Cunha et al. 1999) and where boundaries and minimal constraints are defined so that experimentation can occur (Vera and Crossan, 2005). When team members perceive their environment as interpersonally nonthreatening and tolerant of, or even supportive of, taking risks and trying new approaches, higher levels of psychological safety and engagement in innovative processes. This means that experimental climate supports (1) creation, and new combination of knowledge, and (2) change in the technology and/or in the way of doing things by users. Indeed, technology is an effective knowledge medium and relates to the transformation of experimentation to performance (Walsh and Ungson 1991). This aspect is also corroborated by recent research outlined that team-level climate may affect the degree of experimental behaviors performed in interacting with technology for integrating knowledge (Maruping et al. 2008) Formally,

\[ \text{Hypothesis 2: Increasing experimental climate at team level will have a positive impact on IT Knowledge Integration Capability.} \]

Previous research has emphasized the importance of team-interna processes, and it has recently analysed processes between teams (e.g. Hoegl et al., 2004), but it has overlooked the analysis of team capabilities (Tiwana and McLean, 2005). In particular, we know little about the role of technological
capabilities in influencing both team effectiveness and efficiency. Resource-based view literature in the IS field has analyzed the role of technological capabilities, and the link between these capabilities and performance (e.g., Bharadwaj, 2000; Wade and Hullan, 2004), but the main focus of these studies is the firm level. They have overlooked the role of technological capabilities at team level. The second objective of this paper is to understand the relationship between IT knowledge integration capability and team effectiveness, and efficiency.

Teams conduct knowledge-based complex tasks (e.g. new product development) (Faraj and Xiao, 2006; Hoegl et al. 2003). Therefore, IT knowledge process management capability may positively influence (1) output effectiveness because it supports the use of knowledge within the team (Faraj and Sproull, 2000), (2) efficiency because it supports the storing, and the sharing of knowledge enhancing the team members’ ability to exploit process and procedures (Faraj and Sproull, 2000). IT knowledge integration capability as a combinative capability may enhance these positive effects. Formally,

Hypothesis 3: Increasing IT Knowledge Integration Capability at team level will have a positive impact on team effectiveness.

Hypothesis 4: Increasing IT Knowledge Integration Capability at team level will have a positive impact on team efficiency.

3 METHOD

3.1 Study context

Data were collected in two large European companies in the retail and insurance industries which introduced a new communication technology - voice over IP (VoIP). The technology was introduced to manage all technology-mediated communications among individuals in an integrated manner. The technology was needed to support activities such as agenda sharing, information sharing, mobility management, and event coordination. In addition to offering more information that can be accessed and managed by users, this system embodied the convergence of different communication capabilities, enabling individuals to communicate with their colleagues. This is particularly relevant because individuals, through a unique platform, are allowed to choose among different communication channels that match their synchronicity needs (e.g., voice, instant messaging, conference call, and e-mail). In this particular case, while the use of the system was strongly encouraged, there was no policy in place for non-compliance and no actions were being taken as a result of the usage reports, suggesting that system use was voluntary. Data were collected using a survey methodology. The questionnaire was developed using a multi-stage iterative procedure. First, an initial set of items was constructed drawing upon prior work that measured the constructs adopted in our study (Langfred, 2005; Vera and Crossan, 2005; Chan et al. 1997; Pearce and Sims, 2002. Next, we conducted interviews with the IT managers responsible for the implementation project. This helped ensure that the questionnaire was appropriate for the organizational setting and the technology introduced. One week before the launch of the survey, CIOs at the participating organizations sent an e-mail memo explaining the importance of the study to all potential respondents. This procedure has been already adopted in previous studies (Teo et al. 2003).

Data were gathered through a web survey containing five-point Likert-type scales. To obtain more reliable ratings of the team-level constructs under consideration, multiple respondents from each team participated: the team leader and at least three team members. To ensure data validity, only teams returning at least three questionnaires (the team leader and two team members) were considered. Of a total of 810 individuals and 129 teams targeted for the survey, 410 usable surveys referring to 69 teams were completed, yielding response rates of 50.6% (individuals) and 53.4% (team). The teams were functional long term teams involved in the commercial processes of their organization.
3.2 Measures

In order to obtain reliable team-level ratings for the variables in the study and to avoid potential common source bias, we collected responses from multiple sources in each team, including the team leader. Because some of the data from this team-level study were collected from multiple individuals within each team, it was necessary to justify the aggregation of individual-level within-team ratings to team-level scores (Klein and Kozlowski 2000; Rousseau 1985). This included a one-way analysis of variance (ANOVA) based on team membership to test the between-group variation, and the computation of ICC (Interclass Correlation Index) for assessing the stability of the team-level means (Bliese 2000).

**Team autonomy climate.** We used a five-item scale adapted from Langfred (2005) which captures the shared perception of the degree of autonomy that team has in the decision making process. Results of a one-way ANOVA indicated significant between-team differences in ratings of team autonomy climate \((F = 1.44; p < .05)\). The ICC was 0.79 indicating stable team-level means for this construct. A team-level score for autonomy climate was computed by averaging within-team responses to the scale.

**Team experimental climate** was measured through a three-item scale derived from Vera and Crossan (2005) tapping the shared degree that experimentation is allowed within the team. Results of a one-way ANOVA indicated significant between-team differences in ratings of team experimental climate \((F = 2.01; p < .01)\). The ICC was 0.68 indicating stable team-level means for this construct. A team-level score for team experimental climate was computed by averaging within-team responses to the scale.

**IT Knowledge Integration Capability** was measured adopting five items from Chan et al. (1997) related to IT capability and 3 items from Gold et al. (2001) for assessing the knowledge-oriented facet of the construct. The scale had a reliability of .84. Results of a one-way ANOVA indicated significant between-team differences in ratings of IT Knowledge Integration Capability \((F = 2.13; p < .01)\). The ICC was 0.58 indicating stable team-level means for this construct. A team-level score for this construct was computed by averaging within-team responses to the scale.

**Team effectiveness** was measured through four items derived from the output effectiveness scale by Pearce and Sims (2002) which assesses the quality of the output delivered by the team. The scale had a reliability of .84. Results of a one-way ANOVA indicated significant between-team differences in ratings of team effectiveness \((F = 1.78; p < .01)\). The ICC was 0.63 indicating stable team-level means for this construct. A team-level score for effectiveness was computed by averaging within-team responses to the scale.

**Team efficiency** was measured adopting a two item scale developed for the study in order to assess whether the team delivers its output respecting the budget and the deadlines. The scale had a reliability of .84. Results of a one-way ANOVA indicated significant between-team differences in ratings of team efficiency \((F = 1.35; p < .05)\). The ICC was 0.84 indicating stable team-level means for this construct. A team-level score for efficiency was computed by averaging within-team responses to the scale.

**Controls.** Following Hoegl et al. (2003), we included team size as a control variable. Larger team sizes have been associated with both increased and decreased performance. Larger teams are argued to give team members access to a broader array of resources. However, larger teams also create greater coordination complexity, thereby hindering the ability of individuals to collaborate and perform effectively. Since the teams involved in our study are characterized by varying degrees of task interdependence in team members’ work, the behavior of each team member has an impact not only on the effectiveness of that individual, but also on the effectiveness of the team as a whole (Griffin et al. 2007). Thus, we included task interdependence as control variable. Interdependence was measured using a three-item scale adapted from Campion et al. (1993). This scale had a reliability of .67. A one-way ANOVA revealed significant between-team variation on individual
ratings of this scale (F = 1.75; p<.01). The ICC for task interdependence was 0.80. Tenure was measured through an interval scale with the following anchors: 1= less than 3 years; 2=3 -5 years; 3= 6 – 10 years; 4= more than 10 years.

4 ANALYSIS AND RESULTS

We adopted partial least square (PLS) method to analyze the data. PLS is a structural equation modeling technique which use a component based approach to evaluate the relationship within, and variance explained by a structural equation model. PLS is a technique that is increasingly being used in IS research because it requires minimal sample size and it places minimal demands on residual distributions (Chin, 1998; Fornell and Bookstein, 1982). According to Agarwal and Karahanna (2000) data analysis process was divided in two steps. During the first phase we established the psychometric validity of the adopted measures, while in the second we tested the hypotheses.

4.1 Measurement Model Assessment

The psychometric properties of the scales were assessed in terms of items loading, internal consistency, and discriminant validity. As can be seen from the factor analyses results reported in table 1 all items loaded respectfully on their corresponding factor (Fornell and Bookstein, 1982). In order to assess the discriminant validity the average variance extracted (AVE) should be higher than the interconstruct correlations. As indicated in table 2 all the constructs share more variance with their indicators than other constructs. Moreover, as depicted in table 2, Cronbach alpha and composite reliability for each construct exceeds 0.7, confirming the internal consistency. These results are consistent with previous studies (e.g. Karahanna et al., 2006; Tiwana and McLean, 2005) allowing us to conclude that the measures testing the model all display good psychometric properties. Table 2 reports also correlations, descriptive statistics, and ICC indexes.
<table>
<thead>
<tr>
<th></th>
<th>Autonomy</th>
<th>Efficiency</th>
<th>Experimental</th>
<th>IT KMCap</th>
<th>Efficacy</th>
<th>Interdep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aut1</td>
<td>.581</td>
<td>.123</td>
<td>.434</td>
<td>.186</td>
<td>.121</td>
<td>.073</td>
</tr>
<tr>
<td>Aut2</td>
<td>.735</td>
<td>.255</td>
<td>.321</td>
<td>.206</td>
<td>.295</td>
<td>.066</td>
</tr>
<tr>
<td>Aut3</td>
<td>.776</td>
<td>.355</td>
<td>.376</td>
<td>.295</td>
<td>.350</td>
<td>.105</td>
</tr>
<tr>
<td>Aut4</td>
<td>.630</td>
<td>.104</td>
<td>.303</td>
<td>.247</td>
<td>.090</td>
<td>.205</td>
</tr>
<tr>
<td>Aut5</td>
<td>.739</td>
<td>.270</td>
<td>.363</td>
<td>.308</td>
<td>.274</td>
<td>.204</td>
</tr>
<tr>
<td>Efficiency1</td>
<td>.275</td>
<td>.858</td>
<td>.295</td>
<td>.296</td>
<td>.625</td>
<td>.073</td>
</tr>
<tr>
<td>Efficiency2</td>
<td>.278</td>
<td>.858</td>
<td>.320</td>
<td>.343</td>
<td>.584</td>
<td>.137</td>
</tr>
<tr>
<td>Efficiency3</td>
<td>.348</td>
<td>.269</td>
<td>.749</td>
<td>.303</td>
<td>.307</td>
<td>.173</td>
</tr>
<tr>
<td>Expc2</td>
<td>.458</td>
<td>.336</td>
<td>.909</td>
<td>.358</td>
<td>.362</td>
<td>.222</td>
</tr>
<tr>
<td>Expc3</td>
<td>.486</td>
<td>.285</td>
<td>.880</td>
<td>.328</td>
<td>.293</td>
<td>.226</td>
</tr>
<tr>
<td>IT KMCap1</td>
<td>.269</td>
<td>.326</td>
<td>.317</td>
<td>.821</td>
<td>.300</td>
<td>.249</td>
</tr>
<tr>
<td>IT KMCap2</td>
<td>.243</td>
<td>.261</td>
<td>.298</td>
<td>.807</td>
<td>.241</td>
<td>.269</td>
</tr>
<tr>
<td>IT KMCap3</td>
<td>.299</td>
<td>.377</td>
<td>.348</td>
<td>.827</td>
<td>.359</td>
<td>.237</td>
</tr>
<tr>
<td>IT KMCap4</td>
<td>.262</td>
<td>.284</td>
<td>.287</td>
<td>.810</td>
<td>.219</td>
<td>.227</td>
</tr>
<tr>
<td>IT KMCap5</td>
<td>.336</td>
<td>.315</td>
<td>.316</td>
<td>.853</td>
<td>.312</td>
<td>.258</td>
</tr>
<tr>
<td>IT KMCap6</td>
<td>.311</td>
<td>.229</td>
<td>.319</td>
<td>.718</td>
<td>.282</td>
<td>.213</td>
</tr>
<tr>
<td>IT KMCap7</td>
<td>.292</td>
<td>.196</td>
<td>.247</td>
<td>.694</td>
<td>.233</td>
<td>.211</td>
</tr>
<tr>
<td>IT KMCap8</td>
<td>.276</td>
<td>.193</td>
<td>.281</td>
<td>.683</td>
<td>.193</td>
<td>.234</td>
</tr>
<tr>
<td>Efficacy1</td>
<td>.313</td>
<td>.714</td>
<td>.309</td>
<td>.296</td>
<td>.871</td>
<td>.048</td>
</tr>
<tr>
<td>Efficacy2</td>
<td>.272</td>
<td>.748</td>
<td>.243</td>
<td>.232</td>
<td>.788</td>
<td>.004</td>
</tr>
<tr>
<td>Efficacy3</td>
<td>.293</td>
<td>.661</td>
<td>.332</td>
<td>.304</td>
<td>.885</td>
<td>.035</td>
</tr>
<tr>
<td>Efficacy4</td>
<td>.268</td>
<td>.662</td>
<td>.342</td>
<td>.309</td>
<td>.837</td>
<td>.139</td>
</tr>
<tr>
<td>Interd1</td>
<td>.161</td>
<td>.104</td>
<td>.212</td>
<td>.247</td>
<td>.103</td>
<td>.919</td>
</tr>
<tr>
<td>Interd2</td>
<td>.150</td>
<td>.024</td>
<td>.183</td>
<td>.251</td>
<td>.030</td>
<td>.739</td>
</tr>
<tr>
<td>Interd3</td>
<td>.190</td>
<td>.069</td>
<td>.216</td>
<td>.288</td>
<td>.039</td>
<td>.781</td>
</tr>
</tbody>
</table>

Table 1: Factor analysis

<table>
<thead>
<tr>
<th></th>
<th>Cronbach Alpha</th>
<th>CR</th>
<th>ICC</th>
<th>Mean</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Autonomy</td>
<td>.73</td>
<td>.82</td>
<td>.79</td>
<td>3.32</td>
<td>.65</td>
<td>.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Efficacy</td>
<td>.83</td>
<td>.90</td>
<td>.63</td>
<td>3.83</td>
<td>.66</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Efficiency</td>
<td>.76</td>
<td>.86</td>
<td>.84</td>
<td>3.57</td>
<td>.74</td>
<td>.78</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Experimental</td>
<td>.80</td>
<td>.88</td>
<td>.68</td>
<td>3.30</td>
<td>.82</td>
<td>.51</td>
<td>.38</td>
<td>.35</td>
<td>.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 IT KM Capability</td>
<td>.91</td>
<td>.92</td>
<td>.58</td>
<td>3.41</td>
<td>.73</td>
<td>.37</td>
<td>.35</td>
<td>.36</td>
<td>.39</td>
<td>.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Interdependence</td>
<td>.79</td>
<td>.86</td>
<td>.80</td>
<td>3.29</td>
<td>.85</td>
<td>.20</td>
<td>.10</td>
<td>.25</td>
<td>.30</td>
<td>.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Team size</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>9.39</td>
<td>5.8</td>
<td>-25</td>
<td>-10</td>
<td>-22</td>
<td>-25</td>
<td>-15</td>
<td>na</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Tenure</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>3.38</td>
<td>.97</td>
<td>-02</td>
<td>-05</td>
<td>-03</td>
<td>-09</td>
<td>-09</td>
<td>.06</td>
<td>na</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Correlations, descriptives and reliabilities. CR=Composite reliability.
4.2 Structural Model Assessment

According to Agarwal and Karahanna (2000), in a PLS analysis, “paths can be interpreted as standardized beta weights in a regression analysis”. As Figure 1 illustrates, team autonomy and experimental climate factors explain 19% of the variance in IT Knowledge Integration Capability, while IT Knowledge Integration Capability explains 13.4% of the variance in team effectiveness and 12.4% in team efficiency. Results provide support for hypothesis 1 stating a positive relationship between autonomy climate and IT Knowledge Integration Capability (coeff. = .23, p < .05). Team experimental climate has a positive effect on IT Knowledge Integration Capability, providing support to hypothesis 2 (coeff. = .27, p < .01). Hypothesis 3 and 4, which posit that IT Knowledge Integration Capability have a positive influence on effectiveness and efficiency are supported (coeff. = .35, p < .01 and coeff. = .38, p < .01, respectively). Since data collection at the same point in time and using the same instrument the potential for common method variance exists. We controlled for common method variance running a Harman’s one factor test, which demonstrated that common method variance was not a threat to our findings. Indeed: (1) the unrotated factor structure, explaining 69.6% of the variance, did not show a single factor; (2) the first factor did not account for all of the variance (38.9%) (Podsakoff and Organ 1986).

![Figure 1: PLS results. * p < .05; ** p < .01](image)

**5 DISCUSSION**

**5.1 Theoretical implications**

This research contributes to the literature on team climate, IT capability, and team performance in two ways.

First, our results show that team autonomy climate enhances the IT knowledge integration capability. This means that teams with high level of autonomy, in terms of freedom, independence, and discretion in scheduling the work and determining the procedures, may positively benefit the IT knowledge integration capability. Moreover, team experimental climate enhances the IT knowledge integration capability. Thus the existence of room for experimentation and tolerance of “competent” mistakes...
within the team, may positively influence the IT knowledge integration capability. Our paper goes beyond the traditional climate literature pointing out a direct effect of team level climate on team level capabilities. In particular, our research answers existing call for developing a better understanding of the contextual factors at the team level of analysis which may influence the way through which members exploit resources and technology for developing processes and procedures for acquisition, combination, creation, and sharing of knowledge (e.g. Tiwana and McLean, 2005). Thus, our examination of the drivers that affect IT knowledge integration capability represents an important step in overcoming the knowledge and capabilities at the team level of analysis.

Second, our results underscore the positive effect of IT knowledge integration capability on team performance (i.e. effectiveness, efficiency). Therefore, our results are worthy because we shed some light on the role of IT capabilities on team outcomes. Literature in the IS field has explored the role of IT capabilities, and the link between these capabilities and firm performance (e.g. Bharadwaj, 2000; Wade and Hullan, 2004), while scant research examined these issues at the team level. In particular we answer previous research calls that underscored a lack of knowledge about the link between IT capabilities and team performance (Tiwana and McLean, 2005).

5.2 Managerial implications

The results of this study have substantial implications for organizations that adopt technologies for supporting team activities. It has long been recognized that the introduction of new technologies is not enough for realizing gains in performance. Thus, to the degree that it facilitates the discovery of new sources of value for the technology, active experimentation and autonomy are desirable contextual factors.

Based on our findings about team climate, managers may consider creating a climate that is supportive for exploration and autonomy. Team structures that emphasize a tight integration of technology use into employee work practices. While previous research underscores the need to design team-based structures, our results point out that IT knowledge integration capability is influenced by a climate that facilitate information exchange and learning among team members.

5.3 Limitations and future research directions

As with any work our research has limitations that should be addressed in future studies. Because of the cross-sectional nature of the study we were unable to test for true causality, although causality is theoretically implied in some of the proposed relationships. A longitudinal study can provide some more relevant considerations and implications. Therefore, this study should be reiterated over time in order to catch the temporal effects of depicted variables. Moreover, we did not have access to the demographic data of non-respondents and were thus unable to verify the existence of any significant differences between respondents and non-respondents. Some issues for future research emerge from this study. Although the system we examined embodied characteristics that are common to other systems, future research should validate our results in other settings in order to increase the generalizability of our findings. Moreover, the results are based on the Italian context, suggesting the need for future research in other national and cultural settings.

5.4 Conclusion

In conclusion, the primary contribution of this work is the empirical validation of team level climate as factors influencing IT knowledge integration capability. From a theoretical perspective, a climate-based perspective provides new opportunities for extending the research on team capabilities related to
IT. Clearly, an understanding of the effect of IT knowledge integration capability on team performance is of significant importance to practitioners who are attempting to fully exploit the potential of new information technologies and team knowledge-related processes.

REFERENCES


Proceedings ECIS 2009


IT-ENABLED CHANGE INTO THE STRUCTURES OF HEALTH INFORMATION SYSTEMS IN AFRICA: A CASE STUDY IN KENYA

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0402.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Health information systems / medical record systems / care delivery /, IT artifact, Institutional theory, Developing countries</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
IT-ENABLED CHANGE INTO THE STRUCTURES OF HEALTH INFORMATION SYSTEMS IN AFRICA: A CASE STUDY IN KENYA

Bernardi, Roberta, Warwick Business School, University of Warwick, CV4 7AL, Coventry, UK, Roberta.Bernardi06@phd.wbs.ac.uk

Abstract

In the recent years most African countries have embarked on a series of reforms involving the decentralisation and also integration of health information systems in order to allow improved efficiency and effectiveness in the health care. However, although the discourse around these issues are reflected in global policy documents of almost twenty years ago, IS are still fragmented and weak at the lower levels of the health service. The paper takes a multivocal and multilevel institutionalist perspective to analyse the role of information technology in shaping these shortfalls between institutional accounts and enactments of reforms. Based on the case study of two divisions of the ministry of health in Kenya, it aims to better understand the change implications of information technology for the structures of a health information system in Africa. This is meant to improve the understanding of the way technology-mediated human interactions produce variance between planned organisational change envisaged in donor-driven reforms of the health care service and unplanned HIS structures emerging from the local institutionally-embedded usage of IT tools.

Keywords: Health information systems, IT artefact, Institutional theory, Developing countries
1 INTRODUCTION

Information technology is one of the key components in the implementation of health service reforms supported by international development partners. These reforms are meant to improve the performance of the health sector mainly through the decentralization and integration of health care. Health Information Systems constitute the main core for the implementation of these reforms. Firstly, decentralised information systems can provide decision makers at the lower levels with timely and accurate information for management and planning. Secondly, integrated health information can support coordinated and effective actions in health care delivery.

Still, there is evidence that public IT investments can lead to unexpected organisational impacts that are distant from the original plans or reform objectives, particularly, in developing countries (e.g., Ciborra 2005). Contradictions have been reported, for example, in the usage of IT in the implementation of New Public Management reforms (Hood 1991). Instead of reducing bureaucracies into flatter and more information efficient structures (Osborne et al. 1992), IT implementations within NPM have actually led to the fragmentation of information systems (Kimaro et al. 2005; Dunleavy et al. 2006) increasing, rather than reducing, complexity.

The divergent findings on the implications of IT in the implementation of public sector reforms (e.g., Ciborra 2005) raise the issue of the gap between the organisational impact of information technology as envisaged in formal restructuring plans and unplanned (or informal) practices or structures emerging from IT usage.

Likewise, although the political discourse on the decentralisation and integration of health information systems has engaged development policies for almost twenty years, health information systems are still fragmented and centralized, at least in Africa (Kimaro and Sahay 2007). Most research in this area has actually found that major differences across health information systems depend on competing rationalities underpinning donor-driven reforms, public administrators’ procedures and system users’ practices (Chilundo and Aanestad 2004), bearing strong connections with the “resource-deprived” environment (Kimaro and Sahay 2007). Yet, although past research in health information systems in Africa and other developing countries has clarified the mismatch between formal triggers of change (e.g., public sector reform policies) and informal constraints to change (e.g., professional roles, social expectations, IS users’ routines) (Kimaro and Sahay 2007), little is known about the role of the IT artefact in relation to this formal-informal divide.

The case study illustrated in this paper is meant to address this gap by answering the question of how the IT artefact influences the structures of health management information systems in an African country in relation to its institutional and resource-material environment. The research question rests onto the assumption that the institutional environment both shape and is shaped by users’ enactment of IT (Orlikowski and Barley 2001). The mutual relationship between the institutional environment and IT enactment also accounts of the influence of institutionally-embedded resources (Scott et al. 2000), including, not only financial and human resources but also the technical properties of the technology.

Thus, the proposed case study takes a multilevel and multivocal institutionalist perspective to analyse how new and old rules, norms, and cultural meanings influence technology innovation and the usage of IT in information processing within two national health information systems of the Ministry of Health in Kenya.

The theoretical perspective used is “multivocal” as it takes into consideration the complex interlinkages between varied rationalities and sets of meanings (Lounsbury 2007) underpinning the actions of different actors (e.g., international donor partners, national ministries, public employees, etc.) involved in the reforms of the health systems in Africa (Kimaro and Sahay 2007). In particular, the proposed perspective recognizes two main logics: the “New Public Management” (Hood 1991) embedded in imported reform models and the “Old Public Administration” (Lynn 2006) of traditional African bureaucracies. By considering institutions as enablers of “entrepreneurs action” under an
institutional entrepreneurship perspective (Dacin et al. 2002), the clash between different logics may cause the “translation” (Zilber 2006) of new imported practices into localised “hybrids” (Hood 2000) depending on the modalities of adaptation of the new model (e.g., New Public Management).

The frictions between multiple rationalities and processes of hybridizations can be better understood through the “multilevel” perspective which views the context of African Health Information Systems as characterized by three levels of action: the macro or policy level (e.g., formal policies), the meso or organisational level (e.g., management structures) and the user or agentic level (e.g., IS users’ routines). The three-level perspective allows to zoom in onto the main sets of meanings underlying international and national development policies at the macro level. It highlights how these are mediated or moderated (Dada 2006) by organization level structures, including for example, the influence of different professional norms (e.g., medical doctors, data managers, etc.), and, eventually, interpreted and enacted by institutionally embedded IT users at the micro-level.

In addition, the enactment of institutions at different levels is associated with the means enabling the mobilization of interpretive resources of the institutional environment (Sewell 1992). These means are represented by the technical features of the technology and also the physical and human resources available. In the first instance, the degree of “interpretive flexibility” or “malleability” of IT designs (Orlikowski 1992) influences may either facilitate or constrain new meanings and practices, thereby, reshaping the institutional order of the public health sector.

Under this perspective, the case study presented in this paper provides a historical longitudinal analysis of change opportunities (or constraints) enabled by the institutional environment and by the resource-material environment (IT technical features, human resources, information tools). This has been achieved by focusing upon changes into the information processing behaviour of IS users following the introduction of a specific technology. In particular, the analysis has evidenced the way they use the technology to process information by taking into account its technological properties and how they make sense of it (Orlikowski 2000; Davidson 2006) by drawing from rules, norms and beliefs of a specific institutional context (Powell et al. 1991; Barley et al. 1997).

The empirical lens adopted draws from the information processing view of organisations, whereby the designs of organisations is associated with their information needs (e.g., Galbraith 1977) and the information flows between the different parts of an organisation (Mintzberg 1979). The main assumption here is that changes in IT-enabled information processing practices reflect changes into the configuration of the organization.

Acknowledging the misalignment between formal policies and designs and unplanned or informal outcomes of their implementation (Piotti et al. 2006), the case study seeks to provide a more systematic understanding of the interplay between top-down pressures (e.g., donor-driven development policies) and bottom-up responses (e.g., data management practices). In particular, by explaining those institutional mechanisms giving rise to multiple rationalities and how these are enacted at different institutional levels, it will clarify the processes influencing the adoption and usage of IT in the restructuring of health information systems.

2 MULTIVOCAL AND MULTILEVEL INSTITUTIONALIST FRAMEWORK

The proposed theoretical framework is characterised by three levels of analysis: macro or policy level, meso or organisational level, and micro or agentic level (Figure 1). For each of these levels, different institutions have been classified under three major institutional pillars: regulative, normative, cultural-cognitive (Scott 2001). This classification allows to better understand what are the major institutional mechanisms influencing the encoding and enactment of typified behaviours by IT users. In particular, the framework views cultural cognitive institutions as informing all the others (Scott 2004).
Figure 1. Multilevel institutionalist framework

The macro or policy level focuses on the influence of different actors (e.g., Governments, donor agencies) in the diffusion of new public sector models and technologies (Ciborra et al. 2005). At this level the main institutional mechanisms are “legal and authority systems” including the international pressure carried through donor funding (regulative pillar), “Civil service norms” (normative pillar) (Scott 2001), and the two main sets of meanings (“cultural-cognitive pillar”) characterising the multivocality of the public sector in Africa: the New Public Management embodied in the reforms (Hood 1991), and the Old Public Administration (Lynn 2006) of local public service models (Table 1). Institutional logics at the macro level influence the institutional elements at the meso or organisational level encompassing governance and management systems, information processing mechanisms, human and technological capacity (regulative pillar), norms and legitimacies dictated by the professional field of the organisation (e.g., health) (normative pillar), patterns of typified information

Proceedings ECIS 2009
processing behaviours (or “scripts”) encoded in different institutional logics (NPM or OPA) (Barley, 1986) and influenced by the material properties of the IT tool (e.g., software source) and institutionally-embedded enactment of the IT tool (“IT-in-practice”, Orlikowski 2000) at the micro-level.

<table>
<thead>
<tr>
<th>NPM Logics</th>
<th>OPA Logics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Managerialism</strong></td>
<td><strong>Bureaucracy and politicisation</strong></td>
</tr>
<tr>
<td>Increased responsibility and decision-making power at managerial level (disaggregation, agencification)</td>
<td>Decision-making is concentrated at top of hierarchy</td>
</tr>
<tr>
<td>Devolution (decentralisation of power to the periphery)</td>
<td>Input controls, rules and procedures</td>
</tr>
<tr>
<td>De-politicisation of implementing structures and functions</td>
<td></td>
</tr>
<tr>
<td><strong>Accountability</strong></td>
<td><strong>Political affiliation</strong></td>
</tr>
<tr>
<td>Result and performance-oriented management system</td>
<td>Political rewarding system</td>
</tr>
<tr>
<td><strong>Market</strong></td>
<td><strong>Monopoly</strong></td>
</tr>
<tr>
<td>Competition</td>
<td>Internalisation of service delivery (insourcing)</td>
</tr>
<tr>
<td>Externalisation of the public service to market</td>
<td>Weak market economy</td>
</tr>
<tr>
<td><strong>Customer-orientedness</strong></td>
<td><strong>Politicalisation of service</strong></td>
</tr>
<tr>
<td>Responsive, diversified and “exclusive” service</td>
<td>Central (external) political control</td>
</tr>
<tr>
<td>Customer identity of beneficiaries</td>
<td>Public service complies with</td>
</tr>
<tr>
<td></td>
<td>international/national policy priorities</td>
</tr>
<tr>
<td></td>
<td>Bureaucratic/administrative public service delivery</td>
</tr>
</tbody>
</table>

Table 1. Institutional logics of the New Public Management and Old Public Administration

At this level, IS users mobilise IT resources by drawing from this set of institutionalised patterns of information behaviour encoding institutions informed by either NPM or OPA institutions. Depending on how these are sensed (Weber and Glynn 2006), IT users build their perceptions and assumptions on information needs (e.g., uncertainty and equivocality reduction, Daft et al. 1981) and their choice of the best actions, processes and structures to meet them (Malhotra 2000). Hence, by grounding the information processing view of organisations within an institutionalist perspective, it is assumed that the information behaviour of individuals depends on how they legitimate their information needs over time under specific point-in-time institutional and technological pressures. This, in turn, influences their choice of which patterns of information behaviour or scripts to enact. Typified behaviours or scripts are viewed as encoded in policy or macro-institutional mechanisms, organisational structures, and technical features of the IT tools available. The revision of typified behaviours may destabilise the institutional order (Chreim, Williams et al. 2007) giving rise to hybrids that inform new organisational structures.

3 THE CASE STUDY OF TWO HEALTH INFORMATION SYSTEMS IN KENYA

3.1 Methodology

The case study consists of two units of analysis: the central Division of Health Management Information Systems (D1), and the Division of Vaccines and Immunisation (D2). The divisions represent two of the units of analysis of the case study of three health management information systems in Kenya including the HIV/AIDS information system of the National AIDS Programme. Data were collected during a field visit at the Ministry of Health in Kenya. The case study comprehends a sample of thirty-eight semi-structured interviews (one hour average each), 13 from the Division of Health Management Information Systems (D1); 11 from the Division of Vaccines and
Immunisation (D2), and the rest from the National AIDS Programme, which is not illustrated in this paper.

Staff members with the earliest date of deployment were selected in order to capture the longest longitudinal accounts possible. Moreover, the sample included not only health records officers, the direct users of the information system, but also medical management and technicians, i.e., the indirect users of the information system. The inclusion of different organisational roles into the sample of informants was part of a “comprehensive sampling” strategy (Miles & Huberman, 1994, p. 38) in order to have a more systemic view of the evolution of the information system and technological changes.

The questionnaire was structured according to the theoretical framework in Figure 1. Questions were asked regarding their main functions in the division, changes in their working practices, professional values, technological changes and institutional reforms. Interview data on public sector reforms were also triangulated with main policy documents of the Government of Kenya and Ministry of Health covering a period from 1990s up to date. Relevant international agencies’ policy and project documents were also retrieved online.

Data analysis was also framed within the theoretical framework of Figure 1. Interviews were first transcribed and coded in NVIVO 8. Starting from a pre-defined set of codes for each level (macro, meso, and micro) and pillar (regulative, normative, and cognitive), new codes were created along the process of data analysis (Miles & Huberman, 1994 p. 58). Time-series codes were also created starting from 1980 with a five-year interval. After coding each set of interviews for one division, codes were reviewed and similar codes were grouped under common categories (“pattern coding”, Miles & Huberman, 1994, p. 69). As a consequence, all data from the previous division had to be reviewed and analysed again to take into account the new concepts.

### 3.2 Overview of context

The two divisions have developed separate information systems, each characterized by different technological, organisational and policy histories that have affected their evolution. In the last two decades, the health care system in Kenya has gone through a series of public sector reforms mainly under the pressure of multilateral financial institutions such as the WB and the IMF (e.g., “Structural Adjustment Programmes”; “Poverty Reduction Strategy Paper”). In reaction to these reforms, the Ministry of Health issued the National Health Policy Framework (GOK 1994) followed by two major five-year implementation plans in 1999 and 2005 (MoH 1999; MoH 2005). The health sector reform had among its objectives the decentralisation of sector management to the lower levels of the health systems (e.g., districts) and the integration of the health service packages for a more cost-effective and accessible health care.

Thus, the implementation of the reform required the creation and strengthening of health management information systems at the provinces, districts, and health facilities. It is also supposed to involve the merging of health information systems scattered across the various vertical health programmes (e.g., HIV/AIDS) into the central health management information systems of the Ministry of Health.

It followed that in the 80s the Ministry started decentralising the local planning of the health care provision to the districts (GOK 1994) including the creation of health management information systems in district offices. However, apart from this, not much has been achieved in the implementation of the reforms. On the contrary, programmes have developed separate information systems exacerbating the lack of integration and coordination across programmes. This did also negatively affect the decentralisation of health information systems by overloading reporting systems of districts and health facilities.

Since 2005, however, the introduction of result-based management and monitoring and evaluation systems has given a new boost to the decentralisation and integration of health information systems (NSSSP II 2005-2010, MoH 2005). These reforms represent the apex of the restructuring of the public
sector under the logic of the NPM, marking the formal shift from a bureaucratic output-based to a managerial outcome-based approach. As a result, more emphasis has been put onto an integrated well functioning Health Management Information System at all levels. This is supposed to work as a key component of the monitoring and evaluation system of the health service.

3.3 Case study analysis

3.3.1 Division of Health Management Information Systems (D1)

The HMIS Division (D1) is responsible for the collection and analysis of routine health data (e.g., outpatient morbidity, inpatient morbidity and mortality, service utilization, hospital administrative statistics, etc.). Data are managed through Clarion a dos-based data base management system developed in the early 90s through donor assistance. However, given the low usability of the system and the lack of specialised assistance for maintenance, in 2004 most data were migrated to Excel for data analysis. Due to the low data storage capacity of Excel, the only data still stored in Clarion are the in-patients which need to be exported to Excel sheets for the analysis.

Thus, after more than fifteen years Clarion is still in place, despite the fact that health information officers in the division feel quite uncomfortable to work with an obsolete dos-based system that they consider unfriendly and cumbersome. Moreover, the technical features of Clarion (incompatible with Windows- and web-based platforms) have limited its diffusion in the health service and, consequently, its legitimization. This system could not be easily and cost-effectively adopted by the lower levels of the health system posing limitations to the decentralization of the division’s information processing structures. As a consequence, apart from a few districts sending their reports in Excel sheets, most data entry is still done at the central level enacting the OPA cultural-cognitive institution of “centralised information processing”.

This is mainly due to the lack of legitimacy of the division within the health system (“macro-normative”). In fact, the contradictory views over the nature of the mission and tasks of the division in the health system caused no little confusion in the governance arrangements of the division (meso-regulative) with serious implication for its financial support (macro-regulative): “we were under preventive and promotive, then I realised another time we were told no […] we should report to […] planning cause you provide information for planning […] that is why even when we asked for support we could not be supported”. It is only after the introduction of performance contracts with the public sector reforms in 2005 and the official recognition of HMIS as the Ministry’s Monitoring and Evaluation system that the division’s staff could rely on a stronger institutional back-up, a new governance structure with a clear mission legitimized by the top-ranks of the Ministry.

Since then, the management and staff’s decisions and actions have also been characterized by the logics of managerialism and accountability triggering the strengthening of its information processing capacity. Only recently have the provinces and districts been equipped with new IT tools, which could bring to a partial decentralization of its information processing structures. However, this has not happened with government funding, as one could have expected, but still with donor resources.

Hence, the formal acknowledgement of the role of HMIS by the government reforms has not been followed by a substantial legitimization of the division (“Rhetoric of change”). Its mission is still output-oriented and passive with respect to the new overall vision and mission of the Ministry of Health aiming towards a more efficient and inclusive health service The lack of support and legitimacy at the macro level and the absence of a strong and focused mission at the meso level hindered the entrepreneurship of health records and information officers and the medical leadership in their attempts to strengthen the information system.

Management and health information staff are still struggling to gain major legitimacy and attract more government support notwithstanding the more enabling policy environment. Under these circumstances, they have been unable to put into place technological improvements and coherent information processing mechanisms and procedures (e.g., different data entry processes per system)
that could have streamlined the division IS through a stronger integration across central and peripheral levels of the health information system. Hence, the information processing structures of the division are still centralised and cumbersome.

3.3.2 Division of Vaccines and Immunisation (D2)

Like in the division of HMIS (D1), in the Division of Vaccines and Immunisation (D2), computerisation happened in the early 90s with the adoption of CEIS (Computerised Epidemiological Information System), a dos-based system developed and implemented through donor assistance. Like Clarion, CEIS was very cumbersome to use especially because of the lack of programmers that could fix its bugs and upgrade it to accommodate new information requirements. However, towards the end of 2003 the division experienced a more radical technological shift with the upgrading of the old system into Epi-Info, a more user-friendly, self-sufficient Windows computerised system, compatible with a web-based working environment. This innovation enabled a major integration between the central information system and data management offices at the periphery.

This technological shift is the result of an institutional process that started in the 80s when the division was created under the international mandate of the World Health Organisations Expandable Programme on Immunisation (WHO 1989). Thus, based on the major legitimacy of its mission in the global health strategies (macro-normative), the division was supported by conspicuous donor funding. Informed by the OPA logic of “politicisation of service”, the exogenous pressures of donor funding triggered a series of actions and events at the micro level, that, by contrast, were mainly informed by the NPM logics of “accountability” and “customer-orientedness”.

One action in particular was the initiation of a performance-based grant programme of immunisation at the beginning of the year 2000. This constituted an incentive for the division to improve its routine monitoring systems to provide data evidence that could award them with further funding. Thus, under the NPM logic of “managerialism”, the management of the division engaged in a series of activities to streamline the information system. It started by asking support from international partners to design EPI-Info, a new Windows-based network system. By the end of 2003 EPI-Info was rolled out to the provincial offices which started performing data entry. In this way, the workload of processing information was more evenly distributed between the centre and the provincial offices (“partial decentralisation of information processing”). Later on, policy guidelines on data management were also introduced at all levels bringing more consistency in routine reporting.

Although the latest improvements have happened under the pressure and with the help of international partners, the need to attract donor funding does not seem to have been the major reason behind it. In the last decade, most donor funding has been diverted from immunisation to other international health priorities (e.g., HIV/AIDS). Hence, while in 1999 donor contribution to the routine immunisation system was 83% against the 17% from Government budget, in 2001 Government financing equalled 53% of total expenditures after the Government increased considerably its contributions to vaccines procurement (WHO 2001).

This means that besides the attractiveness of donor funding, the structural improvements in the health information system that followed the development and implementation of EPI-Info rested onto the new legitimation that immunisation and disease surveillance data acquired within the country. Aware of this, management sought as much support as they could to improve their monitoring and evaluation systems in order to provide evidence of the division’s achievements and needs to secure funding from the Government budget.

Hence, the Government ownership (and legitimacy) of national immunisation initiatives facilitated the “translation” (Zilber 2006) of the value of health information for planning instilled by international partners. In this way, the management of the division was facilitated with respect to other departments such as HMIS (D1) in the implementation of the Second National Health Sector Strategic Plan (2005-2010).
Hence, the logics of accountability and customer-orientedness have been the major triggers towards the adoption of a new computerised system for the restructuration of the health information system following the legitimacy of the division in the national health service. These logics survived the considerable withdrawal of donor contribution. In contrast, they were reinforced by the need to compete for the meagre government resources. The appropriation of these endogenous logics into the national immunisation services allowed a more consistent action of management towards an efficient and effective M&E system within the framework of the new reforms.

4 DISCUSSION OF FINDINGS

Findings show that the adoption of Information Technology has led to distinct institutional shifts and organisational outcomes. In the Division of Health Management Information System (D1) cultural-cognitive institutions informing central information processing still persist in contrast with the Division on Vaccines and Immunisation (D2) where the process of decentralisation of their information systems has been more effective. The different impact of the technology onto the structures of the health information systems of the two divisions is due to both institutional and the embedded material resource environment, namely, the technical properties of the technology and available human capabilities.

Although decentralisation policies started far back in the late 80s, the decentralisation of health care delivery has never been fully achieved. The case study shows how different approaches towards technological change and the usage of IT in the two divisions have actually produced opposite outcomes in the decentralisation of the structures of two health management information systems informed by the same policy environment.

In particular, in the first division the weakness of endogenous macro-policy institutions fails to translate the new logics of the NPM informing the reforms into the local setting. The mediator function of the meso-level institutions in maximising the benefits of IT-enabled change is thus hindered. The initiative of management and records personnel did not find the necessary support at the policy level lacking access to the resources necessary for the enactment of the new logics. Although the normative pressures at the meso level from the health records profession to improve the efficiency of the HIS might have identified the need for new IT systems and information processing practices, the lack of a viable institutional, and consequently, material resource environment has prevented these demands from being turned into a necessity and a clear response action.

In contrast, an increased commitment by the central Government in the provision of immunisation services brought about the necessary legitimacy of the Division of Vaccines and Immunisation (D2) in the national health system. The new legitimacy instilled a stronger motivation by management towards the enactment of imported logics informing donor-driven monitoring and evaluation systems into the local context. This motivated management towards the employment of information technology for the partial decentralisation of the information system. Management has been an important catalyst of change towards this direction by focusing on the provinces as the strategic layer of coordination of the health information system and by standardizing data collection practices. Hence, although donors were instrumental by providing the financial and technical resources necessary for the set up of the new computerised system (Epi-Info), the main trigger of the IT-enabled decentralisation of the information system came from endogenous pressure to acquire national legitimacy that would secure Government funding.

Although the alignment between political administrative, management and user levels (Kimaro & Sahay 2007) is important for attaining results, this alignment is easier to achieve when there is a strong support and legitimation from the political administrative level. Institutional forces at the macro-policy level are the main triggers of IT-enabled organisational change. This change is as much consistent and effective as the main cultural-cognitive institutions informing it are supported by endogenous institutions at the macro level.
Finally, further important considerations can also be drawn from the material resource environment, specifically, in relation to the technical features of the IT systems. Epi-Info (D2) and Clarion (D1) were both exogenous technologies of western origin introduced to improve efficiency (timeliness) and accuracy (completeness) of information (Health reports). However, this happened in two different and distant periods of time with different technological advancements. In the early 90s, the technical properties of information technologies did not enable easy networking and communication channels for the integration of information processing. Moreover, being Dos-based, its lower level of usability and user-friendliness implied high deployment costs to the data management offices at the periphery. Lastly, off-the-shelf data management packages were not as diffused as today (e.g., Excel) and the development and enhancement of integrated application suits for data storage and analysis, especially, with poor in-house capacity, was quite demanding. Given the technological advancements of the time, Clarion was among the best technologies to achieve “timeliness and completeness of reporting” by enabling fast data entry and increased data storage capacity at the central IS office.

Hence, although the institutional pressures and sources for change are more or less the same for both systems (donors want more efficient M&E system “accountability”) and such institutional pressures or mechanisms encode the same meanings (“cultural-cognitive institutions”) of change, the enactment of these meanings is influenced by material technological environments (entailing different technological properties) spanning across two different technological eras. Eventually, this gives rise to two different structural and institutional outcomes (centralisation vs. partial decentralisation) and of course performance outcomes.

Likewise, what is it that turns the users’ perceptions of Clarion as an efficient system of the 90s to an outdated and cumbersome system of the year 2000? Most likely, the trigger of a change of cognitive framing (Orlikowski et al. 1994; Davidson 2006) of Clarion has been the exposure towards more advanced technologies. This must have increased the awareness that there are more efficient ways of processing data through IT tools. Hence, it can be inferred that the major factor influencing the changes in the framing of Clarion has been the evolution of the technological environment and the availability of more advanced data analysis features. This has contributed to a more limited usage of Clarion and, on the other hand, a more frequent usage of Excel, which contributes to the enactment of different IT-enabled IP behaviours (“IT-in-use structures”, Orlikowski 2000) and also distribution of tasks.

5 CONCLUSIONS

The case study has shown that the impact of information technology onto the structures of a health information system in Africa is mainly influenced by the support of endogenous macro-level policy institutions and resources. These stimulate the enactment of exogenous cultural-cognitive institutions informing donor-driven reforms by management and key professions at the meso-organisational level towards the usage of technology for sustainable change. Moreover, the material resource-environment represented by the institutionally-embedded technical properties of IT and relevant human expertise is instrumental for the translation of exogenous logics informing change. Hence, the IT artefact can be used to reduce the gap between formal planned structures and informal routine or practices (Kimaro & Sahay 2007) of IT users, if new rationalities embedded in imported reform models are translated to the local context through stronger Government commitment at the macro/policy level. This can facilitate the task of management and key professional staff in their effort to leverage information technology for the realignment between policy discourses at the macro level and their enactment at the micro-level.

The material resource environment plays a considerable role in shaping the institutional forces underpinning IT innovation and usage. Donors’ resources are still instrumental in the transfer of new technologies and accountability systems to the health systems of countries like Kenya. In fact, decision makers within Governments are not familiar with the usage of information technology as powerful tools for planning and monitoring the health system. Moreover, specialised solutions such as
computerised epidemiological systems are more available from the global market than the local market.

Management and professionals at the meso level of the public administration can reduce the gap between global and local “technological frames” (Orlikowski and Guah 1994, Davidson 2006). Being more exposed to international experts, and new innovations, and building onto past experience of IT investments, they can master IT-enabled transformations for a sustainable and effective restructuring of the health information system.

Moreover, the technological environment, in terms of different technological properties, also contributes to the definition of institutional pressures and management and IT users enactments. Hence, expectations of structural changes and performance depend onto the opportunities offered by technological advancements. In the same way, the enactment of same logics (e.g., efficient data processing) can lead to different structural outcomes depending on the material properties of the technology and the influence of the material technological environment (e.g., dos-based systems are kept for centralised systems, while Windows- and Web-based technologies are introduced to decentralise systems). These considerations offer a contribution towards a better conceptualisation of the IT-artefact in the human enactment of information technology.

References


Eric Montiero was the accepting Associate Editor for this paper.


## Microemancipatory Practices in Information System Development

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0250.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Critical theory, Information Systems Development (ISD), IT-enabled social innovations, Socio-technical theory</td>
</tr>
</tbody>
</table>
MICRO-EMANCIPATORY PRACTICES IN
INFORMATION SYSTEM DEVELOPMENT

Cecez-Kecmanovic, Dubravka, School of Information Systems, Technology and Management, Australian School of Business, UNSW, Sydney NSW 2052, Australia, dubravka@unsw.edu.au

Janson, Marius University of Missouri-St. Louis, College of Business Administration, Department of Information Systems, One University Boulevard, St. Louis, MO 63121, USA, Email: janson@umsl.edu

Abstract

The paper revisits the notion of emancipation in Information System Development (ISD) that seems to have lost a battle against functionalist and managerialist approaches dominant in information system (IS) research and practice. Unlike functionalist and managerialist views, the emancipatory view of ISD, informed by Critical Theory, considers ISD as a site of organizational innovation, self-reflection and a struggle for humanization of work and liberation from different forms of domination. Critics of emancipatory project in IS and management literature question the very possibility of the emancipation and deplore its intellectualism, naivety and negativism. The purpose of this paper is to re-consider the notion of emancipatory ISD in the face of these criticisms and develop a more refined and nuanced view of micro-emancipation in ISD that is meaningful in practice. Informed by Alvesson and Willmott (1992, 1996) we explore, question, redefine and ground the micro-emancipatory ISD processes based on a longitudinal (15 year) study of a retail company. Our analysis and critical reflection demonstrate that micro-emancipatory ISD processes have real substance for the people involved, and that their meanings are neither fixed nor universal, but rather local, emergent, uncertain, and sometimes contradictory. This paper contributes an empirically grounded and practically relevant reconceptualization of micro-emancipatory ISD projects which reveals both its benefits and risks for all involved.

Keywords: Micro-emancipation, Information Systems Development (ISD), Emancipatory Practices in ISD, Emancipatory ISD
1 INTRODUCTION

An emancipatory approach to information systems development (ISD) was proposed to counteract an increasingly narrow and functionalist focus on organizational productivity and efficiency dominant in both the mainstream information systems (IS) research and IS practice. Informed and inspired by Critical Theory several prominent researchers exposed new forms of control and domination emerging through implementation of IS disguised under the ‘objective’, ‘impartial’ and ‘scientific’ ISD methodologies (Weizenbaum, 1976; Klein and Lyytinen, 1985; Lyytinen and Hirschheim, 1985, 1988; Lyytinen and Klein, 1985; Lyytinen, 1992; Myers and Young, 1997). An alternative view of IS as social communication systems that have a potential to freeing employees from ‘repressive social and ideological conditions and thereby contributing to the realization of human need’, was proposed by Hirschheim and Klein (1994, p. 87). They argued for an emancipatory ISD methodology that assumes a systematic and meaningful user involvement, open and non-distorted communication, and reasoned argumentation based on cooperation and mutual understanding among IS developers, managers and employees affected by the system.

However, emancipatory ISD ideas and approaches have been dismissed as unrealistic and naïve, incapable of addressing real-life power struggles or preventing colonizing effects of IS (Wilson, 1997). More broadly the emancipatory project of Critical Theory as such and its application in management (including IS) have been criticised by poststructuralists for implying a rationalist bias, for harbouring essentialist assumptions about unified, autonomous subjects and for imposing all-embracing frameworks or ‘metanarratives’ that reduce the complexity and heterogeneity of phenomena (Lyotard, 1984; Calas and Smircich, 1987).

While much of these criticisms, primarily from poststructuralist quarters, arise from a considerable misunderstanding of Critical Theory and its concept of emancipation – as demonstrated by Alvesson and Willmott (1996) – some are justified and worth heeding. Acknowledging the disillusionment with grand programme of emancipation, they propose a more modest concept of micro-emancipation that describes partial, temporary and precarious forms of liberation and resistance to oppressive organizational practices. Their proposal, however, has not been explored further in the literature. In the IS literature in particular, the emancipatory ideas and ideals seem to be forgotten and the whole debate about their meaning, reality, and relevance relegated to history.

In this paper we aim to revive the debate around the notion of emancipatory ISD. More specifically we aim to revise the concept of emancipatory ISD as proposed by Hirschheim and Klein (1994) and explore its relevance in the face of major criticisms in the literature. Inspired by the Alvesson and Willmott’s (1992, 1996) proposal for examining micro-emancipation projects we aim to develop a more refined and nuanced view of micro-emancipation in ISD that is meaningful in practice. While ISD projects are typically firmly driven by desired performance improvements (efficiency, effectiveness) we suggest they can also be seen as sites of organizational transformation where status quo is questioned and existing forms of control problematized and resisted, and where workplace conditions are challenged.

We achieve our aims by first presenting a short literature review and then a longitudinal study (15 years) of an IS development in a Belgian retail company. In the following section we trace and critically investigate the development practices of their most important IS (a corporate information dissemination and groupware system they call ISID. Based on this analysis we develop a framework that defines a refined and empirically grounded conception of micro-emancipatory ISD processes and practices. In the Conclusion we discussed its contribution and implications for the theory and practice of ISD.
2 EMANCIPATORY IS DEVELOPMENT IN THE LITERATURE

Emancipatory ideas in ISD have their origins in the Scandinavian participatory design approach and the socio-technical design movement in the UK in the 1960s. In response to the increasing use of information technologies (IT) serving an economic rationalist agenda at the expense of further bureaucratization and dehumanization of work, participatory design and socio-technical principles were based on humanist ideals, workers autonomy and workplace democracy. They argued for the use of technology to achieve both efficiency objectives and improved ‘quality of working life’ (Mumford and Weir, 1979; Bjerkins et al. 1987; Bodker et al., 1987; Bjerknes and Bratteteig, 1995; Mumford, 1983, 2000, 2006). Both these approaches assumed an underlying belief in technological progress and human knowledge and enthusiasm for computer applications that would replace boring, repetitive and dehumanizing jobs, increase job satisfaction and thereby eliminate workers’ alienation. Proliferation of projects that adopted and advanced participatory design in Norway, Sweden, Denmark – referred to as the collective resources approach – together with projects that adopted the socio-technical design in the UK, followed by other European countries, Canada and the USA, in the 1970s and early 1980s, raised hopes in the democratizing potential of IT.

However, their promises of humanization of work, workplace democracy and workers empowerment to make design and work-related choices were not long lived. The economic pressures in the late 1980s and 1990s and the raise in unemployed labour changed market and employment conditions leading to the revival of computer-aided neo-Taylorism (Moldaschl and Weber, 1998). The deployment of IS to cut costs, downsize workforce, increase managerial control, and achieve lean and efficient production based on standardized work processes, went counter to socio-technical and participatory design principles and practices. Socio-technical design, as Mumford (2006) conceded, ‘moved from success to failure’ as ‘[t]he attraction and validity of bureaucracy was seen as stronger and safer and the new humanistic approaches as over-risky’ (p. 321). Researchers are raising their critical voices against the narrow view of IS as a means of furthering economic rationalist agendas, the view that obscures repressive social conditions and ‘the continued destruction of the human potential’ (Asaro, 2000; Saravanamuthu, 2002; Howcroft and Wilson, 2003).

A distinctly critical approach to IS development was influenced by Critical Social Theory (e.g. Lyytinen and Hirschheim, 1985; Lyytinen and Klein, 1985; Cecez-Kecmanovic et al., 2002; Alvarez, 2008). Hirschheim and Klein formulated four conditions for an emancipatory IS development methodology (1994, pp. 87-88): 1) Providing support for an active process of individual and collective self-determination; 2) Providing support for critical self-reflection and associated self-transformation; 3) Inclusion of a broader set of institutional issues relating particularly to social justice, due processes and human freedom or more concretely in ISD to employees’ ethical needs, quality of work life, personal autonomy, and the linkage between participation and democracy; and 4) Inclusion of the principle for critical evaluation of claims or rational discourse during the systems development processes.

The emancipatory ideas in ISD and more broadly in management, however, have been criticised for neglecting the reality of business conditions, the pre-eminence of shareholders’ interests and robustness of organizational power structures. The alleged utopian nature of emancipatory projects was a reason for claiming their disconnection from reality and the mundane practices of management (Alvesson and Willmott, 1992). Furthermore, critical management studies and critical approaches to IS have been accused for being one-sided (anti-performative and anti-management), negativistic and unconstructive (Alvesson and Willmott, 1992). Moreover, Wilson (1997, p. 196) in response to Hirschheim and Klein (1994) questions their “arbitrary ideological position” and criticises the conditions for emancipatory ISD processes. Being suspicious of the real meaning and agenda of the humanistic approach he contends that an ISD process conceived of as a rational discourse that counteracts distorted communication, can be seen as a formula for totalizing discourses, that view organizations as homogeneous entities.
Empirical studies are called for to examine and respond to these claims and criticisms. In this paper we draw from a longitudinal case study of ISD with the aim to 1) revise the concept of emancipatory ISD as originally proposed by Hirschheim and Klein (1994), and 2) to develop a more refined and practice-based view of micro-emancipation in ISD (informed by Alvesson and Willmott, 1992, 1996) and discuss its relevance in the face of major criticisms in the literature.

3 RESEARCH SITE AND METHODOLOGY

Our motivation to study participatory and emancipatory IS developments initially came from practice. While conducting a longitudinal case study of informatization processes and organisational development in the Colruyt company (from 1993 to 2008) we observed unique practices of user participation in the development of IS with explicit democratic and emancipatory intent. Intrigued by these observations we decided to study more in-depth the meaning and nature of user participation as it emerged and continued throughout the company’s history. We focused on the Information System for Information Dissemination (ISID), which is a groupware and document management system that was developed by the company’s IT department. ISID’s development commenced in the early 1970s and it has been under continuous development ever since. ISID proved to be an essential component of participatory decision making and the cooperative culture that the company has been developing since its inception until today. It was therefore particularly intriguing to investigate how the Colruyt company actually developed ISID, how it attracted and engaged users and how users felt about their participation in the systems development and use. Given the criticisms and controversies found in the literature we were particularly interested in examining participatory and emancipatory ISD processes and theoretically explaining the controversies around non-emancipatory and emancipatory discourses and practices.

The Colruyt company was founded in Brussels, Belgium, in 1965 as a single food discount store - a revolutionary concept in Europe at that time. Selling its products 10% below prices charged by competitors, paying employees 10% above industry average wages, and realizing a 1% profit margin on sales remains the company’s business strategy (IS Manager, Interview 1993, 2000). Today, the Colruyt company is the third largest food retail chain that operates 205 food discount stores in Belgium and 45 stores in France with annual sales revenue of US $8 Billion (company Annual Report, 2007-2008).

Mr. Jo Colruyt’s (CEO until 1993) relied on Information Technology (IT) to support and innovate all business processes. IT was adopted not only to automate tasks and improve efficiency, but also, and more importantly, to support workers in their operations and work simplifications as well as participation in decision-making. Jo Colruyt recognized that IT affects people and influences social conditions of work. He stated: ‘Existing conditions, social structures, relations between individuals, and relations between social classes in a company change when new technologies are introduced’ (interview 1993). To enable workers to embrace IT and meaningfully contribute to company’s continuous innovation it invested heavily in staff training and education (e.g. in 2007/8 training and education budget was Euro 9.5 mil or 3.2% of after tax profit). Seminars were available on communication, self-actualization, self-empowerment, self-expression, decision-making and assertiveness.

Our longitudinal study of the Colruyt company started in 1993 when a co-author first visited company headquarters in Brussels to interview founder and then CEO, the late Jo Colruyt. Since then regular visits to company Headquarters involved informal discussions and formal interviews with employees and managers, and attending and observing official meetings in stores, warehouses and headquarters and observing the use of ISID. On-site audio taped interviews were conducted with Jo Colruyt, the former CEO of the company and his son Jef Colruyt who became CEO in 1994, the chief information officer, the marketing manager, middle level managers, IS personnel, workers in stores and warehouses, and union representatives. In total twenty five interviews were completed over the period 1993 until 2008 and ten company meetings attended. Semi-structured interviews were
conducted and later transcribed for further analysis. Data collection also included researcher’s observation notes, company documents, policy statements, work procedures and rules, meeting documents, most available via ISID, as well as company annual reports (1975, 1985, 1988, and 1990-2008), union reports, and newspaper articles.

Our interpretation of the nature, meaning and challenges of ISID development and use drew on the narratives of actors, employees, managers and IS specialists and the ways these individuals made sense of and reflected on events and on-going changes. The narratives and meaning making processes were an integral part of their social construction of reality, their social actions as well as social and cultural (re)production (Czarniawska, 1998). The narratives reflected what was important to these actors and what was problematic and challenging in ISID development. Furthermore, we analysed the documents created as part of ISID development, relevant for the company as a whole and communicated via ISID. We adopted thematic analysis that was informed by theoretical concerns and interrelated with the analysis of interviews.

For the analysis of empirical material collected during ISID development we adapted Alvesson and Willmott’s (1992, 1996) two-dimensional framework. This framework was applied as a guide and at the same time used for critical reflection that refined it. The framework proposes a distinction between the type of emancipatory project and the focus of its intent in order to enable a more refined analysis and understanding of emancipatory projects in practice. The type of emancipatory project may range from questioning, to incremental transformation or reformist, to utopian type. While questioning involves critiquing, challenging and at times resisting dominant forms of thinking and social arrangements, without proposing the desired (or the ideal), the utopian type advocates alternatives to existing conditions. The utopian element is important, as Alvesson and Willmott (1992, p. 450) explain, when participants envision alternative arrangements, social relations or ends. Between these two opposites, an incremental or reformist type of emancipation involves gradual change towards the desired forms or systems.

Concerning emancipatory intent a distinction is made between means, social relations and ends. The emancipation of means concerns distorted discourses and oppressive organizational practices that are assumed to be necessary to achieve organizational and managerial ends. The emancipation of ends, on the other hand, ‘is concerned with unfreezing institutionalized priorities and, thereby, opening up debate about the practical value of economic growth, consumption, the quality of life and so on’ (Alvesson and Willmott, 1992, p. 450). Finally, the focus on social relations draws attention to social and power structures, relations of domination and control, and the ways these limit autonomy, creativity, self-determination and self-realization.

The Alvesson and Willmott (1992, p.450) framework introduces analytical distinctions that are relevant for examining and clarifying the nature and scope of emancipatory ISD. Inspired by their work we adopted this framework in the analysis of practices during ISID development in the Colruyt company, which we present next.

4 EXPLORING EMANCIPATORY PRACTICES OF ISID DEVELOPMENT

To explore practices during ISID development and examine specific dimensions of the macroemancipatory framework as proposed by Alvesson and Willmott (1992, 1996) we shall follow the project of ISID development from its beginning in 1970 till 2008. We present our findings along these dimensions.

Questioning, challenging, and critiquing
ISID was initiated based on the criticism of existing processes, working conditions, and inefficiencies, made by both managers and employees. Employees were especially dissatisfied with poor access to information. In the 1960s and early 1970s company relied on face-to-face communication and workgroups (written records from group meetings were widely distributed) in an attempt to build a culture of open communication and free expression. However as the company grew it faced increasing problems with communication and information sharing among hundreds of its distributed employees. The first phase of ISID development in the 1970 started in response to criticism of poor information and lack of support for distributed groups of employees and various forms of participative decision making (meetings, workgroups) that could not rely any more on paper-based dissemination of information and reports. Although Colruyt was known as the first company in the 1970s to set up a computer system that store documents and disseminate them in printed form to its employees, employees did not considered it good enough for the cooperative work and decision making processes already established when the company was smaller and working face-to-face. Employees requested faster information channels, they criticised the lack of timely and relevant information they needed to competently and efficiently perform their jobs (in stores, warehouses, sales), coordinate their actions and take part in decision making.

In the second and third phase of ISID development narratives changed with the focus shifting toward more subtle issues such as equity in terms of rights to access to information. From the first ideas ISID provided unrestricted access to information to all employees. However, an incident of information misuse of (documents from ISID leaked by Union Stewart and misinterpreted in a TV program in 1984 with a significant damaged to Colruyt reputation) sparked a wide-ranging debate about the conflicting demands between completely open access to ISID and preserving confidentiality of sensitive information in the interest of all. If ISID was going to realize Colruyt principle of ‘access to information as a right’ for all employees, the question was on what bases this right could or should be curtailed? An anti-emancipatory step was taken – considered necessary but regrettable by the majority of employees – to restrict access to confidential information. Mr. Jo Colruyt was among those who preferred to train the community in responsible ISID use but he also realized that protecting confidential information was inevitable at a time:

> It will always remain a delicate balance between confidential ISID documents and a broad access to information. I prefer to keep the number of confidential documents small in relation to non-confidential documents. It appears useful to teach employees to use information judiciously and to instruct them in normative behaviour with respect to information. Employees have the right to a very broad range of information [stored in ISID]. (Jo Colruyt, 1984)

It was widely accepted by managers and workers that democratic rights and open access to information in ISID needed to be balanced against the risks of misuse and the harmful disclosure of confidential information outside the company. However, the views how to achieve this differed.

Questioning the means, namely, ISID development discourse and practices in addition to the ends to be achieved with ISID implied user-developer and worker-manager social and power relations. Open access by anyone to ISID led to equalization of power that some managers felt threatening and pockets of resistance surfaced more or less overtly. This ‘relational friction’ can be seen to arise from an inherent labour-capital conflict that lies in the foundation of any capitalist company. By supporting and enabling more effective management and control of processes and employees, IS often exacerbate this conflict. However, the Colruyt company’s distinct feature is its continuous effort to reveal and openly confront labour-capital conflict manifested in various relations and processes. By way of questioning, challenging, arguing, critiquing and requesting changes ISID development and implementation led to incremental transformation of discourses and practices (means), power relations and ends to which we turn next.

**Incremental changes and transformation**

Starting during the early 1970s ISID development impacted on and was impacted by changing relationships between users and developers as well as workers and managers. There was an awareness
among the workers that ‘having information means having power to act in an informed manner’. With IT training and with experience in ISID development and use, workers became increasingly aware of their role in the ISID development as part of their participative rights.

Furthermore, incremental changes were experienced in worker-manager relations. Due to access to ISID and workers’ informed participation in decision making, power relations became less hierarchical and more cooperative. However some managers resisted information sharing and devolution of power. Transformation occurred as part of culture change and personal development through training:

We have a culture of [personal] growth within the company. I myself have experienced these [culture and personal growth] at all managerial levels, it is evident, and you know this. I think that only individuals who can accept sharing information with subordinates and with others are promoted to managers. (Walter de Hertog, interview 2003)

Transformations towards sharing of information – primarily through ISID – and devolution of power were not without conflicts, but ultimately, the few managers who obstructed these processes retired.

An important question concerns the emancipatory transformation of ends, that is to say, the extent to which IS development and implementation contributed to the articulation/transformation of ends and their achievement. Transformation of ends was directed toward a more balanced articulation of economic versus social or humanist aims. The former CIO Marcel Lengeler and other employees mention ‘increasing individual performance’, ‘job improvement’, ‘commitment to hard work’, while at the same time emphasising ‘enhancement of job satisfaction’ and ‘enjoyment of work’. Similarly ISID is seen as a major contributor to both better company performance and community building. In fact the economic and social/humanist ends are often seen as conditioning and supporting each other. One employee, for instance, pointed out that ISID enabled him to enjoy ‘freedom to make decisions and to share responsibility within [his] company’. Furthermore, the value of workplace democracy and power decentralization enabled by ISID also is understood as contributing to organizational flexibility as well as its capacity to adjust to new situations as stated by Mr. Jo Colruyt:

Power decentralization has the enormous advantage of organizational flexibility to instantly adjust the organization to new situations. … To communicate [via ISID] means that as a group we are capable of greater achievements. Moreover, in this way we experience greater satisfaction from our work and experience the joy of an increased work engagement. (ISID document, April 1984)

The evidence shows that transformation of ends achieved through ISID involved a continuous struggle to balance personal and organisational achievements, personal performance and enjoyment of work, and to harmonize the Colruyt company’s economic prosperity with community building.

Utopian vision of ISID development

The ideal of open access to information was an important utopian element in the vision of the ISID development. When confronted with existing conditions and risks of leaking confidential information this ideal could not be sustained. Having such an ideal as a guiding vision for ISID however was and still is beneficial as it indicates a desired state of affairs and reminds company members how far or close they are from achieving it. It also counteracts tendencies to make further restrictions.

Furthermore, ISID development can be seen as a ‘rational discourse’ as defined by Habermas’ (1984) Theory of communicative action. Being very pragmatic Colruyt company members do not talk about the ideal of rational discourse but engage in it in practice. This was demonstrated during ISID development team meetings that one of the authors attended where participants showed excellent questioning skills, providing arguments and counter arguments, and conducting an open and well argued debate. The outcomes of ISID meetings were made public via ISID thus increasing company-wide awareness of its results and seeking responses and involvement by others.

The development of ISID can thus be seen as embracing a utopian vision of the company and its communication. Although in the early years (1970s) this vision seemed indeed quite utopian it was
nevertheless powerful and inspirational: it gave company members a sense of direction and purposefulness. Confronted with existing conditions, the vision stimulated thinking and debates thus enhancing emancipatory project. Importantly, employees maintained a critical attitude:

Of course not everything in the company comes up rosy, anyway this is the case everywhere else, how else would it be possible to maintain a critical spirit. (Claude Pardonche, 1984)

Similarly we can see a utopian element in the vision of the ISID development process as a rational discourse. Being very pragmatic Colruyt members do not talk about the ideal of rational discourse, which reminds us of the debate about Habermas’ (1984) ideal speech situation. Since early 1970s they have talked about and applied norms and rules regarding employee participation rights, fairness in social interaction and reasoned argumentation in all key problem solving activities and meetings. ISID development practices and discourses reflected these more general norms of broad participation, rational debate and argumentation.

Another powerful utopian vision in the course of the ISID development has been the vision of equality and symmetrical power relations, among managers and employees as well as among the users and developers. The vision of decentralisation of power that required decentralization of information was and remains to be a guiding principle in ISID development. For instance in the first stage of ISID development employees requested to be informed about company performance and current operational issues in order to be able to participate in decision making. Similarly the vision of ISID as social infrastructure and social arrangements to achieve individual and collective self-determination, power decentralization, and reduction of alienation and domination featured for instance early on in debates about the ways information is captured and distributed to all members, and later in discussions about the open access and necessary protective measures (sparked after Unions’ misuse of ISID documents).

An important utopian view of ISID focused on company ends. Ideally the development and use of ISID is seen as enrichment of professional and personal life of all members, contributing to democratic work environment, work enjoyment, individual and collective self-realization, and thereby freeing creative capacities of employees, opening up company opportunities and improving performance. The vision of ISID was essential to realize the company’s philosophy and to implement in practice its ideals, principles, and strategies.

The utopian vision of ISID and its role in determining and achieving company ends can be criticized, especially from a poststructuralist perspective, as totalizing, possibly excluding other voices. There is the risk, a poststructuralist would claim, that the utopian vision of participatory culture, open communication and the ideal of ISID development as rational discourse degenerates into ideology with anti-emancipatory implications. While such a risk cannot be excluded, our in depth analysis of micro-emancipatory events and situations (grounded in the narratives by actors, documents, ISID transcripts) suggests that the utopian element in ISID development stimulated critical thinking and opened up novel alternatives as envisaged by Alvesson and Willmott’s (1992):

The utopian element emerges when the current conditions are confronted with a new form of ideal, which aims at opening up consciousness for engagement with a broader repertoire of alternatives. Utopianism then represents alternative thinking rather then the suggestion of a ready-made, better alternative or the providing of courses of action. (p. 450)

The utopian vision of open communication inspired and opened up desirable models of ISID that drove its technological development. Rather then imposing a solution, the utopian vision released creative capacities of employees, motivating them to search, and continue searching, for innovative ways of communicating, working and decision making.
5 MICRO-EMANCIPATORY ISD FRAMEWORK

The above analysis reveals a large range of foci and approaches in ISD practices that potentially have emancipatory implications. These emancipatory implications do not necessarily correspond to grand views of liberation. Instead we identify numerous micro-emancipatory processes and practices that are situated, local and meaningful for the people involved. Macro-emancipatory processes are defined within the framework in Table 1, adapted from Alvesson and Willmott (1992).

<table>
<thead>
<tr>
<th>Type of emancipatory project</th>
<th>Foci of emancipatory intent</th>
<th>Questioning, challenging, critiquing, requesting</th>
<th>Incremental changes and transformation</th>
<th>Utopian views and vision</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Focus on means:</strong></td>
<td>Discourses and practices of ISD as a means to achieve individual, group and organizational ends</td>
<td>Has ISD questioned, critiqued and challenged current practices and unspoken assumptions, and identified employee needs?</td>
<td>Have discourses and processes of ISD, implementation and use transformed and in what ways?</td>
<td>Have any utopian views of discourses and practices of ISD been proposed?</td>
</tr>
<tr>
<td><strong>Focus on social relations:</strong></td>
<td>Developers-users, Workers-managers</td>
<td>Has ISD questioned or problematized user-developer and worker-manager social and power relations?</td>
<td>Has ISD impacted on transformation of worker-manager and user-developer social and power relations?</td>
<td>Have utopian views or vision of social relations influenced or inspired ISD?</td>
</tr>
<tr>
<td><strong>Focus on ISD ends:</strong></td>
<td>Individual performance and job satisfaction, Democratic workplace and participative decision-making, Organizational performance and community building</td>
<td>Have participants or users in ISD debated and criticized the purpose of the system?</td>
<td>To what extend has ISD and implementation contributed to articulation and/or transformation of ends and their achievement?</td>
<td>Have utopian views or vision of ends been proposed to be achieved by ISD?</td>
</tr>
</tbody>
</table>

Table 1. The ISD micro-emancipatory framework (adapted from Alvesson and Willmott, 1992)

First the object or focus of emancipatory change may range from the means to the ends:

- **Means** refer to discourses and practices of IS development and use that enable achievement of individual, group and organizational ends
- **Social relations** affected by IS development including those between developers and users, and workers and managers
- **Ends** supported and enabled by IS development include:
  - Individual performance and job satisfaction
  - Democratic workplace and participative decision-making
  - Organizational performance and community building

Each of these foci of emancipatory practices in the observed ISID development in Colruyt is examined in terms of the nature or type of emancipatory change, that is, whether or to what extent it involved:
• **Questioning, challenging and critiquing** current work and information practices as well as **arguing** and **requesting** new and higher quality of information, better access to information, user-developer and worker-manager power relations, articulation of ends, etc.

• **Incremental changes and transformation** of discourses and processes of ISD; worker-manager and user-developer social and power relations; and individual or organisational ends;

• **Utopian views and vision** of discourses and practices of ISD (means), social and power relations and the ends to be achieved with ISD.

These dimensions determine the ISD framework (Table 1) that analytically distinguishes 9 classes of micro-emancipatory changes. Based on our analysis we formulated a generic question to explore micro-emancipatory processes for each class. For instance, the question (second row, first column): *Has ISD questioned or problematized user-developer and worker-manager social and power relations?* is intended to focus attention on social and power relations in an organisation and how ISD questioned or challenged them. These questions may assist researchers to examine specific, local micro-emancipatory practices and investigate the meaning of emancipatory ISD.

### 6 THEORETICAL CONTRIBUTION AND LESSONS LEARNED

The paper contributes to the debate about the emancipator ISD practices. It revisits the concept of emancipatory ISD as originally proposed by Hirschheim and Klein (1994) and develops a more refined and practice-based view of micro-emancipation in ISD informed by Alvesson and Willmott (1992, 1996). Grounded in empirical data from the longitudinal case study of ISID development in the Colruyt company (1993-2008) the paper discusses and provides arguments for the micro-emancipatory ISD framework. The framework identifies classes of micro-emancipatory practices illustrated by examples from ISID development and use. The case analysis shows that micro-emancipatory practices are local, situated and meaningful for the people studied.

The proposed framework to study micro-emancipatory ISD practices grounded in empirical data in many ways responds to criticisms of emancipatory ideas in the literature. Chief among them is the charge that emancipator ISD neglect the reality of business conditions, the pre-eminence of shareholders’ interests and robustness of organizational power structures (Wilson, 1997). The alleged utopian nature of emancipatory projects was a reason for claiming their disconnection from reality and the mundane practices of management. The paper demonstrates the reality of emancipatory ISD practices and their meaning in context, however local or contentious they might appear. The company is both an industry leader in efficiency and in democratic social relations. The emancipatory ISID development in Colruyt reflects its participatory culture, devolved decision making, autonomy and responsibility of workgroups, as well as individual and collective strive for excellence and superior performance. There ISID development processes were equally driven by concerns for work improvement, efficiency, effectiveness, and sound technical solutions, as they were by improvement of working conditions, employees’ work satisfaction, individual and collective self-realization and community well-being.

Furthermore, increasing efficiency and effectiveness of business processes and overall company performance – seen as economic objectives – are not necessarily opposed to workers’ participation, their greater autonomy, responsibility and emancipation – perceived as social and humanist objectives in ISD. Achieving social and humanist objectives does not need to be at the expense of economic ones, as is widely assumed. Similarly, increasing levels of economic performance do not necessarily require increasing control, diminishing autonomy and stringent subordination. In fact in Colruyt by acquiring greater autonomy and responsibility and by increasing their participation in decision-making enabled and supported by ISID, employees became more innovative in simplifying work processes and more successful in increasing their efficiency and effectiveness. Participatory rights, access to information and empowerment of employees have been practiced as values of their own, which in turn became the drivers of the company economic prosperity. While facing tough economic conditions and
ever harsher competition, the Colruyt company has invested in ISID prospered and was continually expanding: it is the third largest food retail chain in Belgium with stores expending in France and Germany. The longitudinal study of ISID at Colruyt demonstrates not only the realism of the emancipatory ISD in practice but also confirms the reality of Hirschheim and Klein’s proposals:

It can be seen that emancipation produces positive consequences for effectiveness and efficiency concerns: stable, self-confident personalities are the pillars of a stress-resistant work force; individuals confidently expressing ideas is the bedrock of creativity to meet competitive demands; and only people accustomed to autonomous, responsible action can be expected to make initiative when things go wrong, which increases organization’s flexibility and capacity to deal with uncertainty (1994, p. 98)

Another major lesson from the Colruyt company perhaps is that in order to advance the emancipatory ISD, and make it more relevant for IS practice, one can use not only a humanist argument but an economic one as well. By adopting numerous micro-emancipatory practices in ISID development Colruyt continuously transformed its processes and practices leading to equalizing developer-user and manager-employee power relations. ISID development was part of the decentralization of decision making, increasing employees’ discretion and autonomy, and enhancement of workplace democracy. Importantly though the company’s emancipatory transformation produced commercial success. In almost every aspect of its performance – sales per store, sales per square meter or per employee; profit per employee or square meter, etc. – Colruyt shows superior results compared to other similar retail chains. This is despite additional costs of coordination and reaching agreement inevitable in decentralized management. Linking emancipatory ISD development practices with company business success and more broadly linking the practices of workplace democracy and employees participation in decision making (enabled and supported by ISID) to the company superior performance and commercial success, may be a story that would catch managers’ and IS practitioners’ attention.

By using an economic argument we may have some chance to draw managers’ and companies’ attention to the unrealized potential of the emancipatory ISD project and the dangers of neglecting it. We have to make it clear that we are not advocating here yet another form of instrumental use of humanist values. We aim to draw attention and raise consciousness about the micro-emancipatory practices as liberating, unleashing individual and collective creative potential for humane and more democratic forms of work and social relations as well as for company economics success. These results open up new space for further examination and critical assessment of micro-emancipatory ISD in practice.

References

and M, Kyng (eds), Computers and Democracy—A Scandinavian Challenge, (pp. 251-278), Avebury, Aldershot.


A PRAGMATIC APPROACH TO IS DEVELOPMENT AND SOCIO-TECHNICAL EVALUATION

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0256.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Information Systems Development (ISD), Job satisfaction, Action research, Socio-technical theory</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
A PRAGMATIC APPROACH TO IS DEVELOPMENT AND SOCIO-TECHNICAL EVALUATION

Sabine Madsen, Roskilde University, DK-4000 Roskilde, Denmark, sabinem@ruc.dk
Richard Vidgen, University of Bath, Bath, BA2 7AY, UK, mnsrtv@management.bath.ac.uk

Abstract

This paper provides an action research account of why and how UML use cases and socio-technical analysis were combined and used to support the development of an information system, using Multiview/WISDM as the framework of ideas. Significant learning points from, and the motivation for, the research is that a focus on task satisfaction and socio-technical evaluation, rather than on the broader concept of job satisfaction and a full, traditional socio-technical analysis, was more acceptable in practice as it related more clearly to the IS domain. The paper presents a socio-technical evaluation (STE) process, reflects on its application to a particular web-based project, and makes recommendations for its use in future ISD projects. The STE process consists of a task satisfaction survey (TSS), which draws on use cases as a template for investigating users’ perceptions of task efficiency, effectiveness, and enjoyment, supported by qualitative interviews in order to systematically identify, verify, and discuss the users’ problems and improvement proposals before and after IT implementation. The use cases constitute the context-specific, development-oriented link between the work system, systems development, and the STE process, which in turn helps ensure that the main goal of supporting and improving the work system through continuous development and deployment of an IT system is kept in mind.

Keywords: Socio-technical analysis, UML use cases, Action research, Task satisfaction
1 INTRODUCTION

The literature on information systems (IS) is very clear that the number of IS project failures are unacceptably high; the main reasons for this being 1) the failure to predict and manage the IS Development (ISD) process and the impacts and outcomes of introducing a new IS into an organization and 2) a lack of explicit concern for human and organizational aspects throughout the IS’s entire developmental and operational life cycle (Doherty & King, 2005). The lack of concern is often explained by a prevalence of a ‘hard’ mindset, and therefore a focus on processes, engineering oriented methods, and IT; that is a focus on technical development rather than socio-technical change (Clegg, 2000; Doherty & King, 2005; Mumford, 1997). Moreover, it is explained by limited use of socio-technical approaches outside of an academic context, among other reasons, because practitioners are unfamiliar with these approaches or because they find them too cumbersome, time-consuming, and ‘soft’ (Clegg, 2000; Doherty & King, 2005; Kautz & Pries-Heje, 2000; Mumford, 1997).

The aim of this paper is to develop, through action research, a pragmatic approach to socio-technical analysis of information systems that draws on and can be used alongside traditional, engineering based approaches to ISD. To this end, we suggest that focus be shifted from a full, traditional socio-technical analysis and its belonging concept of job satisfaction to a time and resource efficient socio-technical evaluation that emphasizes task satisfaction. The research contributes to ISD literature and practice with a socio-technical evaluation (STE) process, reflections on its application to a particular web-based IS project, and recommendations for its use in future ISD projects. The STE process consists of a task satisfaction survey (TSS), which uses UML use cases as its organizing structure, supported by qualitative interviews. The process is used both before and after IT implementation. The use cases provide the concrete, development-oriented connection between the particular work system, the IT system under development, and the content and outcome of the STE process.

The structure of the paper is as follows. In the next section we review the literature on traditional, engineering as well as socio-technical approaches to ISD. In the third section the research approach is described. Section four introduces the case organization and describes the action research intervention. In section five the organizational impact and the lessons learnt from the application of the STE process is discussed and implications for future projects drawn out. The final section provides a summary of the research and its main conclusions.

2 BACKGROUND

The number of formalized, engineering oriented ISD methods is vast. Many of these have been developed through a process of taking successful programming strategies and broadening them out into analysis and design. This was seen first with structured programming, which was expanded to structured analysis and design, and now with OO programming and OO analysis and design, where a convergence of ‘best of breed’ OO methods has led to development of the unified modelling language (UML) notation. The language has become widely accepted as a modelling standard for OO software development (Dobing & Parsons, 2006). It provides a number of techniques for OO analysis and design, i.e. for requirements specification and verification via use case, class, activity, collaboration, sequence, and statechart diagrams as well as use case narratives. The literature often prescribes that UML should be combined with use case-driven iterative and incremental programming.

Despite the predominance of traditional methods concerns have been voiced about their effectiveness and methods, such as ETHICS and Multiview, which combine a focus on the social and the technical have been proposed. ETHICS (Mumford & Weir, 1979; Mumford, 1983, 1995) - Effective Technical and Human Implementation of Computer-based Systems - aims to improve an organizational work system by increasing both job satisfaction and organizational efficiency and effectiveness. ETHICS provides a systematic, analytic approach for work design that prescribes thorough and well-informed
definition of both social and technical objectives and alternatives as well as evaluation of alternatives in accordance with goal fulfilment, costs, constraints, and resource consumption and emphasizes continuous user involvement throughout the project. It is assumed that high job satisfaction exists when there is a good ‘fit’ between the employees’ expectations and needs and the requirements of the job. The ‘fit’ can be measured via a questionnaire survey, such as Mumford’s (1983) 58 item questionnaire that addresses the overall job satisfaction as well as five areas: knowledge, psychological needs, efficiency, task structure, and ethics. For each of these areas, the ‘fit’ can be assessed and used as input for formulating design objectives.

Multiview (Avison & Wood-Harper, 1990) is a contingency method that incorporates a focus on both hard and soft aspects of ISD by drawing on and combining known hard and soft methods such as Structured Analysis (De Marco, 1979), Soft Systems Methodology (SSM) (Checkland, 1981; Checkland & Scholes, 1990), and ETHICS (Mumford & Weir, 1979; Mumford, 1983, 1995). Multiview has been continuously developed through action research to accommodate critique and lessons learnt as well as to address emerging application areas (e.g., web development) and new ISD methods (e.g., OO analysis and design via the Unified Modelling Language (UML)) (see, e.g. Avison et al., 1998 for an account of Multiview2 and Vidgen et al., 2002 for a description of Multiview/WISDM). However, despite these efforts an empirical study among former Multiview students turned IT professionals shows that there were only a few examples of direct adoption of the method, because it was considered too soft and time-consuming (Kautz & Pries-Heje, 2000).

A recent development that draws on socio-technical ideas is Alter’s (2006) work system method, which is a well described, rigorous, but non-technical approach for communicating about how current work systems operate and how they can be improved, with or without the use of IT. The work system method has much promise, but may come to suffer from the same barriers to method adoption and diffusion as ETHICS and Multiview, namely that it is developed in an academic context, that it advocates a systems concept that goes well beyond IT, and that it is perceived as a comprehensive method despite the author’s guidelines on how to use it in a flexible way.

In general, hard methods, such as Structured Analysis (SA) and Object-Oriented (OO) systems development, have had a major effect on teaching and practice, while soft approaches such as SSM and ETHICS have had much influence on the advancement of theoretical ideas and the ISD research agenda, but a much more modest impact on actual practice (Fitzgerald et al., 2002). It is very important to continue the development of, and teaching in, broad and wide-ranging hard and soft approaches, but there also seems to be a need for socio-technical methods and techniques that accept and build easily on the existing hard mind set, thereby allowing a more comfortable route to adoption of (some) concern for the social side, e.g. by supplementing the IS development and deployment processes with a socio-technical evaluation cycle.

3 RESEARCH APPROACH

The research presented in this paper came about as a part of a longitudinal action research (AR) project. The AR project was performed in collaboration with Zenith International, a company specializing in consultancy to the food and drinks industry. It concerned the development of a web-based IS, coined the Research Data Repository (RDR) for the benefit of Zenith’s market research department. The project was performed collaboratively by Zenith and the University of Bath within the Teaching Company Scheme (TCS, a government funded programme that promotes collaboration between industry and universities) and involved the active participation of academic researchers. Officially the project lasted from October 2001 to October 2003, and the IS was in operation when the project finished, but researcher involvement and intervention continued, albeit in a less formal and extensive manner, until April 2005.

A number of different types of AR exist (Baskerville & Wood-Harper, 1998). However, the defining features of all AR is intervention into and change in a practical problem situation for the purpose of 1)
solving the particular problem (problem solving cycle) and 2) contributing to the research literature with new knowledge (research cycle). The research presented in this paper aimed to fulfill this dual imperative (McKay & Marshall, 2001) by 1) improving Zenith’s internal report production process via development of a web-based IS and 2) creating new ways to understand, support, and improve ISD processes and method usage based on practical experience (Mathiassen, 2002).

Multiview/WISDM (MV) was chosen as the framework of ideas (Checkland, 1991) for guiding the research and problem solving cycles because it offers a structure for web-based ISD that facilitates the construction of a situation-specific method. Thus, MV allows for a research focus on process emergence and method usage as well as practical deployment of ISD methods for the purpose of problem solving. The MV framework does not introduce new methods and techniques. Instead it relies on pre-existing methods that are categorized in a method matrix.

The method matrix comprises a collection of formalized methods and techniques organized according to five different aspects of ISD. It aims to support a socio-technical approach by including methods that emphasize design and construction of technical artefacts as well as methods that address the human and organizational aspects of ISD (Vidgen, 2002). The archetypical methods in the matrix are SSM (Checkland, 1981; Checkland and Scholes, 1990) for organizational analysis, UML (Booch et al., 1999) for information analysis and technical design, ETHICS (Mumford, 1983, 1995) for work design, and web usability (Nielsen, 2000) for the human-computer interface. MV is not supported by a pre-specified process model. Instead the method authors suggest that the focus of attention changes during the process as the IS developers move in and out of the five aspects of the method matrix and that while one aspect might be the focus of attention at a particular time, the other aspects can – and should - still be present in the minds of the developers (Vidgen et al., 2002).

The research presented in this paper is qualitative in nature as we are concerned with questions of how and why, not how many or to what extent. Moreover, the number of people employed in the market research department (5-6 people) does not allow for quantitative analyses. Even though a questionnaire survey was used to collect pre- and post-implementation data about task satisfaction it was considered a qualitative technique relevant for getting a quick indication of employee perceptions and the results were analysed as qualitative data. However, in another setting, given a sufficient sample size, the survey approach could be used for conducting quantitative research and the questionnaire and its results could be submitted readily to statistical tests for reliability and validity.

<table>
<thead>
<tr>
<th>Research cycle, M_R</th>
<th>Problem solving cycle, M_PS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interviews:</td>
<td>Information analysis:</td>
</tr>
<tr>
<td>7 interviews with 5 project participants and 2 future end users (Nov 2002).</td>
<td>4 Interviews with Market researcher director and 3 Market researchers, as input for requirements analysis, (May 2002).</td>
</tr>
<tr>
<td>4 interviews with Market researcher director and 3 future end users (July 2004).</td>
<td>5 TSS follow-up interviews with Market researcher director and 4 Market researchers (April 2005).</td>
</tr>
<tr>
<td>Documents:</td>
<td>Work design:</td>
</tr>
<tr>
<td>Original project proposal, minutes from steering committee and technical meetings, company documents as well as project reports and deliverables, e.g. the requirements specification.</td>
<td>Pre-implementation: task satisfaction survey, 6 participants, (July 2002).</td>
</tr>
</tbody>
</table>

Table 1: Data collection in RDR project

The empirical material is made up of interview transcripts, collected documents, and questionnaire survey data (see Table 1). To write the case description we have read and analysed the empirical data using a narrative-oriented approach, where we continuously asked “and then what happened?” (Pentland, 1999). Moreover, to gain a deeper, more nuanced understanding of the TSS findings we discuss the results in relation to the requirements specification from May-June 2002 and interview data from April 2005. The results of the analyses are presented below.
4 THE INTERVENTION

As described in Vidgen and Madsen (2003), the RDR project was undertaken as in-house development. The steering committee involved 6-8 people and the project team consisted of 3-4 people, with one full time developer and the others being involved in the actual development activities to varying degrees.

The setting for the RDR project was Zenith’s market research department, which consists of six full time employees, including the Market researcher director. Each year the department produces a number of market reports, with the two most important ones being the ‘Bottled Water’ and ‘Water Coolers’ reports. The reports are based on data gathered from as many companies as possible in a line of business, e.g., bottled water. The reports are then sold to companies in the drinks industry, such as manufacturers (who provided the original detail data), packagers, and distributors. From initiation to publication, each report takes around three to four months to produce. Each report is led by a single market researcher who gains a deep insight into the data and manages the structure of the report. A large volume of market data has to be collected, stored, processed and formatted. The Market researcher director explains the report production process in the following way: “For the 10-11 years the company has been going, we have produced company profiles in Excel format, we have linked each profile to overview tables and have created charts manually in PowerPoint…Charts then pasted from PowerPoint, pictures into Word and put it all together using these packages. Obviously, this is extremely time-consuming…” (interview, November 2002)

The requirements analysis established that “the main requirement for the new system can be encapsulated as ‘improving the report production process’.“ (requirements specification, June 2002). The four stages of report production are (Figure 1):

- Gathering and checking company profile data, approx. 50% of total report production time;
- Preparation of market summaries from company profiles, approx. 25% of total report production time;
- Analysis of markets and trends, approx. 20% of total report production time;
- Formatting the report for publication, approx. 5% of total report production time.

The requirements analysis showed that the key objectives of the RDR application were to: input and store company profiles in the database (stage 1), combine data from company profiles (stage 2) to form meaningful market analysis (stage 3), and automatically generate a formatted report from the database (stage 4).

From the outset the RDR project was conceived as an initiative that would have a major impact on the work practices in the market research department. An early idea was therefore to collect pre- and post-implementation data based on the generic ETHICS job satisfaction questionnaire (Mumford, 1983, 1995) in order to get an indication of the market researchers’ job satisfaction before and after the RDR application was put to use. Pre-implementation data would be used by the project team to formulate design objectives and guide the development. Post-implementation data would be used by the project team and other stakeholders, such as company management and the TCS representative, to get an understanding of changes in job satisfaction over time and therefore, whether the employees found the project to be a success or not.

Approval to administer the ETHICS job satisfaction questionnaire was sought from the market research director, who in turn asked the personnel director to review the questionnaire. Both directors found it difficult to see the association between the ETHICS questionnaire, the ISD process, and the RDR project. The personnel director was also concerned that the questionnaire was intruding into human resource territory and, furthermore, considered that the questions were somewhat naive, leading, and that more than half were phrased from a negative viewpoint. For example, ‘Senior management is out of touch with the way people feel’. The directors felt that the ETHICS questionnaire had the potential to increase
(and even create) tension between staff and management, undermine the organizational culture, and generate disquiet. At first these concerns came as a surprise for the project team. However, on reflection, the ETHICS job satisfaction questionnaire is a general purpose instrument that does not contain questions directly concerned with IS or ISD, and, clearly, its deployment was not perceived as a meaningful intervention in this situation. The generic ETHICS instrument was therefore abandoned. Instead the project team decided that a way of assessing job satisfaction that was grounded in the specific context of the RDR project was needed.

As a part of the requirements analysis, UML use cases had been developed to understand and document the current work system in a structured and comprehensible manner (Figure 1). After some consideration, it was decided to let the use case diagram form the template for the production of an RDR-specific questionnaire and to create a set of questions that could record user perceptions for each use case in the diagram. With MV as our framework of ideas we looked to the other methods in the method matrix for inspiration for formulating relevant questions. SSM’s (Checkland & Scholes, 1990) three E’s - efficacy, efficiency, and effectiveness - are used to judge the success of a systemic transformation. Efficacy is concerned with whether the system works: does the use case transformation of input to output take place? Efficiency considers the amount of resource needed to execute the use case; in this context market researcher time is the most precious resource. Effectiveness asks whether the use case contributes to the accomplishment of the larger business process, i.e., the production of a market research report. In addition to the three E’s we also ask about task enjoyment, which is related to ethicality and is one of two further “E” evaluations proposed by Checkland (the other being “elegance”). Thus, it was decided to maintain a focus on the market researchers’ satisfaction, not with their jobs in general, but with the particular tasks as well as with the report production process as a whole. Use case activity 5 (Figure 1) was excluded from the questionnaire as this activity is executed solely by the market research director. A set of five questions was therefore created for each of the use cases 1 to 4 in Figure 1:

• The use case activity is easy to do (individual business process efficacy);
• The use case activity takes up too much of my time (individual business process efficiency);
• The use case activity takes up too much of the total process time (business process efficiency);
• Spending more time on the use case activity would add value to the overall process (business process effectiveness);
• I enjoy the use case activity (task enjoyment).

Finally, an area for respondents to make open comments was included in the questionnaire.
4.1 Pre-Implementation

In July 2002, shortly after the requirements analysis, the pre-implementation TSS survey was conducted. The results of the survey and the implications for the future development of the RDR can be summarized as follows:

Use case 1: there was strong agreement that the task of gathering and checking company profiles is difficult, time consuming, and not very enjoyable. The RDR must improve the efficiency of this task;

Use case 2: feelings about producing market summaries was mixed, being given low and high scores;

Use case 3: users enjoyed analyzing markets and trends and the RDR must free up more time for this activity;

Use case 4: formatting the report for publication is time-consuming, unenjoyable, and error-prone. The RDR must automate this task as far as possible.

The overall business process: there was strong agreement that the process is thorough but that there are too many non-value adding tasks, i.e., company profiles and (use case 1) and report formatting (use case 2).
The results from the TSS were in line with the findings from the requirements analysis, which had already helped the project team understand the report production process and its deficiencies. Thus, the survey confirmed the project team’s understanding and helped establish confidence in the ‘correctness’ of the requirements specification within the project team and among the other members of the steering committee. Enjoyment and Efficiency objectives were explicated, discussed, and on the minds of the project team during development, but the requirements specification was not updated to include the TSS results.

4.2 Post-Implementation

In January through March 2004 the RDR was used for the first time to produce the report for the “UK Bottled Water” market and then again in March through April 2004 for the “UK Water Coolers” market. The UK Bottled Water and UK Water Coolers reports were produced for a second time in 2005. By April 2005 work had begun on producing further reports for among others West Europe Water Coolers and East Europe Water Coolers. In April 2005, the questionnaire was re-administered and used as a vehicle for discussing RDR use, challenges, and potential changes with the individual market researchers than had filled out the survey. The results from the survey and the interview sessions are summarized below.

**Use case 1**: The RDR has made the task of gathering and checking company profiles easier and less time consuming, but not much more enjoyable.

The interviews reveal a much more complex picture. In the first year (2004) much time was spent inputting, checking, and re-checking the quality of the historic data taken from the existing Excel files. Moreover, data inputting is, and has become even more laborious with RDR as it takes place via a cumbersome form-oriented interface with many drop-down menus and entry fields and as it is hard to get an on-screen overview. Suggestions for improvement are to allow for inputting and exporting of data into RDR via Excel and to develop an overview of what has already been inputted and what is missing for each company profile as well as a note facility for storing general and company specific facts and comments.

**Use case 2**: the RDR has improved the task of preparing market summaries from company profiles.

The market researchers have had to learn how to run the queries or they have to ask the Developer to extract the market summaries they need. Both ways, extracting the market summaries is much easier and they are better than before, because they are based on better and more data. However, developer intervention is not unproblematic, and the market researchers raised concerns about loss of ownership and the relationship between ownership and level of incentive (or satisfaction): “I feel very much like the whole process has been passed over to [IT support]… in the past, you did it from beginning to end... it was completely your baby. I feel like I have lost ownership on that, which in a way takes away some of the incentive …” (Market researcher, interview, April 2005)

**Use case 3**: “Analysis of markets and trends”, the questionnaire data is ambiguous and it cannot be concluded that the RDR application has freed up time for the task of analyzing markets and trends.

During the interviews, the market researchers explained that the RDR provides more - and more accurate - data and makes it easier to extract data for the purpose of analysis, but the IT system has not otherwise changed the way the analysis of markets and trends is carried out. Interpretation and writing are still the main activities and is done by the market researcher. However, both the Market research director and one of the Market researchers see much potential for the analysis of markets and trends over time as the database gets stronger. They explain that when there is about ten years worth of data for most companies in the database it will be possible to make really strong historic as well as trend analyses.

**Use case 4**: the RDR has made the task of formatting the report for publication easier, less time consuming, and more enjoyable.
Report formatting has become much easier as it is done more or less automatically, with some intervention from the Developer. The Market researchers explain that “well…you just press a button…[but] in my experience it still takes some time for the [Developer] to do some tweaking.” (Market researcher, interview, April 2005) and that “formatting a report for publication is easy. I mean, again, on that one certainly the profiles are ten times easier, a hundred times easier than what we used to do…” (Market researcher, interview, April 2005).

The overall business process: the RDR has automated the tasks of gathering and checking company profiles and formatting the report for publication, while maintaining a thorough process.

Again, the interview data reveals a much more complex picture. There are indeed time and efficiency savings being made as a result of the RDR. These are mainly related to the semi-automation of the preparation of market summaries from company profiles (use case 2) and the formatting of the report (use case 4). However, further development is needed to get efficiencies in data entry (use case 1). One of the Market researchers states that “…I think it is more a case at the moment that [the RDR] has changed the burden to different parts really… Inputting is the laborious part of it, where previously it was less so. That is disappointing, but on the flip side, extracting the data once it is in there…when you ran the queries, you had full confidence in that data.” (interview, April 2005).

The increase in thoroughness and competitive advantage are due in part to the improvement in data quality. The market researchers trust the database contents and the reports that are produced from the database. Broader data coverage of the markets is also reported.

The report production has been improved but there is more to do. The interviews reveal that improvement of data entry is the key objective for developing the RDR further, and that even though “there is a general sense of buy-in from the staff…you are never going to get 100% euphoria over everything anyway.” (Market research director, interview, April 2005). One of the Market researchers also state that “in general am I satisfied with our process of producing research reports? Not yet. I think it is difficult to get to that stage, where you are 100% happy with it.” (interview, April 2005).

5 DISCUSSION

The organizational impact of the RDR can be summarized as follows. The questionnaire results show that the ‘fit’ for the tasks of gathering and checking company profiles (stage 1) and formatting the report for publication (stage 4) had improved greatly. The interviews reveal that data inputting (stage 1) had become more cumbersome and onerous, while improvements were due to the building of a strong database (stage 1), easier data extraction to form market summaries (stage 2) and the final, formatted report (stage 4). Thus, the two data sources are somewhat contradictory and even though some tasks and sub-tasks have improved a lot, the burden seems to have shifted to other parts of the work process. This suggests that questionnaire data, whether interpreted as quantitative or qualitative data, is best used in conjunction with other in-depth, qualitative data collection techniques, such as interviews, to identify, verify, and discuss the users’ problems and ideas for solutions before and after IT implementation. Moreover, it means that it is difficult to assess whether the RDR is a success or not. We suggest that in this case - and perhaps more generally? - it is more accurate to talk about the IS’s successes and failures by focusing on how well it supports the different phases and tasks within the particular work system, here Zenith’s report production process. IS development is also better pictured as a sequence of interventions, where each intervention has both intended and unintended consequences as well as predictable and unpredictable outcomes. This goes for systems use too where, even without the intervention of IS developers, the system in use will drift over time as users find new ways of working and work-arounds.
In Figure 2 we illustrate systems use, development, and evaluation as a set of interlocking activities. First, requirements are extracted from the current work system (Use), transformed into an IS requirements specification in the form of UML use cases, and used as the basis for the IS design activity. At the same time, evaluation data collected via a survey and interviews, organized around the use case structure of the requirements specification, is elicited from users of the current system and used to evaluate user satisfaction with the current work system (STE). Then, an intervention (ISD) is made on the basis of improvements identified by the STE in conjunction with the specified as well as emerging requirements from the work system. The process then continues in cyclical fashion moving between evaluation and ISD ad infinitum. The work of ISD is never finished as each intervention “solves” one set of problems while introducing new ones, as we found with the Zenith case. Many organizations will take a pragmatic view about what constitutes a “good enough” solution and cease development. However, environment changes and drift happens and even if no additional ISD is planned the STE process could/should be conducted regularly to verify that the current work system and IT are indeed, or still, good enough.

The lessons learnt from the application of the STE process is that it was a powerful tool before implementation, for verifying and building confidence in the requirements specification within the project team and in relation to the steering committee and after implementation, for identifying areas for improvement and for a getting a nuanced understanding of systems success as each task is investigated, not just the broad picture. A particular strength of the approach is the integration with UML use cases which provide the concrete connection between socio-technical evaluation, IS development, and systems use. However, the TSS should be extended to include questions about the need for variety, challenge, discretion, and autonomy (Mumford, 1995) at least at the overall process level to get an indication of which social characteristics the market researchers value.

We conclude that in future projects it is necessary to increase the STE process’s concern for the human and organizational aspects further although continuing to recognize the need for parsimony and pragmatics if the approach is to be adopted by practitioners. We propose that this can be achieved with relatively few resources and by building on the prevailing technical way of working in practice, e.g., by:

- **Pre-implementation:** conducting a more comprehensive TSS that covers both technical and social task characteristics and combining it with other qualitative data collection techniques to identify, verify, and discuss technical and social system requirements. Both efficiency and enjoyment requirements could/should be documented in the requirements specification.

- **Post-implementation:** re-administering the TSS and conducting follow-up interviews at least once and preferably on a periodic basis, e.g. once a year, to assess systems success and drift.
and to continue to ‘grow’ the IT system (Truex et al., 1999) and the support it provides for the work system (Alter, 2006).

6 CONCLUSION

The ISD literature recognizes that many IS developers and stakeholders prefer technical rational ISD methods and either do not know the socio-technical approach or are likely to find it too ‘soft’ and time-consuming. A significant learning point, and motivation for, this research is also that in the case of the RDR project it was more meaningful to consider the constrained but more grounded, IS oriented term of task satisfaction rather than the broader concept of job satisfaction advocated by ‘traditional’ socio-technical analysis.

We therefore developed and deployed our own socio-technical approach consisting of a task satisfaction survey (TSS) and interviews to identify, verify, and discuss the users’ problems and ideas for improvement before and after IT implementation. We considered the approach a separate evaluation cycle that supplemented the development activities, and coined it the Socio-Technical Evaluation (STE) process. The STE process builds on the prevailing ‘hard’ mindset towards ISD by taking a task perspective and by using UML use cases as a template for structuring the TSS around a focus on the users’ perceptions of efficiency, effectiveness, and enjoyment of each use case as well as of the overall process and general level of satisfaction. Pre-implementation, the evaluation was instrumental in establishing confidence in the requirements specification within the project team and among the other stakeholders. Post-implementation it facilitated a nuanced understanding of project success, process shifts, and within-task changes as well as identification of areas in need of further IS development. However, too much emphasis was placed on technical characteristics related to the efficiency and effectiveness of tasks and processes, while social characteristics concerned with the variety, challenge, discretion, and autonomy of these activities to some extent were overlooked, both in the application of the STE process and in the design of the IT system.

We recommend that in future ISD projects an even greater concern for the human and organizational aspects of the work system is incorporated by: conducting a more comprehensive TSS that covers both technical and social task characteristics, including both efficiency and enjoyment objectives in the requirements specification, and re-administering the TSS and conducting follow-up interviews on a periodic basis to assess the success, drift, and improvement opportunities of the IT system as these change over time.

References


TOWARD THE CONCEPT OF POCKETS OF CREATIVITY IN BUSINESS PROCESSES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0339.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Creativity management, Business Process Management, Media industry, Workflow</td>
</tr>
</tbody>
</table>
TOWARD THE CONCEPT OF POCKETS OF CREATIVITY IN BUSINESS PROCESSES

Seidel, Stefan, Institute of Information Systems, University of Liechtenstein, Fürst-Franz-Josef-Strasse 21, 9490 Vaduz, Principality of Liechtenstein, stefan.seidel@hochschule.li

Müller-Wienbergen, Felix, European Research Center for Information Systems (ERCIS), University of Münster, Leonardo-Campus 3, 48149 Münster, Germany, felix.mueller-wienbergen@ercis.uni-muenster.de

Karow, Milan, European Research Center for Information Systems (ERCIS), University of Münster, Leonardo-Campus 3, 48149 Münster, Germany, milan.karow@ercis.uni-muenster.de

Rosemann, Michael, Queensland University of Technology (QUT), Brisbane QLD 4000, Australia, m.rosemann@qut.edu.au

Abstract

Creativity-intensive processes such as the development of marketing campaigns or the production of visual effects increasingly find their way into the agenda of process managers. Such processes often comprise of both well-structured, transactional parts and creative parts that often cannot be specified in terms of their process flow, required resources, and outcome. Moreover, the processes’ high variability sets boundaries for the possible degree of automation. In this paper we introduce the concept of pockets of creativity as an analytic device which is hoped to support process managers in their efforts to identify and describe creative sections in business processes. We argue that this step of identifying and describing is imperative to successfully allocate resources, integrate creativity into the overall process, and introduce process automation for those parts that are well-structured and can actually be automated. Our argument rests in the examination of existent literature as well as in findings from exploratory case studies that were conducted in the film and visual effects industry in order to study processes that rely on creativity.

Keywords: Creativity Management, Business Process Management, Media Industry, Workflow
1 INTRODUCTION

Creative people and their practices play a prominent role in business processes as organizations seek to deploy the merits of business process management to more than just the set of transactional processes (Hall & Johnson 2009). Product development and marketing campaigns are just two examples of such creativity-intensive processes (Seidel & Rosemann & Becker 2008) that increasingly find their way into the agenda of process managers. Besides this, there are entire and quickly growing industries designed around creativity-intensive processes with the entertainment industry being the most prominent example of a creative industry.

To make the merits of business process management available, process modeling has emerged as an important enabler. Existing process modeling languages normally focus on modeling the process flow (e.g. Engels & Förster & Heckel & Thöne 2005). However, business processes that involve creativity are characterized by a high demand for flexibility. Whereas parts of processes in creative environments may be well-structured and easy to model, other parts may be not. For example, a visual effects production process includes tasks such as *receiving materials* where references or scans are received from different sources. Such tasks are well-structured and may even be automated. However, the same process also includes tasks such as *modeling or animation* which require the creativity of the involved artists, and generate creative products. Creative products are characterized by novelty and appropriateness and creativity can be described as the process that leads to the generation of such products (Woodman & Sawyer & Griffin 1993). Particularly two facets of creativity within business processes demand special attention: (a) Creative tasks/sub-processes usually do not have a predefined process structure and (b) creative tasks have outcomes that are – at the minimum to a certain extent – hard to predict (Seidel 2009, Seidel et al. 2008).

It may seem appropriate to view these creative sections as ‘black boxes’ as the underlying processes are complex and hard to predict. However, we argue that this approach would not be sufficient and leaves too much to the individual conducting the creative task. Based on the awareness that the more creative tasks have a significant impact on business processes we believe that it is necessary to create more transparency. We do not suggest to model and prescribe the process flow as in many cases this might prevent people from being creative. However, we propose to identify how creative tasks are characterized, how creative tasks can be supported and how they can best be integrated into the overall business process they are part of.

In this paper, we introduce the concept of *pockets of creativity* as a means to identify and describe parts of business processes that are characterized by creativity. The challenge is to identify what is actually known about a creative section within the business process. We propose a set of constraints that enable process managers to determine whether a process section is indeed creative and to then describe this section.

With this paper we make two primary contributions. First, we shed light on the characteristics of creativity from a business process management perspective and provide researchers with an analytic device that can guide further studies. The device is the result of a design process (Hevner & March & Park & Ram 2004) which is grounded in empirical data. Second, we introduce a means that process managers can use in order to identify and describe pockets of creativity in business processes in order to effectively manage creativity-intensive processes.

We proceed as follows. First, based on the awareness that creativity in business processes becomes manifest in high process variability (Seidel 2009, Seidel et al. 2008) we discuss recent work on variability in business processes and develop a framework that gives structure to our subsequent reasoning (section 2). We then briefly discuss an exploratory study of processes in the film and visual effects production (sections 3.1 and 3.2). Based on three key observations we introduce the concept of pockets of creativity as a means to denote creativity in business processes and relate the concept back to the framework of variability (section 3.3). We then describe an example from the film industry so as
to illustrate how such pockets of creativity can be identified in business processes (section 3.4). In a chapter on related work we link our work back to existent research (section 4). The paper concludes with a discussion of contributions and limitations and provides an outlook to our future research agenda (section 5).

2 FRAMEWORK OF VARIABILITY IN BUSINESS PROCESSES

Recent years have seen a number of studies concerned with variability within business processes. Much of this work has been carried out by scholars who focus on automation and implementation of business processes. In the following we explore some classification schema for process variability that originate from this discussion before we then introduce a framework, which rests in the argument of our colleagues.

Sadiq et al. (2005) identify three dimensions of change in workflow management. The first dimension is labeled dynamism and addresses a workflow’s ability to change when the related business process evolves. The second dimension is that of adaptability, which refers to change that does not affect the underlying business process but instances of such processes. In doing so, this dimension subsumes a workflow’s ability to react to exceptional circumstances, which may or may not be foreseen. The third dimension is that of flexibility which describes a workflow’s ability to execute based on only partially specified models. Thus, the complete specification of the model only becomes available at runtime.

Similarly, Weber et al. (2008) distinguish three different types of process flexibility: built-in flexibility, schema evolution, and ad-hoc changes. While built-in flexibility and schema evolution resemble Sadiq’s dynamism and flexibility dimensions of change, ad-hoc changes always occur on an instance level and, thus, have a meaning different from Sadiq’s adaptability dimension which addresses both foreseen variability at a schema level and unforeseen exceptions that have to be dealt with at an instance level.

Van der Aalst et al. (2005) classify process support systems according to the level of structure of the underlying processes. They distinguish support for unstructured processes, for ad-hoc processes, and for implicitly and explicitly structured processes. According to this framework, support for unstructured processes does not impose any control flow, whereas ad-hoc workflow management systems allow for the modification of workflow specifications during execution of workflow instances. Yet, an explicit process model is required for every workflow instance. Systems that support processes with implicit structure rely on process specifications that do not explicitly define every possible route within a process. In contrast to explicitly structured process specifications, such systems rely on the definition of restrictions and authorizations that merely set out boundaries for altering a process instance.

In the area of creativity-intensive processes (Seidel 2009, Seidel et al. 2008) we are faced with another level of variability which transcends the understanding of an exception as a "deviation from the ideal" (Klein & Dellarocas 2000) and also the notion of flexibility, that denotes last-minute lashing of control flow structures prior to instance execution (Sadiq et al. 2005, Weber et al. 2008). In creativity-intensive processes variability is deliberately injected as these processes rely on divergent thinking and exploration of various options (Runco 2007). Consequently, we advance that creativity is another dimension of change and extends the levels of process flexibility. In creativity-intensive processes, variability remains until a process terminates. The process is not straight-jacketed before execution – neither by explicitly defining every option at build time via a detailed workflow model, nor by deferring decisions as far as possible by providing a loosely specified, flexible workflow model. In a creativity-intensive process it is impossible to specify every decision in advance. Yet, it is not the case that creativity-intensive processes are not eligible for management. As various industry examples exemplify, such processes not only comprise of completely unstructured sections, but also contain sections that indeed may be specified explicitly. In consequence, also the creative parts of a process must heed restrictions and contextual conditions so as to fit into the overall process.
In the remainder of this paper we will apply a two-dimensional framework for addressing creativity in business processes that rests in the above discussion (cf. Figure 1). In order to expose how the occurrence of creativity in business processes relates to the introduced levels of variability, the first dimension distinguishes different points in time in the life cycle of a business process when variability may be eliminated. These are build-time, pre-run-time, and run-time. Although Sadiq et al. (2005) and van der Aalst et al (2005) refer to decision-making at run-time, we advocate to understand these eliminations of variability as to happen prior runtime: Processes or process fragments of these categories have to be explicitly specified before they can be executed. Therefore, merely unstructured and implicitly structured processes as specified by van der Aalst et al. (2005) and process fragments that bear creativity imply de-facto variability at run-time.

The second dimension depicts three key aspects of a business process. Besides the control flow aspect discussed above, there are other facets of a business process to be considered from the vantage point of a holistic perspective (v.d.Aalst et al. 2005, Weber et al. 2008). First, there are resources which actively conduct a business process or are applied or consumed within a process (Russel & v.d.Aalst & ter Hofstede 2006, Weber et al. 2008). Second, there is the process related object or product, which is altered in the course of a process (v.d.Aalst et al. 2005).

Figure 1: Variability in business processes
3 POCKETS OF CREATIVITY

3.1 A study of organizational creative processes

In an exploratory study on creativity in business processes Seidel (2009) and Seidel et al. (2008) investigated processes from the creative industries. The study aimed at getting an in-depth understanding of business processes that highly rely on creativity. Data was collected from three organizations from the field of film and visual effects production. These organizations are very much characterized by the creativity within their processes. Thus, organizations were selected, where “the process of interest is ‘transparently observable’” (Eisenhardt 1989, Pettigrew 1990). The study focused on so-called post-production (Clark & Sphor 1998) and visual effects (VFX) production processes. These processes usually start in parallel to the actual production of a film or TV commercial (the process of filming) and are carried out until the product is finally delivered. The processes that were investigated not only rely on creativity, but are also repetitive. For example, many VFX sequences have to be produced for a feature film. The sourcing strategy involved semi-structured interviews and the use of process modeling techniques. Table 1 provides an overview of case study organizations, interviewed people, and analyzed processes.

<table>
<thead>
<tr>
<th>Organization</th>
<th>Area</th>
<th>Interview Partners</th>
<th>Analyzed Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization 1</td>
<td>Teaching</td>
<td>Management, directors, producers, post production supervisor</td>
<td>Post-production processes</td>
</tr>
<tr>
<td>Organization 2</td>
<td>Post Production: Visual Effects Production</td>
<td>Management, visual effects supervisors, technical directors, artists (compositors, lighters, etc.)</td>
<td>Visual effects production processes</td>
</tr>
<tr>
<td>Organization 3</td>
<td>Post Production, TV Commercials</td>
<td>Management, technical directors, visual effects supervisors, artists</td>
<td>Post-production processes, visual effects production processes</td>
</tr>
</tbody>
</table>

Table 1: Case studies and analyzed processes

In the following section we first discuss some major findings with regard to the nature of the processes that were studied (section 3.2) and then relate these findings to the above introduced framework (section 3.3) before an example case is provided where the application of the concept of pockets of creativity is further illustrated (section 3.4).

3.2 Study findings: process creativity versus product creativity

With regard to the main characteristics of creative sections within the analyzed business processes, three key observations could be made (Seidel 2009, Seidel et al. 2008):

1. In most cases the outcome of a creativity-intensive process is not entirely known in advance. Yet, often it is not the case that nothing is known about the creative product. Just think about a particular animation sequence for a film or a visual effect in the next Hollywood blockbuster. As a producer who needs that sequence to complete the film, you would certainly know the technical format and you might also know its length (not necessarily though). However, some aspects remain unknown until the process commenced. Particularly creative aspects (e.g. what characters occur in a particular sequence, or what do the characters look like etc.) are not known in advance.

2. Not only the process outcome is not known in advance, but also the actual process that leads to this outcome. Required process steps and iterations are not entirely predictable. For example, when producing a particular visual effect for a feature film, the organization does not
know what iterations are necessary for what process step (e.g. creating the skeleton, creating the surface, lighting). Also, required process steps may not even be known in advance as the product and its properties are further developed throughout the process. However, it is not the case that nothing is known about the process structure. For example, particular well-structured sub-processes such as review processes, or aspects of data management may be known in advance.

3. As required process steps and iterations are not entirely predictable, so are resources and involved people not known in advance. For example, within visual effects production processes in organization 2 (cf. Table 1), while the process is conducted, it may become necessary to involve further people with particular skill sets. Again, it is not the case that nothing is known about required resources. Certain resources that are required (e.g. a particular editing suite) may be known as well as resource restrictions (e.g. available time).

Based on these three observations, we propose to distinguish two facets of creativity, namely product creativity and process creativity. The study suggests that the primary reason for process creativity must be seen in product creativity; as certain features of the product are not known in advance it is hard or even impossible to predict what is needed to carry out the process. Product creativity refers to the outcome of a creative process— the creative product of which many aspects are not known in advance (supported by observation 1). If the characteristics of the output are clearly defined and there is no (intended) variance in the output, there is no product creativity. Process creativity refers to the process of creative thinking or the process of creative problem solving. Thus, it refers to the variable structure of the underlying process (supported by observation 2) as well as vague information about required resources (supported by observation 3). If the process is highly repeatable or its structure is at least pre-defined, there is no process creativity. This is usually not the case in processes that rely on creativity.

However, as has been indicated, even though there is a certain degree of unpredictability, processes in organizations undergo certain constraints regarding resources, product, and control flow: a product has to fulfill certain requirements (this is reinforced by the awareness that creative products are always purposeful) and the process is restricted by both the required resources and the availability of resources such as human resources, time, budget, and equipment and by dependencies between different process fragments. Consequently, we introduce three types of constraints that describe what elements of a pocket of creativity are known in advance: product constraints, control flow constraints, and resource constraints. When the process is carried out it must adhere to those constraints. Thus, these constraints limit the degree of freedom of a process due to business imperatives.

By specifying these constraints, that is, defining required and available resources as well as demanded characteristics of the output and required process steps, the process designer can (a) allocate resources as well as identify potential strategies or software tools in order to support a particular pocket of creativity and (b) to better plan for precedent and subsequent process steps. In a visual effects company, for example, the digital format of the deliverable of a creative task will be known so that the equipment that is used in subsequent tasks can be planned. However, knowing that certain characteristics are not known in advance, can also require the process designer to plan for subsequent process steps, such as review cycles to make sure that the (creative) product actually meets the requirements. Also, identifying the required resources of a pocket of creativity is of high importance as it is well-known that a lack of resources may even kill any creativity (Amabile 1998). At the same time, providing more than the required resources may not foster creativity. Yet, in many cases, the demand for certain resources might not be known in advance.

Table 2 provides an overview of the relationships between product creativity, process creativity, and the three types of constraints.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Constraint</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product creativity</td>
<td>Product constraints</td>
<td>Product constraints limit the degree of variance in the outcome of the process. Product constraints are important for review cycles involved in a</td>
</tr>
</tbody>
</table>
process and for subsequent sections of a process. Explicating characteristics of the product enables to define how the process can continue after a particular task. In pockets of creativity not all requirements to the product are specified in detail. That is, there is variance in the outcome of the process.

<table>
<thead>
<tr>
<th>Process creativity</th>
<th>Control flow constraints</th>
<th>These constraints describe how much of the process-flow can be pre-determined and, therefore, be explicitly modeled. In pockets of creativity, often only fragments of the control flow are known in advance.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Resource constraints</td>
<td>These constraints describe what resources are needed to carry out the process; this may involve alternatives. In pockets of creativity, not all required resources are known in advance.</td>
</tr>
</tbody>
</table>

Table 2: Types of constraints in pockets of creativity

3.3 Relating the findings to the framework

Figure 2 highlights how the different aspects of creativity can be related to the above introduced framework and thus describes the main characteristic of pockets of creativity. Process creativity becomes manifest in variable control flow and variable resource requirements at run-time. That is, both control flow and resource requirements are not known before the process is actually executed. Product creativity becomes manifest in variable product specifics. That is, the actual characteristics of the product are not known in advance but evolve while the process is executed. The three types of constraints restrict the degree of variability.

Figure 2: Characteristics of pockets of creativity

Obviously, there are varying levels of both process creativity and product creativity. However, the identification of those tasks/processes that do not involve any process or product creativity at all, may allow organizations to automate these tasks without having to fear killing any creativity.
As the above described empirical research has shown, in many cases process and product creativity occur conjointly. Yet, the data suggests that product creativity can be seen as the primary reason for process creativity. This is particularly the case within industries such as the creative industries. However, it may also be possible that, in other industries such as banking or insurance, the output is predetermined (for example due to legal requirements and other compliance issues) but the process of how to achieve that output may be creative. Notably, it is arguable whether such processes may in fact be framed as being creative as they do not necessarily generate products that are characterized by both novelty and appropriateness.

3.4 Example Case

We use an exemplary creativity-intensive process from the film industry, the so-called post-production process, in order to exemplify the concept of pockets of creativity and its applicability. This retrospective analysis is a first step towards the evaluation of the proposed concept. Within the process of post-production there are several technical and creative elements. One highly creative element is the so-called Offline-Editing. Offline Editing is the stage within the process where different pieces of footage are put together by an editor so as to tell the film’s story. Offline Editing is a complex process that comprises of both well-structured and highly creative tasks. Table 3 provides an overview of product-based, process-based, and resource based constraints that apply for the pocket of creativity of Offline-Editing.

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Example Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product constraints</td>
<td>The different output formats are known as well as the length of the sequences. This is prerequisite to continue the process and meet client requirements (e.g. the length of a TV commercial).</td>
</tr>
<tr>
<td>Control flow constraints</td>
<td>The process is highly creative and requires flexibility. Whereas the actual structure of the offline editing cannot be explicited in advance, the pocket of creativity involves well-structured sub-processes such as technical and creative reviews which have to be executed in a predefined order.</td>
</tr>
<tr>
<td>Resource constraints</td>
<td>Various constraints regarding time, budget, human resources, and equipment apply to this pocket of creativity. However, certain aspects are not known in advance. For example, the process may require the involvement of additional personal with certain expertise in offline-editing.</td>
</tr>
</tbody>
</table>

*Table 3: Exemplary constraints*

The offline edit is creative in nature: Only certain aspects of the outcome of the process are defined. The final product looks different each and every time the process is conducted. The final product is highly subjective as different people have differing understandings of aesthetics and creative quality (Firestien 1993). However, certain aspects of the product, such as format and length, may be known in advance. By explicating these characteristics of the creative product, it becomes possible to plan for subsequent process steps. For example, the format impacts on the equipment that is required to further proceed with the post-production process. The process-flow is not pre-determined. For example, the number of iterations through the creative tasks of offline-editing (rough cut) and offline editing (fine cut) are not known, or there may be need for communication in order to receive additional footage that can be used within offline editing. Even though there are resource-based constraints, such as time and budget, it is hard to predict what resources will actually be needed to carry out the task. For example, depending on client feedback it may be necessary to involve people with additional expertise in post-production into the process. Figure 3 depicts the pocket of creativity. Not that Figure 3 only presents a simplified, small part of the overall post-production process. The pocket of creativity comprises of two creative tasks and a well-structured sub-process.
Summarizing, by expatiating this pocket of creativity the process designer ensures that this creative section of the process can be supported and can also be integrated into the overall process. However, by determining what part is indeed a creative task that is not further broken down in terms of process flow, the designer avoids straight-jacketing of a creative process which might prevent creativity.

4 RELATED WORK

As has been indicated, it is not the case that there have not been attempts in order to support flexible business processes. In the following we provide an overview of the work of fellow colleagues who addressed the phenomenon of variability in business processes. In order to do so, we proceed from the more rigid approaches to the more flexible ones. Obviously, particularly the latter ones may be considered by organizations that seek to support creativity-intensive business processes.

Research on workflow evolution addresses the question of what to do when static workflows have to be adapted due to changes of the underlying business process (Casati & Ceri & Pernici & Pozzi 1996). The challenge is twofold: On the one hand, the process of modifying the existing workflow model has to be managed in order not to necessitate a complete reconstruction. On the other hand, running workflow instances must be adjusted according to the new workflow description. Existent techniques for workflow evolution provide support for occasional changes of business processes, which are generally structured and stable (Seidel & Adams & Ter Hofstede & Rosemann 2007). However, when general process blueprints are absent, there is no evolution to be managed.

Approaches towards exception handling have been discussed in order to handle such events that are not accounted for in the original process model, often referred to as exceptions. Allowing for all possible, and known, exceptions within the workflow model can quickly lead to workflow models that are hard to read and maintain. To tackle this problem, Russel et al. (2006) suggest the application of exception handling systems which separate the handling of exceptions from the main process model. Yet, an exception handler has to explicitly know about the exceptions it may handle. In consequence, such systems are incapable of dealing with flexible or even more so creative processes.

In order to allow for a higher degree of flexibility, Sadiq et al. (2005) introduce the concept of pockets of flexibility. They propose an approach to workflow modeling that enables the specification of loosely defined process sections within highly structured workflow models. These sections comprise a set of
workflow fragments and constraints, which restrict the control flows that are allowed between the fragments. For pockets of flexibility, prior to execution, the control flow is specified for every workflow instance (Sadiq & Sadiq & Orlowska 2001). This approach allows workflow processes to be tailored to individual instances in a so-called late instantiation fashion. But still, right before execution there has to be an explicit workflow model that describes process coordination. Thus, this approach is too rigid to handle creativity-intensive processes.

Note that all of the above mentioned approaches require the process designer to model the process flow in detail – at build-time for all workflow instances or right prior run-time to facilitate individual processes routing for every workflow occurrence. In response to this, over the last few years a new paradigm has emerged. The so-called case handling paradigm seeks to overcome the limitations of rigidity inherent in workflow systems by applying a data-driven approach. Van der Aalst et al. (2005) argue that most workflow systems do not reflect the way work is conducted in most non-manufacturing environments. For this reason, they propose to focuses on the whole case rather than the a single work item related to a task and to follow a data-centric approach rather than merely considering process-flow (v.d.Aalst & Berens 2001). The core features of the case handling paradigm are: provide all information available to avoid context-tunneling, decide on the activities to be executed on the basis of the information available rather than the activities executed before, separate the distribution of work from its authorization, extend the classical ‘execution’ role by additional types such as ‘skip’ and ‘redo’, and allow workers to view and modify process data outside the corresponding activities. The wide range of concepts that may be applied for specifying a case, ranging from explicitly describing the process flow to implicitly defining its structure by merely setting out post conditions related to data objects, may qualify the case handling paradigm – at least in some scenarios – for its application in creativity-intensive processes.

Groupware systems constitute a genre of IT systems that take away the focus from supporting business process coordination but promote cooperation between people. Consequently, these systems are capable of handling variability in business processes, as they do not rely on predefined process structure but foster interaction between people – a primary source of creativity (Gurteen 1998). Thus, van der Aalst et al. (2005) propose the application of groupware systems to support primarily unstructured processes.

5 CONCLUSION

In this paper we introduced an analytical framework that characterizes so-called pockets of creativity in business processes. The analytic framework we presented has been developed based on argument that rests in the examination of existent literature as well as on an empirical study on creativity-intensive processes. We suggested to distinguish between process-creativity and product-creativity and introduced three types of constraints: product constraints, control-flow constraints, and resource constraints. Those parts of a creativity-intensive process that are non-creative can be modeled and automated with conventional process modeling techniques and workflow systems. Those sections that have varying degrees of creativity however, are different: Whereas some parts may have enough structure and constraints to be supported by declarative approaches such as case handling, other parts do not have any structure at all. Here, groupware and knowledge-intensive applications can be means to facilitate creativity and enhance process performance.

It is hoped that process managers can apply these criteria in order to identify pockets of creativity in business processes so as to shed light on the relationships between well-structured and creative parts of business processes. It is further hoped that the explication of these relationships will enable to enhance process efficiency by optimizing and automating the well-structured non-creative tasks while not straight-jacketing the creative parts and, thus, compromising creativity.

This study has some limitations. Most notably, the three key observations that led to the notion of product creativity and process creativity as well as to the formulation of the three types of constraints are based on empirical studies with only three organizations from one particular industry. Thus, future
research must consider creativity-intensive processes from other domains. Future research will also focus on the exploration of strategies and IT systems that can be used in order to support such creativity-intensive processes, so as to ultimately enhance organizational efficiency and effectiveness. Moreover, one may question the practicality of the proposed approach as creativity must often deal with rigid structures and inflexibility. A response to this assertion may be the conduct of further studies in which the proposed approach is further applied and evaluated. Other aspects that have not been comprehensively covered in the present paper are the study’s impact on the business process lifecycle as well as the actual modeling notations/grammars that can be used to model the proposed constraints. In summary, it is hoped that the proposed approach can inform the development of a more comprehensive methodology.

Acknowledgments

This paper is the result of a collaborative research effort in the context of the research project ManKIP (Management of Creativity-intensive Processes, promotional reference 01FM07061) funded by the German Federal Ministry of Education and Research (BMBF) and the European Social Fund of the European Union and the ARC Centre of Excellence for Creative Industries and Innovation (CCI) (www.cci.edu.au) funded by the Australian Research Council. We gratefully acknowledge the support of the Project Management Agency as part of the German Aerospace Center (PT-DLR).

References


THE EFFECT OF INFORMATION AND COMMUNICATION TECHNOLOGIES, WORKPLACE RE-ORGANIZATION AND TRADE ON THE DEMAND FOR EMPLOYEES' SKILLS: A COMPARATIVE ANALYSIS OF GREEK AND SWISS ENTERPRISES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0472.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>IT/IS management, Business Process Reengineering (BPR), IT policy, Cross-national study</td>
</tr>
</tbody>
</table>
THE EFFECT OF INFORMATION AND COMMUNICATION TECHNOLOGIES, WORKPLACE RE-ORGANIZATION AND TRADE ON THE DEMAND FOR EMPLOYEES’ SKILLS: A COMPARATIVE ANALYSIS OF GREEK AND SWISS ENTERPRISES

Abstract

This paper aims at investigating empirically at the firm level the effect of the use of modern information and communication technologies (ICT), and also of two other factors, the adoption of new forms of workplace organization (which is regarded as a ‘complement’ of ICT), and trade (export) activities (a major characteristic of modern economy), on the demand for employees’ (vocational) education and skills. The study is based on firm-level data collected through a common questionnaire from firms’ samples of similar composition (concerning firm sizes and industries) in Greece and Switzerland; from these data econometric models of similar specification have been constructed for both countries. The results of multivariate analysis show that the intensive use of ICT correlates positively with the employment shares of high-educated personnel and negatively with the ones of the low-educated personnel. These findings are consistent with the hypothesis of ‘skill-biased technical change’. Further, the intensive use of “employee voice”-oriented organizational practices correlates positively with the employment shares of high-educated employees in both countries, and also negatively with the employment share of low-educated ones only for the Swiss firms. The results for the “work design” organizational practices are more ambiguous. Thus, there is only partial confirmation of the hypothesis of skilled-biased organizational change. Finally, we found some evidence in favour of the trade effect (export activities) only for the Swiss firms. Our results show both similarities and differences in the above aspects between Greece and Switzerland, and indicate that national context characteristics affect the relationship of the demand for employees’ (vocational) education and skills with ICT use, adoption of new forms of workplace organization and trade.

Keywords: labour(employees) skills (education) demand, labour(employees), information technology, workplace organization, trade

1 INTRODUCTION

Most of the research that has been conducted concerning the impact of information and communication technologies (ICT) on organizations has been focused mainly on their effects on organizational performance, processes and structure (e.g. Melville et al, 2004, Attaran 2004, Pearson and Saunders 2006, Albadvi et al 2007, Wan et al 2007); however, other aspects of the organizational impact of ICT, such as their effects on the demand for employees’ (vocational) education and skills, have not received much attention by the ICT researchers and practitioners, despite their high organizational and social importance. In the last 10 years there is a discussion about the observed shift of labour demand toward high-skilled personnel and its possible explanations (Machin and Van Reenen 1998, Berman et al 1998, Acemoglu 1998, Acemoglu 2002, Arvanitis 2005). One of the proposed explanations is based on the so-called ‘skill-biased technological change’-hypothesis, according to which the reason for the observed upskilling of labour force is the non-neutrality of technological change, which favours the use of skilled labour more than the use of other labour inputs; the basic reason for the most recent acceleration of technological change is the diffusion of ICT, which seem to have given new impetus to the substitution of low-skilled by high-skilled employees. However, this ‘technology’ hypothesis cannot explain the entire magnitude of the observed labour demand shift, so other possible alternative or complementary explanations have been proposed. One of them, highly associated with the introduction of ICT, is the introduction of new forms of workplace organization, such as team work, job rotation, decrease of hierarchical levels, decision decentralization (Caroli 2001); these new forms of workplace organization are regarded as a basic ‘complement’ of ICT (OECD 2003). Another alternative explanation that has been proposed (see e.g.
Wood 1995) is the internationalization of economic activities (trade), which is a major characteristic of modern economy.

The present study explores empirically the hypothesis that ICT and its basic complement, new organizational practices, are important determinants of the demand for labour of different educational levels, and further that the joint use of these two factors leads to a mutual strengthening of their impact on these labour demands. A third hypothesis refers to the influence of international trade (in this case: exports), which characterizes modern economy, on the composition of a firms’ workforce. This allows a comparison of the effects of ICT, with the ones of the new organizational practices and the trade. The analytical framework is that of a demand function for employees with different education levels (heterogeneous labour) at firm level. Also, taking into account that most of the research on the organizational impact of ICT has been conducted in a few large and highly developed countries, we have conducted a comparative empirical study of the above effects in Switzerland (a highly developed country) and in Greece (a country that does not belong to the highly developed ones, though it has made considerable economic progress in the last 15 years). Both the Greek and the Swiss part of this study are based on firm-level data collected through the same questionnaire and from samples of similar composition (concerning firm size classes and sectors), and also use the same variables and models specification, so they are comparable.

The study’s new contribution to empirical literature consists in: (a) taking into consideration not only ICT, but all three main factors forwarded in literature for explaining the shift of labour demand in favour of high-skilled labour, which allows a comparison of their effects; (b) investigating possible complementarities among all these three factors; (c) investigating the above effects in a very different national context from the ones in which most of the research on the organizational impact of ICT has been conducted: the national context of Greece (a country not belonging to the highly developed ones); (d) conducting a comparative study of the above effects between two countries, Switzerland and Greece, which are characterised by different levels of economic development.

The structure of this paper is as follows: In section 2 the conceptual framework and the research hypotheses of this study are presented. Section 3 contains a review of the relevant empirical literature. Then in section 4 the model specification and the data of both the Greek and the Swiss parts of the study are described. The results of the econometric estimates are presented and discussed for both countries in section 5. The final section 6 contains the main conclusions and some policy implications.

2 CONCEPTUAL BACKGROUND AND RESEARCH HYPOTHESES FORMULATION

The “Skill-Biased Technical Change”-Hypothesis: The Role of ICT. The shift toward more highly educated workers, which can be observed since the late sixties or possibly the early seventies in many OECD countries, appears to have accelerated in the last twenty years (see e.g. Berman et al. 1998; OECD 1998). While many factors have contributed to this increase most authors think that this effect is attributable primarily to skill-based technical change. The size, breadth and timing of the recent labour demand shift have led many observers to seek skill-biased technical change in the largest and most widespread new technology of the last years, ICT (see Bresnahan 1999; Bresnahan et al. 2002). On the one hand, high-skilled labour is a precondition for the use of ICT; for example, training in problem-solving, statistical process controls and computer skills can increase the benefits of ICT. On the other hand, highly computerized systems not only systematically substitute computer decision-making for human decision-making in routine work, but also produce a large quantity of data which needs high-skilled workers, managers and professionals to get adequately utilized. ICT capital (a) substitutes for workers performing cognitive and manual tasks that can be accomplished by following given rules and (b) complements workers in performing non-routine cognitive tasks concerning generalized problem-solving and complex communications (see Bresnahan 1999; Autor et al. 2003). Therefore our first research hypothesis is:

Hypothesis 1: There are considerable positive (negative) effects of ICT on the demand for high-educated (low-educated) employees.

The “Skill-Biased Organizational Change”-Hypothesis: The Role of New Forms of Workplace Organization. A further hypothesis put forward in the literature recently refers to the influence of the increasing diffusion and application of reorganization processes. The basic idea is that a gradual shift from rigid ‘Tayloristic’
organization (characterized by specialization by tasks) to ‘holistic’ organization (featuring job rotation, teams, integration of tasks and learning across tasks)” (Lindbeck and Snower 2000, p. 353) is taking place within firms, being an important characteristic of modern economies. The main elements of reorganization at the workplace level are (see Caroli 2001 for a survey of the literature on this subject): (a) decentralization of decision-making by delegation of relevant competences from management to lower hierarchy levels, increased involvement and autonomy of employees at the shop-floor level; (b) new working practices such as team-work (semi-autonomous work-teams, quality circles, etc.), job rotation, other forms of multi-tasking, multi-skilling, etc. Many authors seem to share the idea that changes in work organization towards more “holistic” structures, definitely require an upgrading of the skill content of most jobs related to these changes. Caroli (2001) presents a series of reasons for it. Current organizational changes increase employees’ responsibility for tasks and operations. This is not only the case for operatives but also for supervisors and technicians, whose roles, hence skills, are considerably modified by the new organizational practices. Thus, an important precondition for the successful implementation of most of these new organizational practices is the availability of a higher skilled (or higher educated) workforce. Therefore our second research hypothesis is:

Hypothesis 2: There are considerable positive (negative) effects of new forms of workplace organization on the demand for high-educated (low-educated) employees.

We tested this hypothesis for the two main new forms of workplace organization according to Black and Lynch (2002), which concern “work design” and “employee voice”, as explained in more detail in section 4. Therefore hypothesis 2 is analyzed into the following two sub-hypotheses:

Hypothesis 2.1: There are considerable positive (negative) effects of new forms of ‘work design’ on the demand for high-educated (low-educated) employees.

Hypothesis 2.2: There are considerable positive (negative) effects of new forms of workplace organization characterised with more ‘employee voice’ on the demand for high-educated (low-educated) employees.

The Trade Hypothesis: The Role of Trade. The main hypothesis is that the accelerated growth of world trade and foreign direct investment leads to a new international division of labour: the production of goods (and services) with a high content of low-skilled labour is dislocated to developing countries, while activities with a high content of high-skilled labour are concentrated in the developed countries (see e.g. Wood 2005); this results in a declining relative demand for high-skilled personnel and increasing relative demand for low-skilled personnel. Firms from developed countries with high wages (as compared to developing countries) can become internationally competitive by selling products and services with a high content of high-skilled labour. Therefore our third research hypothesis is:

Hypothesis 3: There are considerable positive (negative) effects of trade, especially exports, on the demand for high-educated (low-educated) employees.

Complementarities. The use of ICT, new forms of workplace organization and human capital build a “complementary system” of activities (Bresnahan et al. 2002, p. 341ff; Milgrom and Roberts 1995, p. 191ff.). According to Milgrom and Roberts (1990 p. 514) “the term ‘complement’ is used not only in the traditional sense of a specific relation between pairs of inputs but also in a broader sense as a relation among groups of activities”. Greenan and Guellec (1994) show in a theoretical paper that the relative efficiency of a centralized mode of firm organization in which knowledge is confined to specialized workers and a decentralized one in which every worker participates in learning depends on the technological level of the firm: “whereas the centralized style is more efficient when the technological level is low, the decentralized one becomes more efficient when the technological level is higher” (p. 173). Also, the exports result in a need for international competitiveness, so they necessitate higher use of ICT and new forms of workplace organization. Hence, our group of complementarity hypotheses are:

Hypothesis 4a: There is a positive (negative) interrelationship between technology and organization leading to a mutual strengthening of the effects of these two factors on the demand for high-educated and low-educated employees respectively (complementarity hypothesis a).

Hypothesis 4b: There is a positive (negative) interrelationship between technology and organization on the one hand and exports on the other hand leading to a mutual strengthening of the effects of these factors on the demand for high-educated and low-educated employees respectively (complementarity hypothesis b).
3 REVIEW OF PREVIOUS RELEVANT EMPIRICAL STUDIES

We concentrate here to empirical studies with a similar setting as our study, i.e. studies at the firm level that investigate the impact of at least one of the three factors ICT, organization and trade on the composition of labour demand. Table A.1 in the Appendix gives a summary of this literature. We can see that most of these studies have been conducted in highly developed countries (such as USA, Germany, UK, France, Italy, Norway, Switzerland, etc.). In most of the German studies have been found positive effects of both the ICT and the organizational factors on the share of the high-educated, mixed results for the middle-educated employees, while for the low-educated ones, in most studies negative effects of ICT have been found. Interaction terms for organization and technology were investigated in some cases, but no statistical significant effects could be identified. The studies in the USA and the U.K. demonstrate clearly the expected positive (negative) effects of ICT on the employment of high-educated (low-educated) employees, however the impact of new forms of organization is less clear than that of technology (but in most cases as expected). For French and Italian firms the influence of ICT on the employment shares of employees with different skills is less important than that of organization. Most studies could not find any discernible effects of ICT, but concerning the organizational factors there is a tendency for a positive impact for the high-educated and a negative one for the low-educated employees.

Concerning the trade effect there are only few studies taking it into account, together with ICT and/or organization, and their results are ambiguous. As we can see in the Appendix one of these studies that has been conducted by Kaiser (2001) based on German data has reached the conclusion that export activities exercise a negative influence on the employment share of skilled workers, but have no effect on the employment shares of academics and unskilled workers. Another study conducted by Maurin et al (2002) found a (partly) positive effect of trade on the employment of high-skilled employees of French enterprises. On the contrary a third one conducted by Salvanes and Forre (2003) found a positive effect of trade on the employment of low-skilled employees for a sample of Norwegian firms.

On the whole, the results of these studies show differences among countries, but should be viewed as indicative and not completely comparable, because some of the observed differences can be traced back to differences with respect to the sectors and industries covered in the studies, the specification of the organizational variables and the nature of the investigations (cross-sectional versus longitudinal approach). Further empirical research is required, including all these three factors proposed in the literature for explaining the shift of labour demand in favour of high-skilled labour (ICT, new forms of workplace organization and trade), and also covering countries of different levels of development based on similar samples, variables and models, so that comparisons can be made.

4 MODELS SPECIFICATION AND DATA

The analytical framework we adopted in this study is that of a demand function for employees with different education levels (heterogeneous labour) at firm level. In particular, we considered in this study three categories of employees: high-educated (employees with education at the tertiary level including universities, technical and business colleges), middle-educated (employees with a formal degree in vocational education) and low-educated (employees with some vocational education but without a formal degree or without any formal vocational education). The employment share for each of these categories was used as dependent variables (variables H_EDUC; M_EDUC; L_EDUC). For each of them we constructed one model; all these three models had the same independent variables, which are described in the following paragraphs. Table A.2 in the Appendix shows the definition of all the variables of these models.

As measures for technology input, particularly ICT input, we used the intensity of use of two important technologies, Internet (linking to the outside world) and Intranet (linking within the firm). The firms were asked to report the share of their employees using Internet and Intranet, not by a precise figure, but within a range of twenty percentage points (1% to 20%, 21% to 40% and so on). Based on these data we constructed two five-point ordinal variables for the intensity of use of Internet and Intranet respectively. In a further step we calculated a composite indicator for ICT by adding together the standardized values (with average 0 and standard deviation 1) of the two constituent variables for Internet and Intranet, which was used as independent variable measuring ICT use. We expect in general a positive correlation of this ICT variable with the employment share.
of high-educated employees, and a negative correlation with the share of low-educated employees; we have no a priori expectation for the relation between ICT variable and the employment share of middle-educated employees.

The measurement of the organizational inputs (extent of adopting new forms of workplace organization) is an issue still open to discussion, since there is not yet agreement among researchers on the exact definition of the ‘organizational capital’ (see Black and Lynch 2002 and Lev 2003 for a discussion of this matter). In order to choose variables for measuring the extent of changes and/or introduction and use of new organizational practices at the workplace level we draw on the definition offered by Black and Lynch (2002). They distinguished two components of organizational capital associated with innovations in “work design” and “employee voice”. The first component (work design) includes various modes of changing the occupational structure of the workplace, the number of levels of management within the firm, the existence and diffusion of job rotation, and job share arrangements. The second component (employee voice) is associated with new practices such as individual job enrichment schemes, employees having more decision competences, etc. So we constructed two corresponding composite variables for organization, one for “work design”-oriented organizational practices (ORG1) and a second one for “employee voice”-oriented organizational practices (ORG2), and used them as independent variables. The variable ORG1 is constructed as the sum of the standardized values of the following three variables: intensity of use of team-work (project groups, quality circles, semi-autonomous teams, etc.) (in a five-points Likert scale); intensity of use of job rotation (in a five-points Likert scale); decrease of the number of management levels (in a three-points scale) (see Table A.2 in the Appendix). The variable ORG2 is calculated as the sum of the standardized values of the following eight variables: decentralization-shift to employees of the competence to determine: work pace; the sequence of performing tasks; the way of performing tasks; to assign tasks; to solve emerging production problems; to contact customers; to solve problems emerging with customers (all these seven variables in a five-points Likert scale); and also overall shift of decision competences from managers to employees (in a three-points scale) (see Table A.2 in the Appendix). We expect an overall positive effect of organizational variables on the share of high-educated employees and a negative effect on the share of low-educated employees. We have no a priori expectations with respect to the middle-educated employees. The trade effect is measured in our specification by the export intensity (exports as a percentage of sales; variable EXPQ), which is also used as an independent variable.

We included in the above three models three more independent variables, which are related to workplace organization, but are not components of organizational capital per se. The first one is referring to the extent of incentive-based compensation (a dummy variable for the extent of applying employee compensation according to team-performance); the second variable measures labour flexibility (a dummy variable for the extent of using of part-time work) (both in a five-points Likert scale). With respect to the team-performance-based compensation variable the sign of the correlation to the dependent variable is not a priori clear for middle- and low-educated employees, but we expect that team compensation is considered as more adequate and/or is more often used for higher- than lower-qualified employees. The relation between part-time work and education level of the employees is in the empirical literature not clear and depends on the overall conditions of the labour market as well as its institutional framework. Further, we included two variables measuring the intensity of price competition (IPC) as well as non-price competition (INPC). In a recent paper Gersbach and Schmutzler (2006) postulate and derive theoretically two hypotheses about the market conditions under which industry-specific training is likely to occur: (a) concentration is high or competitive intensity is low, and (b) product differentiation is sufficiently strong. We consider the intensity of price competition as measured in this study as a proxy for ‘competitive intensity’ in the above theoretical context and the intensity of non-price competition as measured in this study as a proxy for ‘product differentiation’. Thus, according to hypothesis (a) intensive price competition would exercise a negative influence on training propensity; on the contrary, according to hypothesis (b) intensive non-price competition would have a positive effect on training propensity. Given that a firm’s training propensity is generally positively correlated to the demand for high-qualified employees we conclude that the above-mentioned hypotheses could be directly used as theoretical background for the two competition variables in our model. Finally, we use extensive controls for firm size and industry affiliation to account for firm- or industry-specific influences not taken explicitly into consideration. Additionally, in order to test our complementarity hypotheses 4a and 4b (section 2) we proceeded to a second estimation of the above three models (having H_EDUC, M_EDUC and L_EDUC as dependent variable respectively) with five additional interaction terms for the possible complementarities between ICT, ORG1, ORG2 and EXPQ.

The data used in the Swiss part of this study were collected in the course of a survey among Swiss enterprises, which was based on a disproportionately stratified (with respect to firm size) random sample of firms with at least 20 employees covering all relevant industries of the business sector as well as firm size classes (on the
whole 29 industries, and within each industry three industry-specific firm size classes with full coverage of the upper class of large firms); finally filled questionnaires were received from 1710 Swiss firms. The data used in the Greek part of the study were collected similarly through a survey among Greek enterprises, which was based on a ‘similar’ sample to the one of the Swiss part of the study (i.e. with the same proportions of firm size and industries classes), using the same questionnaire (translated in Greek); finally filled questionnaires were received from 281 Greek firms. In both parts of this study a non-response analysis was performed, which did not indicate any serious selectivity bias with respect to the use of ICT and new organizational practices (team-work, job rotation).

5 RESULTS

In Table 1 and Table 2 are shown the estimates of the models for the employment shares of the high-, middle- and low-educated employees in Greece and Switzerland respectively. For the Greek data we used OLS (Ordinary Least Squares) as estimation method; for the Swiss data we estimated a Tobit model in order to take into account the rather high number of observations with the value 0 in the dependent variables. We estimated these three model once without interaction terms (column 1, 3 and 5 in Table 1 and Table 2) and once again with interaction terms (column 2, 4 and 6 in Table 1 and Table 2). Since these results are only cross-section estimates, it is not possible to test directly the existence of causal relations between the dependent variables and the independent variables; nevertheless, some robust regularities come out, which if interpreted in view of our hypotheses 1 to 4 (see section 2) could possibly indicate the direction of causal links.

Technological Factors. From Table 1 we can see that for Greece the coefficient of the ICT variable is positive and statistically significant for the high-educated employees, and negative and statistically significant for the middle-educated employees as well as for the low-educated employees. For Switzerland the estimates for the ICT variable coefficient in Table 2 show clearly that the technological factors correlate positively with the share of high-educated employees and negatively with the share of low-educated employees; technology is not discernibly related to the employment share of the middle-educated employees. In sum, we found a positive effect of ICT on the shares of the high-educated and a negative effect on the share of low-educated employees for both countries. Both effects appear to be robust across all estimates. Thus, hypothesis 1 of “skill-biased technological change” receives strong support for both countries. This is the most important common finding.

Organizational Factors. From Table 1 we remark that the organizational variables appear to have only a weak influence on the composition of the workforce in Greek firms. We found a positive effect of ORG2 (employee voice) for high-educated employees and a negative one of ORG1 (work design) for the low-educated employees. From Table 2 we can see that organizational practices associated to the “work design” (variable ORG1) are positive correlated with the share of high-educated and negatively related to the share of middle-educated employees, while no effect is found for the low-educated employees. Thus, for Swiss firms practices such as team-work, job-rotation and flattening of the overall firm organization show the expected positive effect on the share of high-educated employees but a negative effect for the middle-educated employees. Significantly related to all three employment shares is the variable ORG2 measuring various dimensions of “employee voice”. We found positive effects of this variable on the shares of the high- and the middle-educated employees and a negative one for the low-educated employees. In sum, by comparing the two countries organizational factors seem to be much less important for Greek firms than for Swiss firms. Also, the decentralization of decision making, as it is measured by ORG2, shows a positive effect on the share of high-educated employee for both countries; this is the second common finding that provides some empirical evidence for the validity of hypothesis 2 of “skill-biased organizational change”. On the whole, the results for the organizational variables ORG1 and ORG2 are only partly in accordance with hypothesis 2.

Export activities Hypothesis 3 is of no relevance for the Greek firms: the export intensity does not have statistically significant effect on any of the three education-related employment shares (Table 1) On the contrary, the trade effect seems to be quite important for Swiss firms (Table), due to their much stronger exposure to the international competition than Greek firms, mainly in highly sophisticated sectors, such as pharmaceuticals, electronics/instruments, financial and other business services. The positive effect of export activities on the share of high-educated employees is in accordance with expectations, but not the positive effect for the low-educated. A possible explanation for the latter effect might be found in the fact that to the internationally competitive firms belong also enterprises, which use production techniques based on automation that can be operated by workers with low skills; so an increasing degree of production automation would lead to
an increase of the employment of low-educated employees at the cost of middle-educated employees that are typically involved in more traditional production techniques. The above results provide some evidence in favour of the trade effect (export activities) hypothesis 3 only for the Swiss firms.

Complementarity Effects. For the Greek firms we did not find any complementarity effects (Table 1); on the contrary, interesting complementarity effects were found for the Swiss firms (Table 2). The coefficient of the interaction term ICT*ORG1 is positive (and statistically significant) in the equation for the high-educated employees, significantly negative for the middle-educated employees and insignificantly negative for the low-educated employees. For ICT*ORG2 we found no effect for the high-educated employees, a negative effect for the middle-educated employees and, rather unexpectedly, a positive effect for the low-educated employees. Finally, there was a positive effect of the variable ICT*EXPQ for the high-educated employees. The result for ICT*ORG1 can be interpreted as a hint for the existence of complementarity of ICT and workplace organization, which means that besides the direct positive (negative) effects of ICT and new work design practices on the employment shares of high (low)-educated employees, there also exist indirect effects which can be traced back to the joint impact of these two factors on the employment shares. The positive effect of ICT*ORG2 with respect to low-educated employees is contrary to hypothesis 4a. On the whole, hypothesis 4a receives only partly support from these results. Finally, there is only little evidence in favour of hypothesis 4b: the interaction term EXPQ*ICT has a positive and statistically significant coefficient only in the equation for high-educated employees for Swiss firms.

Other Factors. Part-time work and team compensation are of no relevance for the Greek firms (Table 1). On the contrary, for the Swiss firms (Table 2) the variable for part-time work correlates negatively with the share of high-educated employees and the share of middle-educated employees, but positively with the share of low-educated employees. These results reflect the relative importance of various dimensions of quantitative labour flexibility for different employee categories; seemingly, part-time work is considered adequate primarily for low-educated employees. Further, compensation according to team-performance is not relevant for any of the three employee categories in both countries. With respect to the two dimensions of competition, for Greece (Table 1) we found a positive effect of the intensity of price competition (IPC) on the share of the low-educated employees, a positive effect of the intensity of non-price education (INPC) for the high-educated and a negative effect of the same variable for the low-educated are in accordance with theoretical expectation (hypotheses (a) and (b) in Gersbach and Schmutzler (2006). For the Swiss firms (Table 2) the competition effects are of considerably smaller importance; we could find only a negative effect of IPC on the share of high-educated employees, which is also in accordance to theoretical prediction.

<table>
<thead>
<tr>
<th></th>
<th>H_EDUC</th>
<th>H_EDUC</th>
<th>M_EDUC</th>
<th>M_EDUC</th>
<th>L_EDUC</th>
<th>L_EDUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICT</td>
<td>6.221***</td>
<td>5.920***</td>
<td>-2.325***</td>
<td>-1.878*</td>
<td>-3.896***</td>
<td>-4.042***</td>
</tr>
<tr>
<td></td>
<td>(0.752)</td>
<td>(0.848)</td>
<td>(0.954)</td>
<td>(1.070)</td>
<td>(0.818)</td>
<td>(0.920)</td>
</tr>
<tr>
<td>ORG1</td>
<td>0.042</td>
<td>-0.305</td>
<td>1.178</td>
<td>1.866**</td>
<td>-1.220*</td>
<td>-1.561*</td>
</tr>
<tr>
<td></td>
<td>(0.665)</td>
<td>(0.747)</td>
<td>(0.843)</td>
<td>(0.942)</td>
<td>(0.723)</td>
<td>(0.810)</td>
</tr>
<tr>
<td>ORG2</td>
<td>0.537*</td>
<td>0.604*</td>
<td>-0.068</td>
<td>-0.087</td>
<td>-0.469</td>
<td>-0.518</td>
</tr>
<tr>
<td></td>
<td>(0.276)</td>
<td>(0.309)</td>
<td>(0.350)</td>
<td>(0.390)</td>
<td>(0.301)</td>
<td>(0.335)</td>
</tr>
<tr>
<td>EXPQ</td>
<td>-0.050</td>
<td>-0.035</td>
<td>0.029</td>
<td>0.018</td>
<td>0.021</td>
<td>0.0018</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.063)</td>
<td>(0.068)</td>
<td>(0.079)</td>
<td>(0.058)</td>
<td>(0.068)</td>
</tr>
<tr>
<td>ICT*ORG1</td>
<td>0.564</td>
<td>-0.630</td>
<td>-0.630</td>
<td>-0.630</td>
<td>0.066</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.364)</td>
<td>(0.459)</td>
<td>(0.459)</td>
<td>(0.459)</td>
<td>(0.366)</td>
<td></td>
</tr>
<tr>
<td>ICT*ORG2</td>
<td>-0.014</td>
<td>-0.287</td>
<td>-0.287</td>
<td>-0.287</td>
<td>0.301*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.146)</td>
<td>(0.185)</td>
<td>(0.185)</td>
<td>(0.185)</td>
<td>(0.159)</td>
<td></td>
</tr>
<tr>
<td>EXPQ*ICT</td>
<td>0.020</td>
<td>-0.010</td>
<td>-0.010</td>
<td>-0.010</td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.054)</td>
<td>(0.054)</td>
<td>(0.054)</td>
<td>(0.046)</td>
<td></td>
</tr>
<tr>
<td>EXPQ*ORG1</td>
<td>0.021</td>
<td>-0.046</td>
<td>-0.046</td>
<td>-0.046</td>
<td>0.025</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.036)</td>
<td>(0.036)</td>
<td>(0.036)</td>
<td>(0.031)</td>
<td></td>
</tr>
<tr>
<td>EXPQ*ORG2</td>
<td>-0.005</td>
<td>-0.003</td>
<td>-0.003</td>
<td>-0.003</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.014)</td>
<td></td>
</tr>
<tr>
<td>PART-TIME</td>
<td>1.065</td>
<td>1.141</td>
<td>-2.138</td>
<td>-2.518</td>
<td>1.074</td>
<td>1.377</td>
</tr>
<tr>
<td></td>
<td>(1.355)</td>
<td>(1.374)</td>
<td>(1.718)</td>
<td>(1.734)</td>
<td>(1.473)</td>
<td>(1.490)</td>
</tr>
<tr>
<td>TEAM_COMP</td>
<td>-0.582</td>
<td>-0.611</td>
<td>1.260</td>
<td>1.211</td>
<td>-0.678</td>
<td>-0.600</td>
</tr>
</tbody>
</table>
Table 1: OLS estimates; Greece (***, ** and * denote statistical significance at the 1%, 5% and 10% test-level respectively).

<table>
<thead>
<tr>
<th></th>
<th>H_EDUC</th>
<th>H_EDUC</th>
<th>H_EDUC</th>
<th>H_EDUC</th>
<th>H_EDUC</th>
<th>H_EDUC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1.035)</td>
<td>(1.040)</td>
<td>(1.312)</td>
<td>(1.312)</td>
<td>(1.125)</td>
<td>(1.128)</td>
</tr>
<tr>
<td>IPC</td>
<td>-1.320</td>
<td>-1.486</td>
<td>-1.238</td>
<td>-1.013</td>
<td>2.558*</td>
<td>2.499*</td>
</tr>
<tr>
<td></td>
<td>(1.267)</td>
<td>(1.281)</td>
<td>(1.607)</td>
<td>(1.616)</td>
<td>(1.378)</td>
<td>(1.389)</td>
</tr>
<tr>
<td>INPC</td>
<td>2.791***</td>
<td>2.830***</td>
<td>-0.198</td>
<td>-0.161</td>
<td>-2.593**</td>
<td>-2.669**</td>
</tr>
<tr>
<td></td>
<td>(1.130)</td>
<td>(1.141)</td>
<td>(1.433)</td>
<td>(1.439)</td>
<td>(1.229)</td>
<td>(1.238)</td>
</tr>
<tr>
<td>Middle-sized firms</td>
<td>2.597</td>
<td>2.689</td>
<td>-5.988*</td>
<td>-5.740</td>
<td>3.392</td>
<td>3.052</td>
</tr>
<tr>
<td></td>
<td>(2.935)</td>
<td>(2.955)</td>
<td>(3.721)</td>
<td>(3.728)</td>
<td>(3.192)</td>
<td>(3.205)</td>
</tr>
<tr>
<td>Large firms</td>
<td>-3.488</td>
<td>-3.096</td>
<td>2.632</td>
<td>1.772</td>
<td>0.856</td>
<td>1.324</td>
</tr>
<tr>
<td></td>
<td>(3.258)</td>
<td>(3.288)</td>
<td>(4.131)</td>
<td>(4.148)</td>
<td>(3.543)</td>
<td>(3.566)</td>
</tr>
<tr>
<td></td>
<td>(2.582)</td>
<td>(2.617)</td>
<td>(3.274)</td>
<td>(3.301)</td>
<td>(2.808)</td>
<td>(2.838)</td>
</tr>
<tr>
<td></td>
<td>(6.689)</td>
<td>(6.677)</td>
<td>(8.481)</td>
<td>(5.048)</td>
<td>(7.275)</td>
<td>(7.349)</td>
</tr>
<tr>
<td>N</td>
<td>265</td>
<td>265</td>
<td>265</td>
<td>265</td>
<td>265</td>
<td>265</td>
</tr>
<tr>
<td>R² adj</td>
<td>0.321</td>
<td>0.316</td>
<td>0.171</td>
<td>0.169</td>
<td>0.207</td>
<td>0.205</td>
</tr>
<tr>
<td>F</td>
<td>12.382</td>
<td>8.636</td>
<td>5.915</td>
<td>4.017</td>
<td>7.283</td>
<td>5.275</td>
</tr>
</tbody>
</table>
Table 2: Tobit estimates; Switzerland (***, ** and * denote statistical significance at the 1%, 5% and 10% test-level respectively).

6 CONCLUSIONS AND POLICY IMPLICATIONS

6.1 Conclusions

In sum, comparing the three examined factors, which have been proposed by the literature for explaining the shift of labour demand in favour of high-skilled labour, the technology (ICT) has been found to have a strong and robust effect on the demand for employees’ skills (a positive effect on the demand for high-qualified personnel, and a negative effect on the demand for low-qualified personnel) in both national contexts (i.e. in both a highly developed country, Switzerland, and a country not belonging to the highly developed ones, Greece). On the contrary, the new forms of workplace organization in Switzerland seem to have a strong effect on the demand for employees’ skills (its pattern of effects differs between new work design and employee voice practices, with both of them increasing the demand for high-qualified personnel), but in Greece this effect is much weaker; this indicates that in the context of Greece, a country not belonging to the highly developed ones, firms have not sufficiently learnt to adapt their personnel composition to new forms of workplace organization. Similarly, the export intensity in Switzerland has a strong effect on the composition of firms’ personnel (showing a pattern different from the ones found in the limited previous empirical literature on this topic (see section 3): increasing the shares of high-educated and low educated personnel, and decreasing the shares of middle-educated personnel); however, in Greece, a country with much lower export activity, and therefore a much lower exposure to the corresponding international competition, the export intensity does not affect the composition of firms’ personnel.

By comparing the results from the two countries, both similarities and differences in the above aspects can be identified. In Greek firms the main drivers of shift to high-qualified employees is the use of ICT and the decentralization of certain competences; also, the use of ICT is the main driver of reductions in the shares of middle-qualified and low-qualified personnel. On the contrary, in Swiss firms there is a wider range of significant drivers of shift to high-qualified employees: use of ICT, adoption of new work design practices, decentralization of certain competences, export activity, and also the interaction between ICT and new work design practices, as well as the interaction between ICT and export. Swiss firms seem to be able to take a maximum out of the potential of technology and decentralization through the combination of them with appropriate human skills; in order to remain internationally competitive, Swiss firms also tend to increase the skill content of their products. These results indicate that national context characteristics affect the relationship of the demand for employees’ (vocational) education and skills with ICT use, adoption of new forms of workplace organization and trade.

6.2 Policy Implications

The results of this study have significant implications for firms’ management and government policy. Firms’ management before making ICT investments, adopting new forms of workplace organization and intensifying their export activities, should adapt their personnel skills and probably composition. Also, government policy
makers have to pay special attention to the conditions favouring the formation and growth of required human capital in the economy; it is necessary to take into account the results of this study, which indicate that nearly all the examined factors (with very few exceptions) have a negative effect on the demand for middle-educated and low-educated personnel, which can result in increasing unemployment for these groups and therefore complex and multi-dimensional social problems. Therefore they should design appropriate policies for addressing this issue. In order to offer more employment perspectives to middle-educated persons government should promote and develop not only tertiary education, but also education institutions that produce this middle-skilled personnel. For instance, in Switzerland, the system of the “normal” vocational education (“Berufsfhehr”), which is one of the two pillars of the Swiss “dual” education system that produces middle-educated personnel, has to be (further) upgraded, especially with respect to the content of education. A more difficult problem is to offer more employment perspectives to low-educated persons. Additional education and/or vocational training is one way of addressing this problem, but it is not a way accessible for all involved persons, particularly not for older ones; for such cases social partners and policy makers have to co-ordinate efforts for specific solutions aiming at the social integration of this category of employees. Finally, taking into account our conclusion concerning the effect of national context characteristics on the above aspects, it should be noted that for designing our corresponding policies both at the firm and government level we cannot just ‘copy’ solutions from other countries; instead it is necessary to exploit the relevant knowledge of other countries, but in combination with a sound knowledge of the particular characteristics and specificities of our own national context.

References


## Table A.1: Results of Previous Relevant Empirical Studies

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>ICT</th>
<th>ORG</th>
<th>EXPQ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>USA:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capelli/Carter (2000):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average wages of:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>managers</td>
<td>positive</td>
<td>pos./neg.(^{(1)})</td>
<td>n.c.</td>
</tr>
<tr>
<td>supervisors</td>
<td>positive</td>
<td>pos./neg.(^{(1)})</td>
<td>n.c.</td>
</tr>
<tr>
<td>technical workers</td>
<td>positive</td>
<td>n.s./neg.(^{(1)})</td>
<td>n.c.</td>
</tr>
<tr>
<td>office workers</td>
<td>positive</td>
<td>pos./neg.(^{(1)})</td>
<td>n.c.</td>
</tr>
<tr>
<td>production workers</td>
<td>positive</td>
<td>pos./neg.(^{(1)})</td>
<td>n.c.</td>
</tr>
<tr>
<td><strong>UK</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caroli/Van Reenen (2001):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>changes in the wage bill shares of:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unskilled manuals</td>
<td>n.s.</td>
<td>negative</td>
<td>n.c.</td>
</tr>
<tr>
<td>semi-skilled manuals</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.c.</td>
</tr>
<tr>
<td>skilled manuals</td>
<td>negative</td>
<td>positive</td>
<td>n.c.</td>
</tr>
<tr>
<td>clerical workers</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.c.</td>
</tr>
<tr>
<td>supervisors/foremen</td>
<td>positive</td>
<td>n.s.</td>
<td>n.c.</td>
</tr>
<tr>
<td>managers/technical staff</td>
<td>positive</td>
<td>n.s.</td>
<td>n.c.</td>
</tr>
<tr>
<td><strong>France:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caroli/Van Reenen (2001):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>changes in the wage bill shares of:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unskilled manuals</td>
<td>n.s.</td>
<td>negative</td>
<td>n.c.</td>
</tr>
<tr>
<td>skilled manuals</td>
<td>n.s.</td>
<td>positive</td>
<td>n.c.</td>
</tr>
<tr>
<td>clerical workers</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.c.</td>
</tr>
<tr>
<td>middle managers</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.c.</td>
</tr>
<tr>
<td>technicians</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>senior managers</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.c.</td>
</tr>
<tr>
<td>Caroli et al. (2001):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>probab. of employment increase for:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>managers</td>
<td>n.a.</td>
<td>n.s.</td>
<td>n.c.</td>
</tr>
<tr>
<td>intermediate workers</td>
<td>n.a.</td>
<td>negative</td>
<td>n.c.</td>
</tr>
<tr>
<td>operatives</td>
<td>n.a.</td>
<td>negative</td>
<td>n.c.</td>
</tr>
<tr>
<td>Maurin et al. (2002):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>change of employment share of:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>skilled non-prod.</td>
<td>n.c.</td>
<td>n.c.</td>
<td>positive</td>
</tr>
<tr>
<td>skilled prod.</td>
<td>n.c.</td>
<td>n.c.</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Germany:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gerlach/Jirjahn (1998):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>employment share of:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>workers with vocational degree</td>
<td>positive</td>
<td>n.s.</td>
<td>n.c.</td>
</tr>
</tbody>
</table>
foremen/technicians n.s. n.s. n.c.
university graduates positive positive n.c.

Bauer/Bender (2001):
employment share of:
blue-collar workers:
- unskilled n.s. n.s. n.c.
- skilled n.s. n.s. n.c.
- high-skilled n.s. n.s. n.c.
white-collar workers:
- unskilled n.s. negative n.c.
- skilled n.s. n.s. n.c.
- high-skilled n.s. n.s. n.c.

Kaiser (2001):
employment share of:
academics positive n.c. n.s.
skilled workers negative n.c. negative
unskilled workers negative n.c. n.s.

Falk (2002):
probab. of employment increase for:
university graduates positive positive n.c.
masters/technicians positive positive n.c.
vocational degree n.s. positive n.c.
unskilled workers n.s. n.s. n.c.

Hujer et al (2002):
employment share of:
high-skilled positive n.s. n.c.
low-skilled negative n.s. n.c.

Italy:
Piva et al. (2003):
log of difference of the number of:
white-collar workers n.s. n.s. n.c.
blue-collar workers n.s. negative n.c.

Norway:
Salvanes/Forre (2003):
net job creation of:
high-educated employees positive n.c. n.s.
middle-educated employees n.s. n.c. n.s.
low-educated empl. n.s. n.c. positive

Switzerland:
Arvanitis (2005):
employment share of:
high-educated employees positive positive n.c.
middle educated employees n.s. n.s. n.c.
low-educated employees negative n.s. n.c.

Notes: (1): positive: team-work, reduction of management levels, regular meetings; negative: job rotation; (2): partly positive, partly negative coefficients; ICT: information and communication technologies; ORG: workplace organization; EXPO: exports (trade); „positive“ ("negative"): statistically significant (at the test level of 10%); positive (negative) coefficient of the variables(s) for ICT, ORG, EXPO; n.s.: statistically not significant (at the test level of 10%); n.c.: not considered; n.a.: not available (for such cases in which the corresponding variables are included in the models, but the results are not explicitly presented).
### Table A.2: Definition of Model Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition and measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variables</strong></td>
<td></td>
</tr>
<tr>
<td>H_EDUC</td>
<td>Employment share of employees with tertiary-level education</td>
</tr>
<tr>
<td>M_EDUC</td>
<td>Employment share of employees with a formal degree in vocational education</td>
</tr>
<tr>
<td>L_EDUC</td>
<td>Employment share of employees without any formal vocational education or with some vocational education but without a formal degree</td>
</tr>
<tr>
<td><strong>Independent variables</strong></td>
<td></td>
</tr>
<tr>
<td>ICT</td>
<td>Sum of the standardized values of the 2 variables INTERNET and INTRANET</td>
</tr>
<tr>
<td>INTERNET</td>
<td>Six-level ordinate variable for the intensity of <em>Internet use</em>: share of employees using Internet in daily work: 0: 0%; 1: 1-20%; 2: 21-40%; 3: 41-60%; 4: 61-80%; 5: 81-100%</td>
</tr>
<tr>
<td>INTRANET</td>
<td>Six-level ordinate variable for the intensity of <em>Intranet use</em>: share of employees using intranet in daily work: 0: 0%; 1: 1-20%; 2: 21-40%; 3: 41-60%; 4: 61-80%; 5: 81-100%</td>
</tr>
<tr>
<td>ORG1</td>
<td>Sum of the standardized values of the 3 variables TWORK, JROT and LEVEL</td>
</tr>
<tr>
<td>TWORK</td>
<td>Ordinate variable measuring how widespread is <em>team-work</em> inside a firm on a five-point Likert scale (1: ‘very weakly widespread’; 5: ‘very strongly widespread’); team work: project groups, quality circles, semi-autonomous teams, etc.</td>
</tr>
<tr>
<td>JROT</td>
<td>Ordinate variable measuring how widespread is <em>job rotation</em> inside a firm on a five-point Likert scale (1: ‘very weakly widespread’; 5: ‘very strongly widespread’); team work: project groups, quality circles, semi-autonomous teams, etc.</td>
</tr>
<tr>
<td>LEVEL</td>
<td>Three-level ordinate variable for the change of the number of <em>managerial levels</em> in the period 2000-2005: 1: increase; 2: no change; 3: decrease</td>
</tr>
<tr>
<td>ORG2</td>
<td>Sum of the standardized values of the 8 variables COMP_OVERALL, COMP_WORKPACE, COMP_WORKSEQ, COMP_WORKASSIGN, COMP_WORKWAY, COMP_PRODUCTION, COMP_CUSTOMER_CONTACT and COMP_CUSTOMER</td>
</tr>
<tr>
<td>COMP_OVERALL</td>
<td>Three-level ordinate variable measuring the change of the distribution of decision competences between managers and employees inside a firm in the period 2000-2005: 1: shift towards managers; 2. no shift; 3: shift towards employees</td>
</tr>
<tr>
<td>COMP_WORKPACE</td>
<td>Ordinate variable measuring the distribution of decision competences to determine work pace (1: 'primarily managers'; 5: 'primarily employees')</td>
</tr>
<tr>
<td>COMP_WORKSEQ</td>
<td>Ordinate variable measuring the distribution of decision competences to determine the sequence of the tasks to be performed (1: 'primarily managers'; 5: 'primarily employees')</td>
</tr>
<tr>
<td>COMP_WORKASSIGN</td>
<td>Ordinate variable measuring the distribution of decision competences to assign tasks to the employees (1: 'primarily managers'; 5: 'primarily employees')</td>
</tr>
<tr>
<td>COMP_WORKWAY</td>
<td>Ordinate variable measuring the distribution of decision competences to determine the way of performing tasks (1: 'primarily managers'; 5: 'primarily employees')</td>
</tr>
<tr>
<td>COMP_PRODUCTION</td>
<td>Ordinate variable measuring the distribution of decision competences to solve emerging production problems (1: 'primarily managers'; 5: 'primarily employees')</td>
</tr>
<tr>
<td>COMP_CUSTOMER-CONTACT</td>
<td>Ordinate variable measuring the distribution of decision competences to contact customers (1: 'primarily managers'; 5: 'primarily employees')</td>
</tr>
<tr>
<td>COMP_CUSTOMER-PROBLEMS</td>
<td>Ordinate variable measuring the distribution of decision competences to solve emerging problems with customers (1: 'primarily managers'; 5: 'primarily employees')</td>
</tr>
<tr>
<td>EXPQ</td>
<td>Exports of goods and services as a percentage of sales</td>
</tr>
<tr>
<td>------</td>
<td>------------------------------------------------------</td>
</tr>
<tr>
<td>ICT*ORG1</td>
<td>Interaction term of the variables ICT and ORG1</td>
</tr>
<tr>
<td>ICT*ORG2</td>
<td>Interaction term of the variables ICT and ORG2</td>
</tr>
<tr>
<td>EXPQ*ICT</td>
<td>Interaction term of the variables EXPQ and ICT</td>
</tr>
<tr>
<td>EXPQ*ORG1</td>
<td>Interaction term of the variables EXPQ and ORG1</td>
</tr>
<tr>
<td>EXPQ*ORG2</td>
<td>Interaction term of the variables EXPQ and ORG2</td>
</tr>
<tr>
<td>Part-time work</td>
<td>Ordinate variable measuring how important is <em>part-time work</em> inside a firm on a five-point Likert scale (1: 'not important'; 5: 'very important')</td>
</tr>
<tr>
<td>Group compensation</td>
<td>Ordinate variable measuring how important is <em>compensation by group or team</em> inside a firm on a five-point Likert scale (1: 'not important'; 5: 'very important')</td>
</tr>
<tr>
<td>IPC</td>
<td>Ordinate variable measuring the intensity of <em>price competition</em> at a firm’s main market on a five-point Likert scale (1: 'very weak'; 5: 'very strong')</td>
</tr>
<tr>
<td>INPC</td>
<td>Ordinate variable measuring the intensity of <em>non-price competition</em> (competition with respect to quality, customer services, etc.) at a firm’s main market on a five-point Likert scale (1: 'very weak'; 5: 'very strong')</td>
</tr>
<tr>
<td>Middle-sized firms</td>
<td>50 to 249 employees</td>
</tr>
<tr>
<td>Large firms</td>
<td>250 employees and more</td>
</tr>
</tbody>
</table>
Counteracting Forces in Implementation of IS-enabled Global Business Processes

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0596.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Cross-national study, Global information systems, Business Process Improvement, Enterprise resource planning (ERP) (packaged systems)</td>
</tr>
</tbody>
</table>
COUNTERACTING FORCES IN IMPLEMENTATION OF IS-ENABLED GLOBAL BUSINESS PROCESSES

Sannarnes, Møyfrid Kårstad, University of Agder, Service Box 422, NO-4604 Kristiansand, Norway, moyfrid.sannarnes@uia.no

Munkvold, Bjørn Erik, University of Agder, Service Box 422, NO-4604 Kristiansand, Norway, bjorn.e.munkvold@uia.no

Andersen, Kim Normann, Copenhagen Business School, Howitzvej 60, DK-2000 Fredriksberg, Denmark, andersen@cbs.dk

Abstract

The international industry of engineering products and services is characterized by high complexity and competition. Corporations that expand globally have experienced that managing interdependent activities and business processes across several countries requires an effective deployment of advanced information technology. Whereas the literature has described implementation of global information systems as a means to coordinate and control the business processes, empirical studies have shown that introducing a large-scale information systems involves several managerial challenges when organizations are geographically dispersed. This paper studies deployment of a global enterprise system to support evolvement of global business processes. On the basis of a qualitative case study of a multinational corporation implementing an enterprise system across several geographical locations, we identify counteracting forces in the process of global standardization of IS and business processes and discuss how the organization try to manage these forces and challenges therein. The findings suggest that global business processes develop through diverse processes of learning and negotiation between local practices of use and infusion of the global enterprise system.

Keywords: Global business processes, Standardization, Enterprise systems, Multinational Corporation
1 INTRODUCTION

The supplier industry of engineering products and services is characterized by high international competition, and success increasingly depends on expanding the global reach of the organization. To support coordination and integration of global business processes and facilitate cross-border collaboration, western firms tend to deploy advanced information technology (Walsham 2001). The resulting business process network needs to evolve dynamically to adapt to continuously changing global and local business conditions. Several authors argue that the global enterprise should balance in a continuous learning and innovation between the local practices and global standardization (e.g. Begley & Boyd 2003).

We use the term global information systems to describe the development, use and management of information systems in a global context. During the last decade many organizations have implemented large-scale enterprise systems (ES), defined as modular information systems that cover key functions of a company and enable integration and coordination of business processes within the company and with business partners (Davenport 1998). Deployment of ES to support cross-border operations in multinational corporations (MNC) is particularly challenging and typically leads to significant changes in organization, technology and business processes (Biehl 2007). Whereas much of the literature has described implementation of global IS as a means to control the processes, empirical studies have shown that introduction of information systems in global organizations often has unintended consequences (Ciborra 2000) and may even lead to less control (Hanseth et al. 2001). Despite a large body of research on ES implementation and use, reviews of this literature show that the majority of these studies are conducted in a domestic context. Most studies attempting to capture differences between different cultures are limited to one or two countries (Moon 2007), and research studies involving a heterogeneous collection of applications and user communities are limited. There is thus a need for more studies on how ES can support cross-border collaboration on business processes. Also, more in-depth research on the effect of competitive environment, business strategy and strategies to resolve misfits in multinational ES environment is asked for (Sheu et al. 2004).

The tension between standardization and customization has been subject for several IS studies. There is an extensive body of research emphasizing the need to adapt to requirements of the local context of use (Davenport 1998). The need for standardization and coordination has been focused in IS management studies in organizations operating internationally (Lehmann & Gallupe 2005). The research question addressed in this paper is how a multinational corporation balances the forces for and against standardization in the process of global IS deployment. We draw upon findings from a qualitative case study in a multinational engineering enterprise headquartered in Northern Europe. The company is part of a global group of organizations providing advanced engineering and construction services and technology products worldwide. The paper explores the evolvement of a global ES that was successively introduced within the global company and its subsidiaries. The study illustrates that introducing and exploiting information systems in a global organization can be seen as a standardization effort, but is always embedded in local situated practices. We use the concept of dialectics of change from Van de Ven and Poole (1995) as a theoretical lens to guide the analysis of the changes and counteracting forces related to global standardization and local adaptation during implementation and post-implementation of the enterprise system.

The paper is structured as follows. The next section discusses the research foundation for global IS and business processes crossing national boundaries. Section three presents an overview of the case research setting, while section four introduces the research approach, data collection and analysis. Findings from the case study on global IS implementation based on analysis of counteracting forces of change are presented in section five. Concluding remarks and implications for further research are presented in the last section.
2 RELATED RESEARCH

This section provides a brief overview of research on IS deployment to support development of global business processes. Further, a theoretical lens for analyzing the process of development and change in organizations is presented.

Globalization of business processes implies that different organizational units need to be coordinated and integrated across geographical and organizational boundaries. Control and governance are core issues underlying IT strategies (Hanseth et al. 2001) and global IS research has traditionally focused on the potential of IS for control and coordination, global efficiency and transfer of learning (Lehmann & Gallupe 2005). However, deployment of large-scale, global IS is constrained by an evolving socio-technical installed base of information, systems, artifacts, organizational structures and practices (Rolland 2000). Establishing a global infrastructure plays an important role, by supporting global business processes and facilitating coordination and collaboration across time and space (Walsham 2001). An empirical study of a global IS infrastructure implemented in a maritime company found that it is necessary to strike a balance between global requirements and being local when designing large-scale IS. This study also illustrates the largely invisible costs to achieve a working solution, locally and globally (Rolland & Monteiro 2002).

A conceptual model for global IS suggested by Lehmann (2001) employs a two-dimensional typology consisting of a common core and local variations of the IS. The model suggests that it is impossible to support global organizations with a one-size-fits-all information system. Instead, there is a core of systems compulsory for all users and local systems to provide functionality specific for one or a group of regional subsidiaries. Lehmann and Gallupe (2005) characterize the implementation of a global IS as a territorial force field between the central/global and local, and a functional force field between business and IT.

![Force field dimensions surrounding a global information system](Source: Lehmann & Gallupe 2005, p. 170).

Enterprise systems represent one category of global IS to support cross-border operations. There have been an increasing number of research studies on ES implementation during the recent years (Lorenzo et al. 2008; Moon, 2007; Sheu et al. 2004). The fit between ES and business processes is regarded critical for
the success of ES implementation (Robey et al. 2002; Soh et al. 2003). Companies that implement ES should thus simultaneously change and improve their business processes to take advantage of new data, but most organizations find it demanding to make process change during implementation (Davenport 2004). ES deployment has been described as a means for increased coordination and control in global organizations by facilitating enterprise-wide integration of information across functional and geographical boundaries. Whereas introducing ES is challenging to any organization, multinational implementation adds new dimensions of complexity due to national, cultural, organizational, and technical differences (Markus et al. 2000; Sheu et al. 2004). Research studies find that changes such as introducing ES often has unintended side-effects in global organizations (Ciborra 2000). For example, a study by Hanseth et al. (2001) illustrates that implementing an ES system to enhance control over a global organization may just as well deliver less control from a managerial perspective.

The majority of previous studies on large-scale ES implementation have adopted a variance approach, typically describing critical success factors or value that companies generate from the ES implementation (e.g. Biehl 2007). Other studies apply a process perspective, focusing on how change develops over time in a sequence of stages (Markus et al. 2000). These stage models match the life cycle motor from Ven de Ven and Poole (1995), who presented four basic “motors” to explain processes of development and change in organizations; life cycle, teleology, dialectic and evolution. The dialectic motor emphasizes colliding forces or contradictory values that compete for domination or control. Organization theories that incorporate “logic of contradiction” have been valuable to explain change in organizations as the result of countering forces, those promoting organizational change and those opposing it. Robey et al. (2002) use the dialectic perspective in a comparative analysis of ES implementations in several industrial firms. In their view, the dialectical theory is valuable to situations where two or more parties have opposing views, values and power, while still working towards a common goal. In this paper, dialectics of change is used as a theoretical guidance to explore how global business processes and global IS develop and change during implementation and post-implementation of a global enterprise system. Introduction and development of global IS can be described as change processes, influenced by countering forces that promote standards and stability, and forces opposing global standardization but promoting local adaptation and innovation.

3 CASE OVERVIEW

The empirical study is conducted in ‘Eng-Construct’ (EC), a global group of organizations providing engineering and construction services and technology products worldwide. EC was established in 2004 after an organizational merger by two of the leading organizations within the industry. The multinational corporation headquartered in Northern Europe serves several industries, including oil and gas, and chemicals. It employs approximately 24,000 people in about 30 countries. TPS (‘Tech Product and Services’) is one of the companies within the EC group. TPS has been in the industry for more than thirty years and delivers high-tech engineering products for drilling rigs, both off-shore and land-based installation. The company has also a large, global support service network to support customers in different countries. During the recent years, the company has experienced a large international growth, which has increased the overall complexity of the organization and has now approximately 2500 employees located at the headquarters and regional offices and subsidiaries in Europe, Asia, Northern and Southern America. The products and services provided by TPS need to be competitive in the global market, leading to a continuous pressure on this and other companies within the EC group to work more effectively and efficiently without compromising on quality and safety. The core global processes of TPS are sales, procurement, supply chain management, engineering, customer service and after sales services.

The EC group initiated a global change program in 2004, aiming to coordinate and standardize strategic information systems and harmonize business processes across organizational and national boundaries. One
part of the program was deployment of global IS to harmonize business processes, globally and locally. An enterprise system, SAP, consisting of application areas or modules such as Finance, Payroll, Knowledge management, Supply chain management and Customer service, was implemented through several waves in companies and subsidiaries within the group. The EC group established a central, large-scale ES implementation project team, with IT/IS personnel from the central organization and a large number of external consultants and project managers. In 2005, TPS started the implementation of the global ES at the headquarters. Since then, it has been rolled-out successively in the subsidiaries over a period of three years, till summer 2008. Other central parts of the global IS are the new corporate Intranet, project system and collaboration applications. The global IS, consisting of global and local applications, serves to support development and improvement of the global business processes. Currently, several new initiatives are thus put into operation in parallel to enhance utilization and reaching the anticipated benefits of the global ES implementation, demonstrating the ongoing need for further improvement in global IS support. These projects address various process improvements related to for instance finance, engineering and supply chain.

4 RESEARCH METHOD

The applied research approach is an interpretive case study, assuming that people create and associate their own subjective meanings as they interact with the world around them. The specific context where the global IS is introduced, the particular business area, and characteristics of the multinational corporation are all considered important elements to focus (Walsham 1995). Previous empirical research on global IS is used to guide the research study, thus iteratively linking theory and data (Eisenhardt 2002).

The data reported in this study is part of a larger research study aiming to explore how global IS and business processes emerge in a global organization. The paper reports from twenty interviews conducted through two phases during fall 2007 and spring 2008. The first phase was a ‘screening’ consisting of three open interviews with functional managers within human relation (HR), IS/IT and global project management. The aim of these interviews, which lasted up to two hours, was to get more background information about the global company and the specific industry in which it operates, core business processes and strategy for deployment of global IS. Also, key informants for the next rounds of interviews were identified. During the second phase, seventeen interviews lasting from one to two hours were conducted. The interviews were semi-structured, and the informants were encouraged to talk freely about their experiences with and perceived outcome of the global ES implementation related to local and global business processes within their function area, challenges and expectations for future use. The interviews were conducted at the informant’s office or in a meeting room. One of the interviews was a group interview. The informants were managers from different business areas in Norway, UK and India, including HR, quality management, IT/IS, procurement and customer service. Also, global process owners, project managers and super users involved in the implementation at four different sites were interviewed. All interviews were recorded and transcribed. In addition, access was given to attend meetings and improvement workshops where notes were taken. Secondary materials such as process models and work-flow diagrams, collected during some of the interviews and on request, provided important contextual information on the process. The analysis of the empirical data proceeded in an iterative way between interpreting interview transcripts, secondary data and notes from meetings (Miles & Huberman 1994). Theory of dialectics of change was used as a ‘sensitizing device’ to analyze the findings from the empirical study.
5 FINDINGS

In this section we present some findings from the case study as counteracting forces for change towards global standardization and cooperation on global business processes.

5.1 ES strategy and motivation

Research on success factors of ERP has documented the need of a clear strategy and goals for implementation. Some of the informants from the headquarters explained that the main motivation for implementing the global ES was a need to replace an old legacy system that did not support international expansion in a satisfying way. According to one IS/IT manager, the new global ES was: “... a prerequisite for being able to manage a transformation from being a company with some international installations to becoming more global and handle the increasing number of projects and activities globally.” Another motivation for global standardization was that the company had been through processes of mergers, and there was a need for enhanced integration of operations. A common, global ES was expected to enable integration of work practices and information flow and integrate functionally oriented applications at different business units.

Informants at the local level, who typically were introduced to the implementation project at a later stage, asked for a more clear and articulated strategy from the group, pointing the direction for the change processes. One of the informants from the subsidiary emphasized the need of balancing between a global vision and local practices: “...there should be a global vision. And each business unit should try to support that global vision. And well, they may have some specific needs on their own, because their business may be slightly different. But it should still try to meet the global vision. And, again, if we step back down the process, our process owners need to make sure that the process meets both the global and the local needs”. The informants from the headquarters reported that there was a focus on communication and information during the first phases of implementation. However, the findings illustrate the need of continuous updated information during roll-out to subsidiaries and in post-implementation phase.

5.2 Towards global business processes

To support increasing international operations and expansion of the company TPS defined and implemented global business processes that should interconnect previous separate practices across various locations. Some of the managers stated deployment of a global ES was fundamental to manage and enhance global business processes such as customer service and supply chain management. The ES project manager said: “What we see, is that the quality is not better than how we manage to coordinate our activities, work in a similar way. Without a global IS infrastructure and routines, there are too many alternatives. By implementing enterprise systems such as SAP we get an opportunity to have a common terminology and interface.”

Some informants experienced that reaching consensus on business processes was challenging at some business areas and organizational units. Establishing the governance model with global business process owners and super users was reported as a critical factor in the process. Dedicated global process owners were appointed during roll-out to subsidiaries and these persons became central in process mapping and modelling and post-implementation activities to harmonize the global business processes. Their involvement as border-crossing facilitators was regarded essential for the success of the implementation, due to their knowledge about the processes and the business. One of the external project managers stated: “The global process owners are very important for establishing collaboration and development of global processes across geographical locations. The person has profound competence on the particular business process, but should also be a facilitator to ensure that the overall goals are achieved”. The global process owners filled the role as “translators” between the consultants, who knew how the enterprise system worked technically, and the key users, who had business and local process knowledge. One of the global
process owners emphasized that development and cultivation of global business processes will be an ongoing learning process: “a job that is started, but by no means completed. By now, it is established in to-be process models, but we still have a way to go…”.

5.3 Configuration of global ES

The aim was to implement an SAP industry solution. Some of the informants emphasized that this was one of the first times that the organization should adapt to a standard application. Still, a central part of ES implementation is technical configuration to customize the solution to the company needs. One manager at the headquarters said that the organization was not prepared on the dimension of the configuration job. The central project team, consisting mainly of consultants with a Scandinavian background, configured a “SAP light” version that was rolled out at subsidiaries. The configured version was thus very dependent of the consultants’ interpretation of the business needs, globally as well as locally. One of the key users from a subsidiary expressed concern with respect to the configuration and the role of the consultants and the central implementation team: “At the courses they presented: This is how SAP works. This is not correct, this is how they have configured it to work”. Training of super users on how the ES is configured is thus important for further use of the system.

Another important part of the ES configuration is definition of roles with access to functions and responsibilities, which proved to be a challenging task. The IT/IS manager emphasized that this will be an important, ongoing activity: “…there has been a walkthrough of the organization’s roles and functions, how we collaborate on business processes. And this deals not only with the system, it is a process that means you have to look at who is responsible for which tasks. And this in many respects has had an ‘upbringing’ effect on the company, centrally as well as locally, to relate to a more structured process”. Moreover, walkthrough of processes in the standard configured solution revealed local differences, e.g. in employee-management relationship.

Some of the informants raised serious concern with the central goal of providing ‘one-size-fits-all’ solution in a heterogeneous MNC. Particularly one manager at a smaller subsidiary emphasized large costs in implementing and learning a complex ES and few benefits for a smaller unit due to what he described as inflexibility in procedures and inability of the global system to adapt to specific local needs.

5.4 Knowledge sharing and learning

Experiences expressed by informants working with the customer service process illustrate the dialectics between old knowledge embedded in the process, “the way we have always worked”, and the new global ES. The process of learning required unlearning old ways of working, which can be challenging. Training is necessary to increase knowledge about the ES, which in turn is expected to enhance utilization and value realization from the ES investment. As commented by one of the global business owners: “In SAP, there is so much important information that could be utilized in a better way, for improved decision making - we do not utilize the potential of the system well enough yet”. There were examples that lack of ES knowledge resulted in “workarounds” where employees found their own way of solving the problems, which again resulted in different work practices. This was explained partly by a need for more training and exercises related to actual work processes. Another explanation was that implementation of “industry best practices” and corresponding routines had not yet been translated into local work practices. Also, knowledge sharing was mentioned as critical for learning and capability transfer. The project manager, for example stated: “One of the most important barriers to organizational learning is that there is not established well-functioning arenas for knowledge sharing. We certainly need to improve our ability and willingness to share knowledge between the organizational units and between headquarters and subsidiaries”.

Proceedings ECIS 2009
Process knowledge is critical for establishing global business processes. Some initiatives were mentioned as important to overcome lack of process knowledge. Especially at the last rollout in Asia, the global process owners and super users were actively involved at workshops where work processes and use of ES were modelled, discussed and ‘negotiated’. Also, written procedures and user guidelines such as “SAP for dummies” for the customer service process available from the Intranet were important for enhancing knowledge of global business processes and improving the processes. Development of e-learning program supporting classroom training is another initiative mentioned.

5.5 Cross-cultural and cross-functional communication

The core activities of the companies are often project driven, and there is a need for cross-border collaborative work. Several of the informants emphasized that cross-functional cooperation and communication is crucial for developing and maintaining global business processes. One of the super users described his as a coach for a team in which everyone has a particular role to play.

Cultural diversity was reported as an important factor in the case. Differences in national culture influenced the process, for example when discussing and modeling new business processes. One of the super users noted: “In our country, if someone disagrees he or she usually speaks out. This is not necessary the case in other cultures”. A cultural awareness initiative has been launched in the company for enhancing cross-cultural work. One line manager emphasized that an attitude to information sharing and collaboration cannot be taken for granted: "There are cultural differences, for example developing an understanding within the organization that we have to share work and knowledge between different countries, locations. That is a new situation for us”. Moreover, the implementation revealed important differences at the geographical locations. For example, some of the informants described that similar roles could have different authority across different locations, demonstrating that the organizational culture has an effect on definitions of roles.

Managing of the IT consultancy process was reported to be most challenging during the first rollout to smaller subsidiaries as the project team was not familiar with the subsidiary’s particular needs and the limited internal project resources. One informant at a subsidiary argued that a more global implementation team should have been established from project start-up.

5.6 Change processes on global and local level

During the implementation project the company experienced changes in roles and responsibilities. The global company has a heterogeneous work force regarding age, and professional and educational background. Technical skill and process knowledge were also critical factors that varied a lot between the different units and locations. The company was not prepared for this challenge, and one of the functional managers commented how the new global ES made work more transparent: “One of the outcomes of the change was that some of the employees became more clever to do their job, whereas others actually became less competent. Because writing skills etc. became more visible, now people had to do things by themselves. That was an important learning experience”. Job design, job enlargement and job simplification were activities during this process. An organizational change management (OCM) role was established at the local business units to ensure information, communication and training. The OCM and line managers played important roles in this phase.

At the group level a shared service model was established covering the functional areas of finance, human resources, payroll, time and travel administration. The findings show that this represented a major organizational change that had an effect on the employees’ and managers’ daily work. The HR manager explained how the new global ES and increased international focus need to be learned and cultivated in TPS, “a process that will take time in an organization, in which delivering engineering product and services has had a local focus with use of mainly locally developed applications. This is a new situation for us, going from being a locally managed organization with international projects, to becoming an international company”. Informants from the headquarters and subsidiaries expressed different
experiences with the shared service model. One informant at a smaller subsidiary expressed the view that the shared service model did not support the smaller units well enough.

After the last rollout in May 2008, the company experienced a need to go back and take a fresh look at the implementation and routines established at the head office in light of the experiences and discussions from the local rollouts. Areas such as training at different levels, further development of global business processes with updating of corresponding routines and guidelines were decided. Moreover, TPS has launched several initiatives aiming at continuous improvement of the global IS. One of the initiatives is the establishment of an improvement board. The main objective of the improvement board, which consists of line managers from the different functions and subsidiaries, is to prioritize changes to improve global business processes and global IS. Evaluation of the global ES through user surveys at different locations is another example of improvement initiatives.

6 DISCUSSION

In this section we discuss some main challenges and counteracting forces for change related to introducing a global ES in a multinational engineering enterprise. A global ES implementation could be characterized by forces at different levels (Lehmann & Gallupe 2005). In this paper the focus is on the global-local tension. From a central coordination point of view there was a need for standardization of modules, processes and routines. From a local standpoint the global IS should support the local needs and requirements. This tension could be described as a ‘territorial force field’ (Lehmann & Gallupe 2005). Table 1 summarizes key findings from the empirical study regarding counteracting forces promoting and opposing changes to globalization and standardization of IS and business processes. While some of the forces could be described at both the local and central level, the table shows where the force predominately occurred.

<table>
<thead>
<tr>
<th></th>
<th>Local level</th>
<th>Company level</th>
<th>Central level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promoting forces</td>
<td>Improved project delivery and control</td>
<td>Economies of scale and scope</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enhanced customer focus</td>
<td>Increased integration, coordination and control</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Access to accurate information/capabilities</td>
<td>Capability transfer and change management</td>
<td></td>
</tr>
<tr>
<td>Opposing forces</td>
<td>Knowledge barriers (global processes, ES functionality)</td>
<td>Knowledge barriers (local processes, roles, practices of use)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Integration costs, local adaptation, established practices</td>
<td>Integration costs, environmental changes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lack of communication</td>
<td>Diversity (national, cultural, organizational)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Implementation of IS-Supported Business Processes: Promoting and Opposing Forces for Global Standardization at Local and Central Level

We discuss some of the main findings from the study, challenges that were encountered and initiatives to overcome these.

Deployment of global ES and business processes can be understood as a dialectic of learning and knowledge (Robey et al. 2002). The findings illustrate that there is a dialectic between existing
knowledge, embedded in business processes, routines and practices associated with legacy systems and manual routines, and new global business processes and routines supported by the global ES. Based on our findings, we suggest that lack of knowledge is a major barrier to overcome. At the individual, local level, the informants emphasized that there was a lack in knowledge of ES functionality, but also in process knowledge related to assimilation of new work processes. The case demonstrates that moving from being a company with international presence towards global integration is a learning journey by itself, for top managers, functional managers as well as employees. Configuration of a complex ES was another reported knowledge barrier. The ES was configured by the consultants in the project team and the organizations experienced that limited knowledge about the configuration was a limitation for understanding and further exploitation of the enterprise system.

Initiatives undertaken to overcome knowledge barriers are courses at different levels, focusing on practical use of the ES and how local practices are related to the overall global business process. E-learning programs and user guidelines are developed in supplement to classroom training for enhanced global services. The project team and consultants played an important role in bringing in new knowledge and distributing knowledge to different business units in TPS. However, the findings demonstrate that an active involvement of global process owners and super users with process knowledge in workshops and training at subsidiaries was necessary in this process.

Dialectics may also describe the tension between established, country-specific practices across the MNC and requirements for new, common work practices related to global business processes. The heterogeneity of TPS, consisting of geographically dispersed units of different sizes, activities and history, involves many different local practices based on functional areas of expertise. Thus, the findings demonstrate that development of global business processes could be characterized as iterative moulding and shaping processes to fit the needs of the organization. There is a need of infusing global strategies into local practices by communication of goals and strategies, as well as mapping local information and practices of use into global structures. Moreover, introduction of a global enterprise system does not necessarily imply that local practices become standardized across different business units, departments and cultures.

There are counteracting forces between a global ES promoting integration and standardization and the local need for differentiation. The findings illustrate that deployment of the global ES was regarded as a prerequisite for managing increasing global operations, but also adding complexity by introducing technical, organizational and business process changes (Markus & Tanis 2000) at the headquarters and subsidiaries. The system needs to be integrated with existing applications, physical networks, organization, culture and work practices (Hanseth & Lyytinen 2004), illustrating that information systems are always embedded in situated practice of use (Lehmann & Galupe 2005). During the period, the company has experienced a large growth, and the new system has to ‘align-in-action’ (Ciborra 1997) with the organization. The findings demonstrate that the need to align and integrate the global ES with the organization and business processes was regarded as particularly challenging and will be an ongoing process due to changes, for example in environmental conditions and mergers and acquisitions. While a stage model for IS implementation (e.g. Markus et al. 2000) seems appropriate for the first phases of the implementation, post-implementation could be characterized as gradual adoption of the ES, consisting of iterations of learning and improvements. Contrary to research treating implementation of ES as distinct stages, we find that implementation of global ES should adopt a long-term perspective and be regarded as long-time improvement and adaptation at different levels.

Establishment of a governance model with global process owners, super users and application managers was regarded essential in the case, both for consensus-making in discussions and modeling of business processes during rollout, ensuring ownership and further development of global IS and global business processes. The improvement board for prioritizing improvements across business units and functions was also reported as an important initiative. Evaluation of the global ES involves an increasing number of stakeholders as the ES is deployed globally.
Introducing a global IS to support global business processes has several implications that need to be discussed to find a balance between the global standards and local adaptation of global information systems and business processes. National and cultural differences such as government regulations, roles and authority need to be dealt with (Sheu et al. 2004). The need for communication of project goals and strategic guidelines seems to have been underestimated when rolling out at subsidiaries and in post-implementation phases of the enterprise system.

7 CONCLUSION AND IMPLICATIONS

In this study we have explored counteracting forces in change processes during deployment of an enterprise system in a MNC, in the context of global business process evolvement. The in-depth study is conducted in a company operating in the oil and gas industry, based on interpretations of informants from the corporate headquarters and three subsidiaries.

Our findings demonstrate the complexity of introducing a global ES into local work contexts, both at the headquarters and subsidiaries. The multinational company includes local sites and local practices of use. The case has illustrated that deployment of ES to support global business processes in a multinational enterprise can be described as a long-term learning process in which lack of functional ES knowledge and process knowledge related to assimilation of new work processes are major barriers to overcome. Standardization of IS and business processes implies that subsidiaries are often faced with changes imposed rather than designed. Our findings illustrate a need for infusing global strategies into local practices as well as mapping local information and practices of use into global structures. Global process owners and super users at local units and headquarters should play a critical role in these activities, which have been characterized by the informants as a continuous learning and negotiating process. Moreover, ownership and collaboration initiatives on global business processes are essential for value realization of the global ES, including ability and willingness to collaborate and share knowledge and capabilities across national and organizational boundaries.

The paper contributes to the body of global IS research by exploring deployment of enterprise systems in an organization that is geographically dispersed. By examining the phenomenon from the perspective of both headquarters and subsidiaries, the paper has explored counteracting forces at the territorial, functional and cultural levels and discussed how one multinational is dealing with these. These findings thus contribute both to research and practice. The case study explores one MNC in a particular industry over a limited period. The interpretive case study has raised several issues that should be considered in future research on global ES and business processes, and indicates the need for more in-depth research studies on how to improve the process of realizing the intended value of the global IS. The effect of the shared service model on heterogeneous units and processes of local involvement in development of global IS are examples of areas for further research.

References


AN APPROACH TO ASSESS THE IMPLEMENTATION OF BUSINESS PROCESS MANAGEMENT IN ENTERPRISES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0670.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Process Management, Capability Maturity Model Integration (CMMI), Management practices, Organisational Change</td>
</tr>
</tbody>
</table>
AN APPROACH TO ASSESS THE IMPLEMENTATION OF BUSINESS PROCESS MANAGEMENT IN ENTERPRISES

Michael Rohloff, University of Potsdam, August-Bebel-Str. 89, 14482 Potsdam, Germany
michael.rohloff@wi.uni-potsdam.de

Abstract

Business Process Management is an important management practice for business transformation and organizational change. This paper presents a Business Process Management implementation approach in a large international company. It introduces a process management maturity assessment which was developed to assess the implementation of Business Process Management and the achievements. The maturity model is based on the assessment of nine categories which comprehensively cover all aspects which impact the success of Business Process Management. Some findings of the first assessment round are pinpointed to illustrate the benefits and best practice exchange as a result of the assessment.

Keywords: Business Process Management, Maturity Models, Process Implementation, Reference Modelling

1 INTRODUCTION AND RELATED WORK

Business Process Management (BPM) is a management practice which encompasses all activities of identification, definition, analysis, design, execution, monitoring & measurement, and continuous improvement of business processes. Consequently Business Process Management encompasses not only the analysis and modelling of business processes but also the organizational implementation, leadership and performance controlling (Becker et al. 2003, Schmelzer/ Sesselmann 2008, p. 7f.). Although it is a well-known and largely used practice there is an ongoing discussion on how to best implement Business Process Management. Due to the comprehensive nature of BPM a variety of different approaches exist (e.g. Business Process Reengineering (BPR); Continuous Process Improvement, Workflow Management, reference modelling and implementation of standard enterprise applications).

Facing the importance and vital role of Business Process Management for the transformation and organizational change of enterprises the question arises how different organizations perform in their development of Business Process Management. The notion of maturity has been proposed in other approaches to assess an organizations state in terms of implementing a specific program or the quality of a process.

A prominent and widely used model is the Capability Maturity Model developed by the Software Engineering Institute at Carnegie Mellon University (Paulk et al. 1993). This model was originally developed to assess the maturity of software development processes. Over the years it was extended to other domains. Today the Capability Maturity Model Integration is an approach for the assessment and improvement of product development processes in general. A number of additional maturity models were developed which cover other areas like the CMMI Acquisition Model (CMMI-AM) or the People Capability Maturity Model (P-CMM) for personal management and development to name a
few. Today, CMMI is widely used in practice to evaluate and to improve (software) development processes (CMMI; Ahern et al. 2004, Chrissis et al. 2006, Foegen et al. 2007, Hofman et al. 2007).

CMMI uses standardized question catalogues and evaluation criteria to assess an organizations product development process and to work out the strengths and weaknesses. It helps to define improvement measures and to plan the implementation in an organization. The CMMI introduces the concept of five maturity levels defined by special requirements that are cumulative.

In recent years a number of maturity models for Business Process Management have been proposed (BPMM; Fischer 2004, Lee et al. 2007, Rosemann et al. 2006, 2004, Rosemann/ de Bruin 2005, Smith/ Fingar 2004). Most of the models focus on only one dimension for measuring BPM maturity and very few applied studies are known. Exceptions are the Business Process Management Maturity Model (BPMM) of the OMG and the maturity model of Rosemann et al.

Rosemann et al. identified five factors which are perceived as covering and characterizing BPM (Rosemann et al. 2006, 2004, Rosemann/ de Bruin 2005, Hüffner 2007). In the progress of defining the model these factors have been restructured and renamed by Rosemann et al. and are finally defined as

- Strategic Alignment: Alignment of process management to strategic objectives
- Governance: Organizational implementation of BPM and responsibilities for assigned tasks
- Methods: Methods for all BPM relevant tasks
- Technology: Technologies e.g. I&C which support and enable BPM
- People: Competencies of people involved in BPM
- Culture: Common values towards BPM and process change

At the end of 2007 the Object Management Group (OMG) released the Business Process Management Maturity Model (BPMM). It is a model to assess the maturity of Business Process Management. The model is structured into five process area threads:

- Organizational Process Management: foundation and development of process management
- Organizational Business Management: planning, steering and resource allocation at enterprise level
- Domain Work Management: management of product and service deployment and delivery
- Domain Work Performance: product and service delivery and support
- Organizational Support: all supporting activities for controlling the core activities

BPMM defines objectives for each process area thread. This is supplemented by practices how to reach these objectives. Overall the BPMM offers a variety of recommendations for Business Process Management implementation. On the other hand it leaves some deficiencies in areas like process organization and process accountability. The important role of IT support for Business Process Management is not covered in this model.

This paper presents the implementation of Business Process Management in a large international company, undertaken as a corporate, company wide project within Siemens AG.

The next section outlines the objectives and the overall approach for implementing Business Process Management. It introduces a process framework including a reference process house and the overall structure and content of the BPM implementation process.

Section 3 gives an overview of the Business Process Management maturity model which was developed in order to assess and to derive improvement measures for the Business Process Management in the company. The assessment process and some results of the assessments are presented to illustrate some benefits of the approach.
2 IMPLEMENTATION OF BUSINESS PROCESS MANAGEMENT

2.1 The Business Process Management Initiative

Siemens is engaged in different business sectors with a very broad and diverse product and service spectrum. It is a global company with regional representations in more than 190 countries (for an overview see Feldmayer/ Seidenschwarz 2005, pp. 124 ff.). Over the years the process and IT landscape has developed differently in the respective business units and regions. With the Business Process Management activities a redesign, alignment and optimization of business processes and a better process standardization and utilization of synergies was intended.

A central element of the Business Process Management Initiative was the development of a Siemens Process Framework (SPF 2005) which consists of a reference process house (RPH) and common methods for process management across the company [Feldmayer/ Seidenschwarz 2005, p. 26, Schmelzer/ Sesselmann 2008, p. 241-252]. The initial company wide activities for process standardization started in 2000 with the E-Business initiative “Generic Business Processes”. The primary focus was on the definition of the Supply Chain Management processes based on the Supply Chain Operational Model (SCOR). In the following years the process activities where extended to the Customer Relationship Management and the Product Lifecycle Management. Finally, the activities were taken up and consolidated under the leadership of corporate CIO and the development of a comprehensive reference process house covering all business processes was accomplished (SPF 2005). The primary objective was to leverage synergies and cost potentials with a common organization and process coordination, and the definition of reference processes.

Reference models are increasingly used in industrial practice and leave the area of research (Becker/Delftmann 2007, Fettke and Loos 2007, see the overview in Brocke 2004, p. 393f., for reference modelling projects see RefMod). In practice reference models for processes have particular relevance (e.g. SCOR, Fettke et al. 2006, Scheer 1994). For the development of the Siemens Reference Process House the Supply Chain Operational Model (SCOR) was a fundamental basis.

The Siemens Process Framework (SPF, figure 1), with its binding set of principles and definitions for the overarching management of processes, provides the basis for a uniform implementation of process management within Siemens. The core component of the SPF is the Reference Process House (RPH). It contains the definitions of all Business, Management, and Support Processes down to the agreed level of detail. The Reference Process House is structured into the following process categories:

- Management Processes
- Customer Relationship Management (CRM) Processes
- Supply Chain Management (SCM) Processes
- Product Life Cycle Management (PLM) Processes
- Support Processes

These generic process definitions are fundamental to process standardization and provide a stable basis for process management. They are subject to the cascaded rollout and refinement in the business groups and regions. Incorporating process definitions, guidelines for documentation and modeling of processes, and a binding decision structure for process standardization, the Siemens Process Framework is the basis for:

- Configuration and design of specific business processes (e.g. CRM, PLM, SCM) and end-to-end business process chains
- Redesign of processes based on commonly defined standards for to-be processes
- Common language and common understanding of processes
- Realization of the saving potentials identified through
  - faster implementation of standard processes and alignment of applications
  - utilization of synergy effects
- Comprehensive benchmarking and best practice sharing.
The process management methods of the SPF represent a comprehensive set of tools, concepts, conventions, procedures, and guidelines which are needed for any implementation and operation of process management in the Siemens units. With the description of all roles and responsibilities required for effective process management the SPF provides a blueprint for the organization of process management in the respective business groups and regions. It ensures clear communication and decision processes.

Figure 1. Siemens Process Framework

The main objective of the introduction of Business Process Management is to increase the effectiveness and efficiency of all value creating processes of the organization. From an operational point of view, process management is about having defined processes, measuring their performance, and improving them incrementally as part of daily business. It is also about defining performance goals for processes “top-down”, based on benchmarking results or strategic goals derived from corporate initiatives, and performing major re-engineering activities on processes to close existing performance or cost gaps. Process standards and a common process framework are a fundamental basis for a systematic design and optimization of results, processes, and resources.

Most efficiency and effectiveness problems in an organization have their origin in non-mastered processes. A proper implementation leads to the mastery of processes with regard to lower non-conformance, as well as to high reliability and safety, and results in reduction of process costs, process cycle times, and improvement of quality. Process standardization affects the strategic levers operational excellence and active management of synergies and supports the vertical and horizontal strategies of Siemens. This is achieved by the cascaded definition and rollout approach of the Process Initiative based on the Reference Process House. The implementation of Business Process Management based on the Siemens Process Framework results in a number of benefits which where pursued with the Process Initiative.

- Establish a process management community within the business units and regions to coordinate and optimize local, regional, and headquarter process improvement initiatives.
- Provide a common reference framework for supporting &coordinating all process related projects
- Present a uniform appearance to customers and business partners through Siemens wide standardized process implementation.
- Provide standard service levels to the global customers.
- Enable best practice sharing across all business units and regions.
- Provide opportunity for shared services and an improved lean IT landscape through process standardization.
2.2 Process and implementation topics for Business Process Management

Experience shows that business transformations are often a consequence of good process management. Thus, the implementation of Business Process Management itself has to be organized as a business transformation program covering all relevant aspects of an organization’s development. These aspects have to be addressed by implementation topics which are dependent on each other with regard to their contents. All these issues are addressed by guidelines for BPM implementation (see Process Management Implementation Guide 2005). The following gives a short overview on the different implementation topics.

- **Process Management Organization**: Establish process management roles & bodies according to the Siemens Process Framework and assign the responsible persons.
- **Process Portfolio**: Select, assess, and prioritize the processes which have to be standardized and optimized.
- **Process Documentation & Standardization**: Develop consistent and organization-wide valid process definitions at least for the portfolio processes. Drive the standardization and alignment of business processes and the management and support processes. Establish a process house based on the Reference Process House with organization-wide binding and where necessary more detailed process definitions addressing at least the portfolio processes. Initiate process improvement initiatives for relevant processes of the process portfolio covering: visualization of as-is processes as required, derivation of improvement potentials & measures, design & implementation of to-be processes.
- **Process Performance Controlling**: Define key performance indicators (KPI) and metrics for the portfolio processes derived from business goals and strategies. Introduce a continuous KPI-based performance measurement and assessment for the processes.
- **Process Management Maturity Assessment**: Conduct process management maturity assessments of the organization. Derive & implement improvement measures. Repeat process management maturity the objectives assessments periodically.
- **Methods & Tools**: Provide standard methods and tools required for the operation of process management and according to the Siemens Process Framework guidelines (e.g. a RPH database and ARIS tools).
- **Communication**: Provide target group specific information about objectives, content, roles & responsibilities, and progress of process management to create awareness and support the implementation.
- **Qualification & Training**: Derive competency development measures for the persons involved in process management. Define and conduct target group specific qualification programs. Verify the success.
- **Target Setting & Incentives**: Check and amend target setting and incentive systems. Define process harmonization/standardization and process performance goals. Implement process target agreements, define related incentives.

Only if each of these topics are planned and implemented to a certain degree and in a coordinated way, the effects necessary for overall success are achieved. The overall maturity degree of a process management implementation is therefore directly linked to the maturity degree of each of the implementation topics (see next section). Of course, the business situation, the cultural environment, and the readiness of an organization are additional boundary conditions which have to be considered in the setup of the contents and the timeframe of the implementation program.

The process for process management is structured into the following generic process steps (compare Becker et al., Schmelzer/ Sesselmann 2008): Set Goals, Analyse, Define; Realize, and Review. The process is part of the support process “Process and Information Management” of the Reference Process House. All implementation topics need to be addressed in each of the 5 generic process management steps resulting in the overall structure and content of the BPM implementation process.
Figure 3 comprises the overall structure and the holistic view and comprehensive content of a BPM implementation process. All topics need to be addressed by a BPM maturity assessment.

![Figure 2. Process for Process Management](image)

### 3 A MATURITY MODEL FOR BPM

#### 3.1 Process Management Maturity Assessment (PMMA)

The assessment of the maturity of all activities related to Business Process Management is an essential element of the BPM implementation process. The so-called “Process Management Maturity Assessment (PMMA)” has its focus on the assessment of the organizational implementation of all Business Process Management activities. In contrast most maturity models solely focus on the performance assessment of a specific business process. The process performance of a business process is addressed as a separate category in the implementation process. In this respect the business process performance measurement is one category among others to be addressed in a BPM maturity assessment.

The Process Management Maturity Assessment provides a methodology for a structured analysis and objective assessment of the achieved implementation status of Business Process Management (process management maturity) and the compliance with the Siemens Process Framework (SPF) standards (Feldmayer/ Seidenschwartz 2005, pp. 107 f., Schmelzer/ Sesselmann 2008, pp. 337 f.). The major objective of the PMMA is the identification of need for action and derivation of measures for Business Process Management improvement, as well as the identification of requirements for further support. It serves as a driver for the process initiative. The following objectives are pursued with the PMMA approach:

- to assess the maturity of Business Process Management and the processes
- to monitor the advancement of the process initiative and to derive further fields of actions
- to reveal the potential for best practice sharing
- to motivate and increase the awareness for process management among the involved parties like management, process drivers, and users.

At the time of implementing the Process Initiative no holistic process management maturity model existed which would cover all relevant BPM implementation issues outlined in section 2. The BPMM model of the OGM and the maturity model of Rosemann et al. evolved in parallel to the own development of the Process Management Maturity Assessment (PMMA).

The PMMA follows the principle structure of the Capability Maturity Model Integration Method of the Software Engineering Institute at Carnegie Mellon University (CMMI) but provides a holistic assessment of all areas relevant for BPM based on a comprehensive set of criteria. As an indicator for process maturity, a five step model is applied in the same fashion as the CMMI model.
The PMMA consists of nine categories with one to three sub-categories each. The PMMA categories and sub-categories correspond to the implementation topics of the Process Management Implementation Guide (see section 2):

- Process Portfolio & Target Setting
- Process Documentation
- Process Performance Controlling
- Process Optimization
- Methods & Tools
- Process Management Organization
- Program Management, Qualification, Communication
- Data Management
- IT-Architecture

For every sub-category, each maturity level 1-5 is clearly defined in a to-be status by a set of criteria. These descriptions, as well as examples for questions and possible deliverables, are combined in worksheets. A tool based on MS-Office products was developed to support the assessment process.

Figure 3 outlines the five overall PMMA maturity levels which consolidate the detailed maturity levels of the categories.

<table>
<thead>
<tr>
<th>Maturity Level</th>
<th>Criteria Catalogue</th>
</tr>
</thead>
</table>
| 5 ———> “Optimizing” | - Processes are analyzed, optimized and adjusted to changes in market requirements systematically  
- Benchmarking and Best Practice Sharing are used continuously in order to identify improvement potential  
- Methods for mistake avoidance are used |
| 4 ———> “Quantitatively managed” | - Continuous measurement and adjustment of process performance (quality & quantity)  
- Process management is subject to a systematic maturity assessment (continuous PMMA)  
- Implementation controlling of initiatives with top+ degrees of implementation |
| 3 ———> “Defined” | - The process landscape is derived from systematically ascertained major components of the value chain, business strategy and binding internal/external guidelines.  
- In order to compile a process portfolio, a comprehensible assessment and prioritization of these processes is conducted  
- The systematically ascertained and strategically relevant processes incl. KPIs are documented according to the SFF in the reference process house of the GRDC, a KVP is established  
- Responsibilities for processes are established (roles, committees)  
- Rules and methods of the process management are defined and implemented |
| 2 ———> “Managed” | - Need for action identified/project manager entitled  
- Particular processes in the GRDC are harmonized/standardized  
- Deployment of process management as needed  
- Situation- and/or event-driven approach |
| 1 ———> “Initial” | - Processes are not defined – ad-hoc approach  
- Success depends on certain specialists  
- Schedule, quality and costs are not predictable |

Figure 3. Overall PMMA maturity level

For a sub-category all defined criteria of a maturity level must be met to achieve the respective level. The overall result of a PMMA will be stated in a maturity level grade (e.g. 3,2). The pre-decimal position states that 100% of all sub-categories fulfil the criteria of level 3 (bottleneck is the lowest value for a sub-category). The decimal place states the percentage of fulfilled sub-categories of the successive level (e.g. 20% of level 4). The achievement of a higher level (e.g. 5) in any sub-category is not reflected in the overall grade.

While the maturity levels of figure 3 document the overall assessment and consolidate the maturity assessment of the different categories, a more detailed look on each of the categories is provided by radar screens (see figure 6). Detailed criteria and a set of questions exist to assess the maturity level for each of the categories. Table 1 summarizes what needs to be accomplished for a maturity level 3 in each category.
## PMMA Scope

### PMMA Content of Maturity Level 3

<table>
<thead>
<tr>
<th>Process Portfolio &amp; Target Setting</th>
<th>In order to compile a process portfolio, a comprehensible assessment and prioritization of these processes is conducted.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process Documentation</td>
<td>The systematically ascertained and strategically relevant processes incl. KPIs are documented according to the SPF in the reference process house.</td>
</tr>
<tr>
<td>Process Performance Controlling</td>
<td>A systematic procedure to identify KPIs out of the numerous metrics is defined.</td>
</tr>
<tr>
<td>Process Optimization</td>
<td>Benchmarks are defined and improvement levers identified.</td>
</tr>
<tr>
<td>Methods &amp; Tools</td>
<td>The process landscape is derived from systematically ascertained major components of the value chain, business strategy and binding guidelines.</td>
</tr>
<tr>
<td>Process Management Organization</td>
<td>Responsibilities for processes and process management are established.</td>
</tr>
<tr>
<td>Program Management, Qualification, Communication</td>
<td>The activities for introduction and further development of process management are coordinated systematically by a program and project management.</td>
</tr>
<tr>
<td>Data Management</td>
<td>Harmonization/standardization of data content and formats, clearly defined responsibilities for data definition, content and consistency.</td>
</tr>
<tr>
<td>IT Architecture</td>
<td>Requirements from process management are definitive for IT target architecture. The migration requirement for the IT architecture is derived from deviations between as-is and target architecture.</td>
</tr>
</tbody>
</table>

### Table 1. PMMA categories and maturity level 3 achievements

In general, most CMMI based maturity models define five maturity levels and associate a higher level with a higher maturity and a better performing organization. Crawford (Crawford 2001) argues that this can be a misleading interpretation. An organization should aim for a particular maturity level in relation to its organizational strategies and objectives. A detailed view on the implications of the current maturity level based on the identified shortcomings and weaknesses is proposed in order to derive strategies for improvement.

### 3.2 Maturity Assessment: Initial study and findings

In addition to the workout of the PMMA, a qualification and training program was set up to build a pool of certified assessors who can conduct the PMMA. A roadmap was defined when to assess each organizational unit, eventually covering the entire organization. The PMMA is designed to be repeated once a year to track and drive the improvement.

Figure 5 outlines the execution steps for a PMMA: Between two and three days are required to prepare, conduct, and evaluate the PMMA. The PMMA will be conducted based on interviews with the head of the units, the Process Owners/Process Executives for the Business, Management & Support Processes, and the Process Framework Executive.

The initial assessment analyzed 14 organizational units from some business groups and regions.

The PMMA result can be documented in a radar chart showing the level achievement for each category. Moreover, with the help of PMMA highlights and lowlights for each category, suitable actions can be derived and initiated to improve the implementation status of Business Process Management (process management maturity).
The results for the analyzed units show an overall maturity level ranging below maturity level 3. Although all units participated in the Process Initiative and have implemented Business Process Management the figures show that it is quite some effort in terms of time, resources, and people involved to achieve organizational performance. Also, one has to keep in mind that due to the method of measurement the overall maturity level cannot be higher than the lowest maturity level in any category.

A more detailed view is provided by a radar chart showing the level of achievement for each category. Figure 6 shows the assessment for two selected units providing insights in strengths and shortcomings; e.g. one organizational unit is quite strong in Process Portfolio & Target Setting (level 4) and in Process Management Organization (level 5) and the other in Process Documentation (level 5).
Figure 6. Detailed PMMA of different categories (example for two units)

<table>
<thead>
<tr>
<th>Category</th>
<th>Strength</th>
<th>Weakness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process Portfolio &amp; Target Setting</td>
<td>Specific tools, e.g. scorecards, as basis for deployment from business</td>
<td>No systematic deployment of process portfolio</td>
</tr>
<tr>
<td></td>
<td>strategy</td>
<td>Individual Training necessary</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Objectives are often monetary</td>
</tr>
<tr>
<td>Process Documentation</td>
<td>Process description contains all relevant information (e.g. Input/Output, Interfaces)</td>
<td>Sometimes lacking parts (milestones, metrics or interfaces)</td>
</tr>
<tr>
<td>Process Performance Controlling</td>
<td>Milestones and metrics are defined and used for controlling of most</td>
<td>No integrated measurement system; focusing on process cost drivers to be enhanced</td>
</tr>
<tr>
<td></td>
<td>processes</td>
<td></td>
</tr>
<tr>
<td>Process Optimization</td>
<td>CMMI Assessments in PLM Process Benchmarking with internal and external partners</td>
<td>Organizational obstacles for end-to-end process optimization (interfaces!)</td>
</tr>
<tr>
<td>Methods &amp; Tools</td>
<td>ARIS often in use</td>
<td>Process description not based on RPH</td>
</tr>
<tr>
<td></td>
<td>Several process management methods are used (e.g. six sigma)</td>
<td>or at least level 4 processes not linked to RPH or documented in ARIS.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Level concept/ conventions not used</td>
</tr>
<tr>
<td>Process Management Organization</td>
<td>Process Management Roles are defined; organization is process oriented</td>
<td>Process responsibility not clearly defined; no systematic job rotation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>between roles</td>
</tr>
<tr>
<td>Program Management, Qualification,</td>
<td>Process Management reports directly to BU Head; communication plan</td>
<td>Roadmap for migration to SPF is missing; no qualification plan available</td>
</tr>
<tr>
<td>Communication</td>
<td>regarding process management</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No internal communication</td>
</tr>
<tr>
<td>Data Management</td>
<td>Responsibility for data content and format defined</td>
<td>No mechanism to check data quality or integrity</td>
</tr>
<tr>
<td></td>
<td>Necessary measures are set up</td>
<td>No alignment with process landscape</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Too few resources</td>
</tr>
<tr>
<td>IT-Architecture</td>
<td>Requirements of process management are fully covered</td>
<td>IT architecture not defined, nor communicated – process to derive the</td>
</tr>
<tr>
<td></td>
<td>Migration measures derived</td>
<td>to-be it-architecture not defined</td>
</tr>
</tbody>
</table>

Table 2. Strengths and weaknesses in the BPM categories
The assessment provides a detailed analysis which helps to identify strength and weaknesses and allows to compare the performance of organizations in a differentiated manner and provides a sound basis for best practice sharing. Table 2 summarizes some strengths and weaknesses for the different categories revealed across the assessed organizational units.

Organizations can learn from one another in terms of good and poor performance by understanding the performance of an organization and the underlying reasons. What proofed to be Best Practice can be adopted by other organizations in order to improve performance.

4 SUMMARY AND OUTLOOK

Business Process Management is an important management practice for business transformation and organizational change. This paper outlined the implementation of Business Process Management in a large international company, undertaken as a corporate, company wide project within Siemens AG.

The paper introduced a Process Management Maturity Assessment (PMMA) which was developed to assess the implementation of Business Process Management and the performance of organizations in this respect. The maturity model is based on the assessment of nine categories which comprehensively and entirely cover all aspects which impact the success of Business Process Management.

The proposed Process Management Maturity Assessment advances most of the maturity models which are based on a limited set of criteria. Only the Business Process Maturity Model of the OMG and the maturity model of Rosemann et al. cover also a broader range of BPM factors. Both were in progress of development at the time of PMMA development. All five factors of the Rosemann et al. model can be mapped to the nine categories of the PMMA. A detailed analysis of the underlying criteria and questions for assessment provided they are made public available would show the common ground, possible differences, and additions.

Since the PMMA is based on the principal structure of CMMI using defined maturity levels, structured questionnaires and work sheets, it is easy to use and an assessment for a respective organizational unit can be undertaken in a limited timeframe. A limitation of the CMMI approach is the consolidation of criteria to a single maturity level which may result in misleading interpretations. It is recommended using a detailed view on the assessment and maturity level of each of the nine categories in order to derive a more differentiated picture for improvement measures and best practice exchange, like it was outlined in the example from the business case.

The PMMA was developed to suit the BPM implementation approach which in parts, like the Siemens Process Framework, is company specific. However, the PMMA approach proved to cover all relevant factors for Business Process Management and can be adapted with little effort to a maturity model for general use. This could go in hand with a detailed cross check with the criteria and questions of the maturity model of Rosemann et al. and the Business Process Maturity Model of the OMG.

Overall experiences using PMMA for the assessments are promising in terms of acceptance, ease of use, and coverage of BPM impact factors. The PMMA fits into the overall BPM implementation process in the company and provides an important link to Business Process Management success.

References

BPMM, Business Process Management Maturity Model (BPMM) of OMG, see http://www.omg.org/docs/formal/08-06-01.pdf (called 2009-03-31)


CMMI: Capability Maturity Model Integration (CMMI) of Carnegie Mellon University, see http://www.sei.cmu.edu/cmmi/ (called 2009-03-31)


SCOR. Supply Chain Operations Reference Model, Version 8, see http://www.supply-chain.org/cs/root/home (called 2009-03-31)


Understanding Suppliers’ Participation in Business-to-Government (B2G) Electronic Auction Markets in the Thai Context

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0242.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
UNDERSTANDING SUPPLIERS’ PARTICIPATION IN BUSINESS-TO-GOVERNMENT (B2G) ELECTRONIC AUCTION MARKETS IN THE THAI CONTEXT

Dolpanya, Kitsada, The University of New South Wales, School of Information Systems, Technology and Management, Sydney, New South Wales 2052, Australia, kitsada@student.unsw.edu.au

Land, Lesley Pek Wee, The University of New South Wales, School of Information Systems, Technology and Management, Sydney, New South Wales 2052, Australia, l.land@unsw.edu.au

Dick, Geoff, The University of New South Wales, School of Information Systems, Technology and Management, Sydney, New South Wales 2052, Australia, g.dick@unsw.edu.au

Abstract

Despite business-to-government (B2G) electronic auction (e-auction) markets being a way for suppliers to create opportunities for market expansion and for trading activities, little has been done to understand the behaviour of suppliers participating in these markets. In this paper, we propose a framework to explain suppliers’ intention to participate, and the level of participation in B2G e-auction markets, which will be tested in the Thai B2G e-auction markets. Low supplier participation has been a major problem in the Thai e-auction markets. We posit that suppliers’ participation depends on organisational motivation, environmental uncertainty, and their capabilities. The conceptual framework draws from the Motivation-Ability Framework, Transaction Cost Theory, Institutional Theory, and Resource-Based Theory. It proposes that four key constructs - efficiency motive, legitimacy motive, environmental uncertainty, and organisational capabilities influence suppliers’ intention to participate as well as their participation level in B2G e-auction markets. The conceptual framework is developed by drawing on our understanding of the Thai electronic auction market as well as from extensive literature. We believe this framework may be useful to better understand the key reasons for suppliers to participate in B2G e-auction markets.

Keywords: Electronic auction markets, Business-to-Government (B2G), Suppliers’ participation.
1 INTRODUCTION

Business-to-Government (B2G)\(^1\) electronic markets can be considered as an inter-organisational information system with which participating buyers and sellers utilize electronic markets for a dynamic price-making mechanism (such as electronic auctions), as well as for the exchange of information related to price, product specification, and terms of the trade (Bakos 1991; Grewal & Comer & Mehta 2001). Electronic auction (e-auction) markets are increasingly being used in B2G electronic markets to procure goods and services for governments; they have been reported to yield significant price reductions and time saving for governments as well as to create opportunity for suppliers to penetrate new markets (Beall et al. 2003; Emiliani & Stec 2002; Smeltzer & Carr 2003).

The majority of research on e-auction markets is focused on developed countries, primarily in North America and Europe (Germer & Carter & Kaufmann 2004). There is very little empirical evidence on how B2G e-auction markets perform in the context of developing countries. Developing countries generally lack resources (e.g. skilled people, proper ICT infrastructure) and they also generally report slow economic progress (Jones 2007), compared with developed countries. Molla and Licker (2005) also support this assertion and further suggested that businesses in developing countries are faced with number of challenges (such as technological, financial, and legal infrastructure constraints) in their adoption of e-commerce. These are example barriers to the development and support for B2G e-auction markets. The literature also reports the need for transparency in e-government procurement, especially in developing countries (Rege 2001). Equity is generally promoted for developing countries (UN 2005). B2G e-markets can result in more equity in supplier participation (MacManus 2002) and therefore allow new suppliers to enter the marketplace using a competitive bidding process. For all the above reasons, this study which will be conducted in a developing country is particularly interesting. It will therefore increase our understanding of how B2G e-market will assist the Thai government in promoting the level of suppliers’ participation.

The Thai B2G e-auction markets have a number of characteristics that make them suitable for this study. Firstly, the study of B2G e-auction markets has become significantly important for the procurement of goods and services in South East Asian countries including Thailand (Jones 2007; Settoon & Wyld 2003). The Thai government shows commitment in promoting the B2G e-auction markets by making it mandatory for all Thai government agencies to procure goods and services through e-auction markets, whenever the procurement value is more than 2 million Baht (US$ 60,000). Secondly, the National Statistical Organization (NSO) of Thailand reported that the participation of suppliers in Thai B2G electronic markets is low. Only 0.3% of businesses participate in B2G e-markets compared to 85.3% in business-to-consumer (B2C) and 14.4% in business-to-business (B2B) electronic commerce (NSO 2007). Thirdly, primary researcher has full access to the Thai e-auction markets which makes this study possible.

Prior studies suggest that a sufficient number of qualified suppliers participating in B2G e-auction markets can lead to a competitive market environment (Beall et al. 2003; Elmaghraby 2005; Smeltzer & Carr 2003). Thus, the number of qualified suppliers plays a significant role for the success of e-auction markets. However, the effort to understand the behaviour of suppliers participating in B2G e-auction markets has been lacking. Prior research in this area has two main foci, both in terms of the type of electronic marketplaces (i.e. B2B electronic marketplaces) (Grewal et al. 2001; Son & Benbasat 2007) as well as the research approach (i.e. qualitative case studies). Qualitative case studies provide a rich picture of specific phenomenon within the chosen context (e.g. Emiliani & Stec 2005; Hackney & Jones & Lösch 2007; Soh & Markus & Goh 2006). However, the results do not allow us to generalize to other settings and they also do not allow us to quantitatively validate relationships between key constructs.

---

\(^1\) Business-to-Government is defined as “business activity that involves a business selling its products or services to the central, regional or local government” (source: http://business.govt.nz).
From a thorough review of the relevant literature, this study proposes a research framework with an objective to extend our understanding of the antecedents of suppliers participating in B2G e-auction markets by drawing from the relevant literature including these four theories: the Motivation-Ability Framework, Transaction Costs theory, Institution Theory, and Resource-Based Theory. Studies of B2G e-auction markets have been relatively rare so there is still a lack of a good understanding of how B2G e-auction markets work, especially from the suppliers’ perspective.

This study contributes to the literature in the following ways: Firstly, this study explicitly focuses on the linkage between B2G e-auction markets and supplier participation behaviour, which has not been done before. Secondly, this study conceptualizes supplier participation behaviour within the Motivation-Ability Framework, Transaction Costs Theory, Institutional Theory, and Resource-Based Theory, to extend our understanding of supplier behaviours in the B2G e-auction markets. Lastly, this study aims to fill a gap in the literature arising from a lack of research in the B2G e-auction markets environment.

2 LITERATURE REVIEW

This literature review is taken from diverse disciplines; marketing, economics, organisation management, strategic management, and information systems. We first introduce the electronic auction mechanism, then summarise four key theories important for this study, followed by a description of the Thai e-auction context, before stating the main research questions.

2.1 Electronic Auction

Electronic auction (e-auction) is defined as a market institution with an explicit set of rules determining resource allocation and prices on the basis of electronically submitted bids from market participants (Beall et al. 2003). The term “auction” is used to represent both selling auctions (bidding to buy) and purchasing auctions (offering to sell) (Kaufmann & Carter 2004). The literature in e-auctions is usually discussed in terms of selling auctions, rather than purchasing auction, for simplification (Kaufmann & Carter 2004). The four basic selling auction types were introduced by McAfee and McMillan (1987): (1) the English auction (ascending-bid auction), (2) the Dutch auction (descending-bid auction), (3) the first-price sealed-bid auction, and (4) the second-price sealed-bid (Vickrey) auction. In the same way, Kaufmann and Carter (2004) suggested that the four selling auction types have mirror images in the context of purchasing auction: (1) reverse English auction, (2) the reverse Dutch auction, (3) the first-price sealed-bid purchasing auction, and (4) the second-price sealed-bid purchasing auction. Moreover, the auction types can generally be defined as the following dimensions: the number of bidders, the number of bids per bidder, and the degree of visibility between bidders (Kaufmann & Carter 2004). For this study, we employ only the first-price sealed-bid purchasing auction (with descending prices, multiple permitted bids and almost full visibility) because it has been used in the Thai e-auction. Thus, the terms “e-auction” used hereafter refer to “the first-price sealed-bid purchasing e-auction”, which refers to the electronic competitive bidding between suppliers that drives prices down, or purchasing auctions from buyers.

2.2 B2G Electronic Auction Markets

This paper is confined to the context of a B2G e-auction market as it pertains to a situation with one buyer (government) and a group of sellers (Kaufmann & Carter 2004). In B2G e-auction markets, a government procuring agency invites pre-qualified suppliers who compete against each other to supply a specified good or service, thus driving down the price. Governments generally find the e-auction process attractive because of the tangible benefit of price reductions and the prospect of a reduced transaction cost (Beall et al. 2003; Hackney et al. 2007; Settoon & Wyld 2003). Similarly, suppliers can obtain benefits from opportunities to bid electronically for new business, to penetrate new markets, to create new low costs sales channels, to lower overall transaction costs for buyers in e-auction markets (Smeltzer & Carr 2003), and also to obtain benefit from auction process time.
reduction between bidding and winning the business (Smart & Harrison 2003). Sometimes the auction results are announced at the end of the event, or a day or two later versus weeks or months under traditional auction processes.

2.3 Relevant Theories

A number of relevant theories have been adopted for the theoretical development. These are explained in this section.

The Motivation-Ability framework was introduced by Merton (1957). It has also been applied in the marketing and strategic management literature to study organisation behaviour in terms of organisations’ movement towards online channel. For example, Grewal et al. (2001) employed the motivation-ability framework to investigate organisational participation in business-to-business (B2B) e-markets, they also developed the efficiency construct from transaction cost economics, the legitimacy construct from the institutional theory, as well as the learning and IT capabilities from the Strategic Management literature. Son and Benbasat (2007) extend Grewal et al.’s (2001) work by identifying the salient factors rooted in efficiency motive and legitimacy motive, which affect organisation buyer’s adoption and use of B2B e-marketplaces.

According to Transaction Cost Economics (TCE), all economic activity revolves around a transaction, which is simply some form of exchange of a good or service between two or more economic actors. To optimize the exchange, an appropriate governance structures must be matched to the nature of the transaction (Williamson 1999). Consequently, transactions may be divided into production and coordination costs (Malone & Yates & Benjamin 1987). Coase (1937) proposed that the use of price mechanisms generates cost such as searching for prices, reaching an agreement and enforcing the commitments. In this research, transaction costs represent coordination costs, which consist of operational costs and contractual costs among economic actors in the market. If transaction costs are high, no or little economic activity from suppliers is likely to occur. Bakos (1991) pointed out that information technology would reduce transaction costs, thereby enabling the emergence of more efficiently organised electronic markets.

The Institutional Theory has been used to study organisations. Institutional environments are important for organisational structure and action (Son & Benbasat 2007; Teo & Wei & Benbasat 2003). The key idea behind institutionalization is that organisational action reflects a pattern of doing things that evolves over time and becomes legitimated within organisation and an environment (Eisenhardt 1988). DiMaggio and Powell (1983) suggested three types of isomorphic pressures - mimetic, coercive, and normative – that cause an organisation to have the same form with their environment (e.g. competitors or government/buyer). Mimetic pressures may cause an organisation to imitate the actions of other structurally equivalent, whereas coercive and normative pressures operate through interconnected relations (DiMaggio & Powell 1983).

The Resource Based View (RBV) of the firm suggests that organisations compete and create value on the basis of resources that are unique, rare, valuable, and not easily imitable or substitutable (Barney 1991). Competencies develop when such resources are combined to create specific organisational ability (Day 1994). Peteraf (1993) also proposed a resource-based model of the theoretical conditions which underlie competitive advantage, namely resource heterogeneity, ex post limits to competition, imperfect resource mobility, and ex ante limits to competition. Hall (1993) suggested the sources of sustainable competitive advantage as being two types of capability differential; namely, capabilities based on assets and capabilities based on competencies.

2.4 The Thai B2G E-Auction Markets

The Thai e-auction markets were introduced by the Thai government in 2002. The Thai B2G e-auction markets are highly decentralized. There is no central procuring authority or control agency, there is no purchasing department or the associated purchasing staff. Each of the Thai government agencies can
procure the goods, services through e-auction markets provided by third-party providers of e-auctions. However, the Prime Minister’s Office (PMO) has authority to issue and update regulations that stipulate procurement procedures and standardized contracts in order to enforce all government procuring agencies and public enterprise to deploy transaction through e-auction markets. Given the authority of the different government administration units in mandating the strict electronic procurement practices of the Thai government, it leaves suppliers no choice but to comply with the set rules and regulations if they wish to expand their business to the government sectors.

The Thai government procuring agencies in B2G e-auction markets face a major problem pertaining to too few suppliers participating in these markets - this could result in a non-competitive electronic auction environment (NSO 2007). Smeltzer and Carr (2003) have suggested that at least four or five suppliers are needed to begin the bid process. Whereas, Elmaghraby (2005) argues that more bidders is not always better. These different views notwithstanding, it is important to understand the suppliers’ behaviour to participate in the Thai B2G e-auction markets in order to facilitate these markets’ success and to make these markets more competitive. Thus, the aims of this research are 1) to investigate the factors that influence suppliers’ intention to participate and the level of participation in the Thai B2G e-auction markets and 2) to examine differential effect of the four groups of factors in participation intention and participation level. The main research questions to be addressed are:

1. What types of precursor factors motivate suppliers’ intention to participate, and to increase their level of participation in B2G e-auction markets?
2. Do these key factors play different roles in explaining suppliers’ intention to participate and participation level?

3 FRAMEWORK DEVELOPMENT

In this section, we describe the theoretical development of a framework for explaining the factors that directly influence suppliers’ participation in B2G e-auction markets. We theorise that suppliers’ participation in a B2G e-auction market depends on a supplier’s motivation and their capabilities. The literature on e-commerce also suggested that environmental uncertainty is inherent in e-markets (Lee & Clark 1997). We propose that four main constructs: efficiency motive, legitimacy motive, environmental uncertainty, and supplier capabilities - influence suppliers’ participation (dependent variable) in B2G e-auction markets. Components of the proposed model (Figure 1) are explained below. While we recognise that independent variables may be interdependent to one another, the purpose of the current framework is an initial attempt to find the immediate antecedents of the dependent variable.

3.1 Supplier Participation

In e-auction markets, suppliers’ participation can be classified into two groups; transaction intention and the level of participation (i.e. exploration, trial, commitment, and passive stage). A supplier firm can only be in one stage at any point in time (Grewal et al. 2001).

3.1.1 Transaction Intention

In the technology acceptance model and e-commerce literature, transaction intention is likely to influence future transaction behaviour (Davis 1989; Son & Benbasat 2007; Teo et al. 2003). Behavioural intention refer to the motivational factors that reflect how people are willing to try to undertake a behaviour (Ajzen 1991). In transaction intention period, suppliers will face higher levels of uncertainty related to evaluating the pros and cons of doing business through the B2G electronic auction market (Son & Benbasat 2007).
3.1.2 The level of Participation

To deal with the varying levels of supplier activities in B2G e-auction markets, the participation level can be classified into the exploration stage, the trial stage, the commitment stage, and passive stage (Grewal et al. 2001; Son & Benbasat 2007). In the **exploration stage**, the supplier has been registered in the B2G e-auction market but has not yet begun to conduct trading activities through the e-auction market. In the **trial stage**, the supplier will have conducted several transactions through a B2G e-auction market, but supplier is still evaluating the pros and cons of this means of doing business. In the **commitment stage**, the supplier has made a full commitment because trading through a B2G e-auction market has become an important part of its operations. In the **passive stage**, the supplier has considered not doing business or terminated conducting business in the B2G e-auction market.

Figure 1 shows the proposed research framework for B2G e-auction markets.

---

### Figure 1. Research Framework for B2G E-Auction Markets.

3.2 Efficiency Motive

Organisations participating in e-commerce would be more tended to obtain both efficiency and effectiveness benefits (Bakos 1991). An e-market can reduce coordination costs, which include setting up a relationship, search costs, and transaction costs, between the buyers and the sellers (Bakos 1991). We draw from the Transaction Cost Theory to study the economic organisation of how suppliers seek to minimize transaction costs (Williamson 1981). Arguments for the move to e-markets were based on expected reduction in the transaction costs between buyers and sellers (Bakos 1991; Williamson 1981, 1999). Malone et al. (1987) proposed that information technology, by reducing the transaction costs of market-based coordination, will lead to increased use of market-based governance structures (such as B2G) than hierarchy-based governance structures (such as EDI). An organisation would choose one of these structures that best fits its economic efficiency rationale. Malone et al. (1987) provide two characteristics of products (i.e. asset specificity and product description complexity) which can influence an organisation to select one of governance structures between electronic markets and electronic hierarchies that minimize their total cost. This research also proposes additional variables that could potentially influence organisation decisions to participate in the e-auction markets; namely, market transaction costs, market transparency, and auction process cycle time.

---

Proceedings ECIS 2009
3.2.1 Product Characteristics

Hackney et al. (2007) suggest that not all products are equally suitable for procuring through e-auction markets. Hur et al. (2007) also support this assertion and further suggest that not all products are auction-suitable and the commodities are most suitable for e-auction markets. The type of products directly impact on its specificity (Hackney et al. 2007) and product description complexity (Malone et al. 1987). Malone et al. (1987) proposed two characteristics of products (i.e. asset specificity and product description complexity) that influence suppliers to participate in a B2G electronic auction market. Asset specificity is the extent to which specialised investments are needed to support an exchange (Williamson 1981) or cannot be easily utilized by other firms (Malone et al. 1987). If products in the e-auction market have high asset specificity, suppliers tend not to participate in this market. Product description complexity refers to the amount of information necessary to describe the attributes of a product (Malone et al. 1987; Son & Benbasat 2007). If complex products are difficult to translate into unambiguous product description, suppliers tend not to participate in a B2G e-auction market.

Proposition 1: Product characteristics (high asset specificity and high complexity) in a B2G e-auction market will negatively influence supplier’s intention to participate and the level of participation in the B2G e-auction market.

3.2.2 Market Transaction Costs

E-markets offer facilities to support communicating information about price and production characteristics, and conducting transactions between buyers and sellers (Bakos 1991). E-markets can also help to reduce transaction costs that occur between buyers and suppliers (Bakos 1991). Market transaction costs is defined as the coordination costs involved in using an outside markets, comprising operational costs and contractual costs (Gurbaxani & Whang 1991). Operational costs refer to the costs for accessing market information and process transaction such as search costs and communication costs. Contractual costs refer to the costs of establishing and maintaining contractual relationships with outside parties, including costs of writing contracts and costs of enforcing contracts. We propose that the transaction costs that occur in B2G e-auction markets will negatively influence suppliers’ intention to participate and the level of participation.

Proposition 2: High market transaction costs in a B2G e-auction market will negatively influence supplier’s intention to participate and the level of participation in the B2G e-auction market.

3.2.3 Market Transparency

Auction Market transparency is defined by Madhavan (2000) as “the ability of market participants to observe information about the trading process”. Suppliers can make decisions to participate in trading through e-auction markets based on the available information regarding trading rule, price as well as information on the transaction process (Granados & Gupta & Kaufmann 2006). More transparency in e-auction markets would help suppliers to reduce the information processing costs and opportunity costs of obtaining information about auction price and auction process.

Proposition 3: More information transparency in a B2G e-auction market will positively influence supplier’s intention to participate and the level of participation in the B2G e-auction market.

3.2.4 Auction Process Cycle Time

The use of traditional processes for government buying goods and service can consume several weeks or months (Beall et al. 2003; MacManus 1991). On the other hand, the use of e-auction markets can decrease auction process cycle times (Emiliani & Stec 2005), which are condensed into a period of a few hours (Beall et al. 2003). The cycle time reduction in B2G e-auction markets can benefit suppliers in that suppliers would be better able to plan production scheduling because time is reduced between bidding and winning the business (Beall et al. 2003), and suppliers also save costs in terms of a
reduction in negotiation time (Smeltzer & Carr 2003). We propose that the longer the auction process cycle time, the less willing suppliers are to participate in a B2G e-auction market.

Proposition 4: The longer auction process cycle time in a B2G e-auction market will negatively influence supplier’s intention to participate and the level of participation in the B2G e-auction market.

3.3 Legitimacy Motive

Much of the institutional literature emphasises that organisational structures and processes tend to become isomorphic with the accepted norms for organisations of particular types (DiMaggio & Powell 1983). According to RBV theory, isomorphism is often used as a mechanism for reducing uncertainty by organisations by adopting innovations (DiMaggio & Powell 1983). For example, Son and Benbasat (2007) studied how legitimacy-oriented factors, which are mimetic pressures, coercive pressure, and normative pressures, influence organisational buyers’ adoption and use of B2B e-marketplaces. They found that two isomorphic processes; mimetic and normative pressures have significant effects on adoption intent, but not on participation level. While, coercive pressures did not significantly explain either adoption intent or the level of participation.

3.3.1 Mimetic Pressures

As with Teo et al. (2003), we focus on two specific types of mimetic pressure: participation among competitors and perceived success of participated competitors. Participation among competitors refers to the participation level of competitors participating in B2G e-auction market. Whereas, perceived success of participated competitors refer to suppliers often closely monitoring their competitor to identify successful practices and imitate their actions to achieve similar benefits.

Proposition 5: Mimetic pressures in a B2G e-auction market will positively influence supplier’s intention to participate and the level of participation in the B2G e-auction market.

3.3.2 Coercive Pressures

Coercive pressures is defined by DiMaggio and Powell (1983, p. 150) as “both formal and informal pressures exerted on organisations by other organisations upon which they are dependent and by cultural expectations in the society within which organisation function”. These pressures may take several forms, such as force, threats, persuasion, and invitation (DiMaggio & Powell 1983). For example, the government is one of the largest customers of the supplier, and the supplier’s well being may very much depend on whether it is being awarded the contract from the government. Thus, the purchasing volume from government can dominate a supplier firm’s need to participate in B2G e-auction markets. We propose the effect of the perceived dominance of government procuring agencies on the supplier’s intention to participate and the level of participation in a B2G e-auction market.

Proposition 6: Coercive pressures in a B2G e-auction market will positively influence supplier’s intention to participate and the level of participation in the B2G e-auction market.

3.3.3 Normative Pressures

Normative pressures implies that strategic processes taken by organisations are subject to the values and norms shared among members of their social network (DiMaggio & Powell 1983). Normative pressures from participation in professional and trade associations may promote transactions through a B2G e-auction market. We posit that the effect of supplier participated in professional and trade associations on the supplier’s intention to participate and the level of participation in a B2G e-auction market.

Proposition 7: Normative pressures in a B2G electronic auction market will positively influence supplier’s intention to participate and the level of participation in the B2G e-auction market.
3.4 Environmental Uncertainty

Organisational theories have suggested that organisations must adapt their environment to remain viable in business (Duncan 1972). The literature on the relationship between organisation and environment shows the link between these two variables. For example, Karimi, et al. (2004) show that managerial decision-making tasks are affected by rapid changes that occur in organisational task environments and that when confronted with environmental uncertainty. Lee and Clark (1997) also claimed that environmental uncertainty is inherent in e-markets. The literature has identified many different environment dimensions, three factors are viewed as particularly important (Kabadayi & Eyuboglu & Thomas 2007; Karimi et al. 2004; Newkirk & Lederer 2006) and have been included in a majority of e-commerce studies. They are dynamism, complexity, and hostility. This is also consistent with Duncan’s work (Duncan 1972), which identifies dynamism and complexity as major sources of environmental uncertainty. Dynamism refers to the rate and unpredictability of environmental change. It is especially challenging for suppliers who need to participate in B2G e-auction markets. Researchers have measured dynamism in terms of the frequency of environmental change and the unpredictability of market factors (Homburg & Workman & Krohmer 1999; Kabadayi et al. 2007). Complexity refers to the number and diversity of competitors, suppliers, buyers, and other environmental actors that firm decision makers need to consider in formulating their strategies (Duncan 1972; Kabadayi et al. 2007). Hostility represents the availability of resources and the degree of competition (Newkirk & Lederer 2006) in e-auction markets. Hostility can be measured in terms of the threats to the supplier’s firm posed by labor and material scarcity, intense competition in price, and product differentiation (Karimi et al. 2004; Newkirk & Lederer 2006).

Proposition 8: High environmental uncertainty in a B2G e-auction market will negatively influence supplier’s intention to participate and the level of participation in the B2G e-auction market.

3.5 Supplier Capabilities

This construct is mainly drawn from Resources-Based View Theory (RBV). In the strategic management literature, there is growing evidence that competitive advantage often depends on the firm’s deployment of capabilities (Barney 1991; Day 1994; Wade & Hulland 2004). Thus, firm’s capabilities enable a firm to compete more effectively in the marketplace (Day 1994). Coyne (1986) suggests that not only do the product and/or delivery system attributes need to be important to customers, a capability differential need to be significant for enduring sustainability. Suppliers with greater efficiency can develop sustainable competitive advantage by using this capability to reduce costs and develop a cost leadership position in their industry (Burney 1991; Porter 1985). Hall (1993) suggests that two types of supplier capabilities – capabilities based on assets and capabilities based on competencies - could influence supplier to gain competitive advantage in markets.

3.5.1 Capabilities based on Assets

We propose two sub-constructs that can influence suppliers to participate in B2G e-auction markets. Economics of scale and excess production capacity would be used as sources for suppliers’ competitive advantage in B2G e-auction markets (Elmaghraby 2005). Cost leadership refers to a supplier can gain sales by offering product and/or services at a price that is low than that of competitors as well as pursuing economies of scale in production. Suppliers can produce goods at low cost only if they produce in large quantities. For example, the government provides a large purchasing volume through a B2G e-auction markets which induces a supplier, who has economies of scale (i.e. cost for producing a second unit is less than cost for producing the first unit) to supply goods or services at a small cost than its competitors. Excess supply capacity infers a supplier may differ in its supply capacity. Excess supply capability can be used to supply products and services as supplier’s competitive advantage in order to pursue additional business (Elmaghraby 2005; Smeltzer & Carr 2003). If excess capacity exists in the supply base, a supplier can allocate this valuable resource in an
e-auction market (Jap 2002). For instance, when supplier’s excessive capacity exists, supplier tend to participate in a B2G e-auction market.

**Proposition 9:** Capabilities based on assets of supplier will positively influence supplier’s intention to participate and the level of participation in the B2G e-auction market.

### 3.5.2 Capabilities based on Competencies

Hall (1993) proposed two types of capabilities based on competencies which can be the sources of sustainable competitive advantage, namely, functional capability (i.e. top management’s IT self-efficacy) and cultural capability (i.e. total quality management). In the context of B2G e-auction markets, **top management’s IT self-efficacy** refers to the perceptions of the owner and/or CEO of supplier to use IT in the accomplishment of a task (Bandura 1986; Compeau & Higgins 1995). This definition is based on the concept of self-efficacy defined as “people’s judgments of their capabilities to organise and execute courses of action required to attain designated types of performances. It is concerned not with the skills one has but with judgments of what one can do with whatever skills one possesses” (Bandura 1986, p.391). For example, top management can use his/her ability to manipulate e-auction system provided by third-party providers of e-auctions. Hulland et al. (2007) also found that the organisation which had a strong IT skill capability was positively influenced to commit to the online channel. **Total quality management (TQM)** refers to the continuous improvement of work processes to enhance the organisation’s ability to deliver high-quality product or services in a cost-effective manner (Beer 2003). Supplier’s firms that implement TQM are better positioned to gain through lowered costs and improved customers’ satisfaction (Beer 2003). In addition, Powell (1995) found that TQM can produce economic value; and it can also be used as a potential source of sustainable competitive advantage for suppliers’ firm.

**Proposition 10:** Capabilities based on competencies of supplier will positively influence supplier’s intention to participate and the level of participation in the B2G e-auction market.

### 4 CONCLUSION

In this paper, we have attempted to derive a theoretical framework for explaining supplier behaviours in a B2G e-auction context, by drawing from multiple disciplines. The outcome is derived from extensive and rigorous literature review. It is anticipated that the examination of four key constructs: efficiency motive, legitimacy motive, environmental uncertainty, and supplier capabilities will help to identify reasons for suppliers’ decision to participate in B2G e-auction markets in Thailand. Overall, we believe that this paper extends the understanding of supplier behaviours in the B2G e-auction markets. We also hope that the outcome of this study encourages new thinking and research into the B2G e-auction markets. Future steps include qualitative interviews and focus groups with suppliers in the Thai B2G e-auction markets. These techniques will uncover empirical evidence for possible relationships amongst the independent variables specifically addressing the Thai context and therefore help us refine the theoretical model. Based on the improved model, we will develop survey instruments, followed by pre-test of the instruments, the main survey, and follow-up interviews (if necessary) to explore unexplained results.

### References


FOSTERING EFFICIENCY OF COMPUTATIONAL RESOURCE ALLOCATION - INTEGRATING INFORMATION SERVICES INTO MARKETS

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0507.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>Grid Computing, Agent-based model, Information exchange, Marketplaces</td>
</tr>
</tbody>
</table>
FOSTERING EFFICIENCY OF COMPUTATIONAL RESOURCE ALLOCATION - INTEGRATING INFORMATION SERVICES INTO MARKETS

Borissov, Nikolay, University of Karlsruhe, Englerstrasse 14, 76131 Karlsruhe, Germany, borissov@iism.uni-karlsruhe.de
Brunner, René, Technical University of Catalonia, 08034 Barcelona, Spain, rbrunner@ac.upc.edu
Neumann, Dirk, University of Freiburg, Platz der Alten Synagoge, 79085 Freiburg, Germany, dirk.neumann@is.uni-freiburg.de
Freitag, Felix, Technical University of Catalonia, 08034 Barcelona, Spain, felix@ac.upc.edu
Navarro, Leandro, Technical University of Catalonia, 08034 Barcelona, Spain, leandro@ac.upc.edu
Weinhardt, Christof, University of Karlsruhe, Englerstrasse 14, 76131 Karlsruhe, Germany, weinhardt@iism.uni-karlsruhe.de

Abstract

The application of market mechanisms for the allocation of computing services is a demanding task, which requires bridging economic and associated technical challenges. Even if the market-based approach promises an efficient allocation of computing services, the wide heterogeneity of consumer requirements and the diversity of computational services on provider side are challenging the processes of finding, allocating, and using an appropriate service in an autonomous way. The focus of the most papers is mainly devoted to the optimization embedded in the allocation process itself. However, we think that the optimization process starts much earlier and contains the information gathering until the final market-based resource allocations.

In this paper we introduce an integrated framework for market-based allocation of computing services, integrating information retrieval of market information, prediction models, bidding strategies and marked mechanisms. As proof-of-concept, we implemented a first prototype of the framework. Furthermore, we propose a methodology for evaluating strategic behavior in market mechanisms with bidding strategies using market information and statistical prediction techniques. First simulation results show strategic behavior in selected market mechanisms by applying the proposed techniques.

Keywords: Market Mechanisms, Market Information, Bidding Strategies, Evaluation.
1 INTRODUCTION

Grid and Cloud Computing become not only in science increasingly popular, but also in industry. Prominent examples are Sun Microsystems's Network.com, Amazons Elastic Compute Cloud (EC2) and its Simple Storage Service (S3) and SimpleDB Service. The companies frequently offer a fixed pay-per-use price for static resource configurations. Fixed prices can lead to inefficient utilization, reward and usability, as it does not reflect the dynamics of the market supply and demand. Efficient provisioning and usage of computational resources as well as pricing in vigorous environment like Grid Computing is not manually manageable. Such processes should be automated with no or minimal human interaction. Hence, market mechanism and strategic behavior play an important role for the design of the environment. Self-organization and automatic adaptation to the changing market conditions are key prepositions for efficient offering and consuming of computational resources.

An efficient allocation of consumers' jobs to providers' resources is a complex task, where agent decisions on resource provisioning, market-based allocation and information processing is included. Moreover, the wide heterogeneity of computational resources challenges the process of finding an appropriate resources for given consumer preferences. Since demand and supply of computational resources fluctuates in the course of time, information about current and future resource utilizations and prices are often not known a priori to the participants. In this case consumer and provider agents try to maximize their utilities by generating bids based on available information. This enables strategic behavior both on provider as well as on consumer side.

The knowledge about a market is essential for the design of efficient bidding strategies. Examples are computational approaches incorporating game theory that allow predicting the future through forecasting or learning rules on former or actual trading information. Bergemann's survey (Bergemann and Valimaki 2007) shows that the economic aspect of information acquisition in market mechanisms got more attention by the economic research community. Moreover, the study demonstrates the importance of the economic information disclosure for market participants. The need for this information lies in both being able to apply efficient economic strategies and to feed business models, which are behind these strategies.

The first contribution of this paper is to survey and bring together results for market-based allocation of computational resources. More specifically, to structure the allocation processes in individual steps, we go further in regard to specify mechanisms for information provision and retrieval, price prediction, bid generation and market-based resource allocations. Based on them, we propose an integrated system for market-based computational resource allocation, which organizes the discussed processes. Finally, we specify a common evaluation scenario and present first results.

Following the structure of the paper, section 2 presents promising market mechanisms and bidding strategies. Section 3 shows the role of market information and prediction models for improving allocation efficiency. Thus, section 4 shows an integrated model of a market-based allocation system. Evaluation methodology and first results are depicted in section 5 and related work in section 6.

2 MARKET-BASED RESOURCE ALLOCATION

Auction and strategy selection are closely connected in the sense that a given choice of strategy should affect the choice of auction, and vice versa. For example, some bidding strategies perform well in a Continuous Double Auction, but not in other auctions such as the Dutch auction. This also implies that the choice which auction to participate in depends on the available strategies. Other factors to take into account are the bidding rules – when, how and what to bid, current and average prices, transaction costs, and pricing schemata in the different auctions.

In this section, we discuss potential market mechanisms for allocation of computing services as well as potential bidding strategies enabling the automation of the provisioning and usage processes.
2.1 Market Mechanisms for Computing Services

Economic models for resource scheduling are widely explored in the literature (Wolski et al. 2001, Parkes et al. 2004, Lai et al. 2005, Nassif et al. 2007). According to the allocation modes, scheduling mechanisms can be distinguished into mechanisms, which execute periodically, called also “offline mechanisms”, and mechanisms which execute continuously, called also “online mechanisms”. Mechanisms like Vickrey, English, Dutch, double auctions (Grosu and Das 2006) and combinatorial mechanisms (Bapna et al. 2005) involve a scheduling problem that is complex. The complexity of such mechanisms drives research to look at allocation models, which can be adapted to real-world scenarios and requirements.

Following section presents promising market mechanisms, implemented within the integrated model (section 4) to evaluate market-based allocation of computing resources as well as strategic behavior using the market information from the proposed market information service.

2.1.1 Continuous Double Auction

The CDA has been widely employed in experimental economic studies, where different agent strategies (Gjerstad and Dickhaut 1998, He et al. 2003, Vytelingum et al. 2008) are investigated for applying automated bidding behavior by the provisioning and usage of resources. The matching in CDA is mostly based on a single value – price, which in a computing resource scenario incorporates the consumer’s preferences for a job i.e. his value $v_j$ per time unit. Based on a preferred bidding strategy (see section 2.2) a consumer generates and submits a bid price ($b_j \leq v_j$) to the CDA market. Respectively, based on a preferred bidding strategy, a resource provider also generates an offer price ($o_i \geq v_j$) and submits it to the market. In the provider case $v_j$ represents the reserve price (minimum price) for using the provider’s machine for a unit of time. In case of a match, the consumer executes immediately his job on the provider’s machine with a payment to the provider, calculated by the winning bid and offer prices. The market price of a match is calculated using the pricing schema $k$-pricing (Satterthwaite and Williams 1989): $p_m = kb_j + (1 - k) o_i$, with $k = 0.5$ in our case.

2.1.2 Decentralized Online Machine Scheduling

In the case of the Decentralized Local Greedy Mechanism DLGM (Heydenreich et al. 2006), each time a job $j$ arrives on the consumer side, his bidding agent creates a request in the form $t_j = \{r_j, d_j, v_j\}$ with its release date $r_j$, duration $d_j$ and valuation $v_j$, and reports this to all known providers. The valuation $v_j$ expresses the costs of the job for waiting one additional time unit in the provider machine’s queue. Based on the received bids, the machines perform real-time planning based on a local scheduling policy – if job $j$ has a higher priority value than $k \Leftarrow \frac{v_j}{d_j} \geq \frac{v_k}{d_k}$, then $j$ is scheduled before job $k$ in the waiting queue. Depending on the current local waiting queue, the machine $i$ reports a tentative (ex-ante reported) completion time $\hat{C}$ and payment $\hat{p}$ to the agent of job $j$. The payment $\hat{p}$ contains the aggregated compensations to all job-agents whose jobs are currently waiting at machine $i$ and are delayed due to allocation of $j$.

Upon receiving information about its tentative completion time and required payments, the job-agent makes a binding decision to queue at certain machine $i$, and pays $\hat{p}$ to the delayed jobs. The decision on which machine to submit the job is taken based on the consumer’s utility function $u(j) = -\sum t_j * \hat{C}_{j,i} + \hat{p}_{j,i} , \text{ which selects a provider machine’s offer } i \text{ with the shortest weighted tentative completion time } v_j * \hat{C}_{j,i} \text{ and tentative compensation payments } \hat{p}_{j,i}$. The providers applying the DLGM mechanism do not behave strategically and do not get compensation for the use of their services. The payments are divided only among the consumers. Heydenreich et al. showed that the plain DLGM mechanism achieves a performance ratio of 3.281 against an optimal offline scheduling mechanism.
2.1.3 Fix-Price Technical Scheduling

In order to compare the outcomes of the introduced market mechanisms, we implemented FIFO as baseline scheduling mechanism. The agents submit their jobs to a central scheduler, which schedules and executes them through the FIFO scheduling policy on the various registered provider machines. For each machine the providers request a fixed reservation price.

2.2 Defining Bidding Strategies

Bidding strategies for market-based scheduling are well explored in the economic literature (Reeves et al. 2005, Li and Yahyapour 2006a, Vytelingum et al. 2008). Wellman et al. 2007 give an overview of the various agents and their strategies taken place in the trading agent competition. In general, bidding strategies can be classified into non-adaptive, where the generated bid price do not depend on past and market information e.g. Truth-Telling, and ZI as well as adaptive strategies, where the generated bid price depends on past and market information e.g. ZIP and GD.

Although a simple strategy, truth-telling is essential in case of strategy-proof mechanisms, where in such mechanisms this strategy guarantees to obtain optimal payoffs, no matter what strategies are adopted by the others. A comparison of state-of-the-art bidding strategies is evaluated by Das et al. 2001. The authors of Adaptive-Agressiveness (AA) strategy (Vytelingum et al. 2008) describe and evaluate a novel bidding strategy, which implements short and long-term learning, considering and responding also to dynamic market fluctuations. Simulation results show that the AA outperforms the ZIP and GDX (Tesauro and Bredin 2002) strategies.

Although mainly designed for financial markets, these strategies are fundamental for the design and evaluation of bidding strategies and market mechanisms for Grid and Cloud services (Foster et al. 2008). Following sections present bidding strategies, which we aim to adopt in a Grid market scenario.

2.2.1 Zero-Intelligence Plus Strategy

Zero-Intelligence Plus (ZIP) agents are widely explored and become a popular benchmark for agents trading on continuous double auctions (Das et al. 2001). In Cliff 1997, the author showed that zero intelligence (ZI) agent strategy is not enough, since the bids are uniform generated between a given interval and not depend on current market information or past bids. They introduced ZIP agents, which use public market information to adapt the bid price of a certain market. Experiments (Das et al. 2001) show that ZIP agents perform better than (non-expert) human traders on CDA markets as well as it converges quickly to equilibrium price by high demand and supply.

In most experiments in the literature (Cliff 1997, Tesauro and Das 2001, Das et al. 2001), the ZIP agents receive information of all bids and offers, whether they are accepted or not. There are, however, experiments showing that ZIP based agents perform well also in auctions where the agents only receive information on the winning bids (Bagnall and Toft 2005). In our scenario we will adopt the ZIP strategy in the same way, considering only the winning bids and clearing price of the market.

2.2.2 Gjerstad-Dickhaut Strategy

The Gjerstad-Dickahaut, GD (Gjerstad and Dickhaut 1998, Tesauro and Bredin 2002) bidding strategy is also mainly developed and evaluated in CDA markets. Compared to ZIP, this mechanism is memory based, thus it is using historic market information of the last M bids and offers, H_M, in order to calculate a “belief” function f(p) estimating the probability for a bid or offer to be accepted at price p.

Provider belief function \( f_o(p) = \frac{\sum_{o \in H_M} l + \sum_{b \in H_M} l}{\sum_{o \in H_M} l + \sum_{b \in H_M} l + \sum_{o \in A_H} l} \)
Consumer belief function $f_b(p) = \frac{\sum_{b \in H_M \cap o \geq p} i + \sum_{o \in H_M \cap b \geq p} i}{\sum_{b \in H_M \cap o \geq p} i + \sum_{o \in H_M \cap b \geq p} i}$

For the provider, $\sum_{o \in H_M \cap b \geq p} i$ represents the number of accepted offers in $H_M$ with offer price $o \geq p$, $\sum_{b \in H_M \cap o \geq p} i$ is the number of accepted bids $H_M$ with bid price $o \geq p$ and $\sum_{o \in A_{H_M} \cap b \geq p} i$ is the number of unaccepted offers in $H_M$, $o \notin A_{H_M}$, with offer price $o \leq p$, respectively in opposite for the consumer. Interpolation is applied for prices at which no bids and offers are registered in $H_M$. The bid or offer price is calculated by the product $f(p) \times g$, where $g$ is the calculated gain for a trade at that price, $g = p - v$ for providers and $g = v - p$ for consumers, where $v$ is the valuation.

Although mainly evaluated in CDA markets, GD is applicable as a state-of-the-art strategy in markets for computing resources to the estimation of bids, considering past bids and market information.

2.2.3 Q-Strategy

Q-Strategy (Borissov and Wiström 2008), is a novel bidding strategy adopting a reinforcement learning approach with an e-greedy selection policy. Using the Q-Strategy, an agent explores the environment, exploitation phase, with a probability of $\epsilon$, learning available provider machines and the reward of executing a job on them. With a probability of $1-\epsilon$, exploitation phase the strategy exploits the collected knowledge (i.e. market information) for generating bids more intelligently.

Using Q-Strategy, the agents can bid and adapt also in markets, where market information is incomplete. Simulations show that Q-Strategy adapts to market dynamics (fluctuating prices and changing utilities from providers machines) in different market mechanisms and against bidding strategies like ZIP and Truth-Telling. A common drawback of reinforcement learning algorithms is that learning the optimal bid price needs “training time”. In worst case, Q-Strategy will perform worse at the beginning, but will converge to optimal values in the time (Watkins and Dayan 1992).

3 THE ROLE OF MARKET INFORMATION

Bidding strategies (section 2.2) and prediction techniques (sections 3.1.1 and 3.1.3) require market information in order to efficiently bid in selected markets. However, in a distributed environment like in open and wide Grid markets like in SORMA (http://www.sorma-project.eu/) such information is not locally accessible. (Bergemann and Valimaki 2007) demonstrate the importance of economic information disclosure and show the increased attention, paid to economic information acquisition. Traders require information that enables them to deduce entry prices for available markets and trading times. Centralized markets are able to furnish the current information necessary for simple or sophisticated bidding strategies - such as ZIP agents (Preist and van Tol 1998) or human traders. Therefore, this is an issue for distributed and segmented markets. Distribution and segmentation of markets result in loss of information such as prices, products and effective supply (Brunner et al. 2008). An efficient information system should allow participants to choose a compromise between exact global information and partial information. Furthermore, aggregated, anonymous and summarized information contribute to scalability and may in most cases be sufficient. We conclude the following hypothesis for the influence of information to distributed markets.

Bidding strategies need information about a state of the market, commonly through the offered resource type and price dynamics in time. The main focus here is the treatment of market information by agent’s bidding strategies applying statistical price prediction techniques.

Statistics are used in markets to measure current conditions as well as to forecast financial or economic trends. Indicators are used extensively in technical analysis to predict changes in stock
trends or price patterns. Economic indicators quantifying current economic and industry conditions are used to provide insight into the future demand for commodity goods. In our approach, we use statistical prediction methods for price forecasting.

3.1.1  Trend Extrapolation

The trend extrapolation indicates new trends within the market. In computation markets this is important when newer or higher performance resources entered the market, which will start a deceasing trend for older resources. We will use different kind of moving averages that are indicators in technical analysis of market charts and show the average value of a security's price over a set period. Moving averages are used to emphasize the direction of a trend and to smooth out price and volume fluctuations, or noise. Most bidding strategies implement learning mechanisms, which use statistical prediction models. The ZIP bidding strategy, for example, uses a Widrow-Hoff gradient algorithm, a learning rule (Preist and van Tol 1998), which is a variation of the moving average rule.

In time series analysis, the Moving Average model is common approach to model univariate time series models. A Simple Moving Average (SMA) is the unweighted mean of the previous n data points. For example, a 10-day simple moving average of closing price is the mean of the previous 10 days' closing prices. If those prices are \( P_M, P_{M-1}, \ldots, P_{M-9} \) then the formula is

\[
SMA = \frac{P_M + P_{M-1} + \ldots + P_{M-9}}{10}
\]

In all cases a moving average lags behind the latest data point, simply from the nature of its smoothing. An SMA can lag to an undesirable extent, and can be disproportionately influenced by old data points dropping out of the average. This is addressed by giving extra weight to more recent data points, as in the weighted and exponential moving averages.

A SMA calculates the average of a whole period to predict the future price, which has poor ability to react to new market events. Therefore we will evaluate our approach with a Weighted Moving Average (WMA) to give more weight to recent prices: 

\[
x = \frac{1}{n} \sum_{i=1}^{n} x_i = \left( \frac{1}{n} \right) x_1 + \left( \frac{1}{n} \right) x_2 + \ldots + \left( \frac{1}{n} \right) x_n
\]

where \( \left( \frac{1}{n} \right) \) are the weights, which sum to 1.

An Exponential Moving Average (EMA) is based on the WMA and applies weighting factors which decrease exponentially. The weighting for each older data point decreases exponentially, giving much more importance to recent observations while still considering older observations

\[
EMA = \frac{p_1 + (1-\alpha)p_2 + (1-\alpha)^2p_3 + (1-\alpha)^3p_4 + \ldots}{1 + (1-\alpha) + (1-\alpha)^2 + (1-\alpha)^3 + \ldots}
\]

3.1.2  Cyclic and seasonal component models

Cyclic and seasonal prediction models are applied to detect deviation and trends according to the properties of computational market, which follows periodic type of behaviour. This can be the property of different prices due to the demand on daytime or weekdays. We evaluated workloads (Feitelson 2008) resolving that e.g. in a university network the demand for computational resources is significantly higher during official working hours than during night times or weekends. In our analysis, we will use the autoregressive moving average model (ARMA) and the seasonal autoregressive moving average model (ARIMA).

Autoregressive Moving Average Model ARMA (p) refers to the autoregressive model of order p. The ARMA(p) model is defined as

\[
X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \epsilon_t,
\]

where \( \varphi_1, \ldots, \varphi_p \) are the parameters of the model, \( c \) is a constant and \( \epsilon_t \) is white noise. An autoregressive model is essentially an all-pole infinite impulse response filter with some additional interpretation placed on it. Some constraints are necessary on the values of the parameters of this model in order that the model remains stationary. For example, processes in the AR(1) model with \( |\varphi_1| \geq 1 \) are not stationary.
3.1.3 Grid specific models

As we investigate computational market we will also apply models, which are successful in predicting the demand of computational resources. Sandholm and Lai (Sandholm and Lai 2008) showed that a prediction based on Chebyshev is better suited for predicting performance guarantees, which need worst case scenarios than perfect forecasting.

4 AN INTEGRATED MODEL – INFORMATION AND MARKETS

We propose an integrated architecture for market-based allocation for computational resources. The components have clear separated capabilities and interactions to achieve an economically efficient allocation of applications to needed computing resources. The model is composed by different independent systems (Borissov and Wiström 2008, Brunner et al 2008), which allow a clear separation of code, functionality and expert domain for an easier development, maintenance and fault-detection. The BidGenerator, the Market Information System, the Trading Manager and the Resource Manager are the main components, building an infrastructure for market-based allocation of computing services.

Real applications of the proposed model are listed and described in detail in (Nimis et al. 2008). The application component expresses batch jobs or real (web or desktop) applications, which has to be executed on demanded computing resources. Therefore, the consumer of the application submits a resource request to the BidGenerator. On the other side, the resource manager manages the resources of the provider and is responsible for the execution of the allocated applications. In order to offer a free resource, the resource manager submits a request to the BidGenerator for the resource provision.

![Figure 1. Integrated architecture of the related components.](image)

The BidGenerator (Borissov and Wiström 2008) implements common and new bidding strategies (section ), that autonomously generate and place the consumer bids and provider offers. The bids and offers are generated based on the consumer's preferences and the provider's business model, both initiated in the requests to the BidGenerator. A consumer specifies its preferences in form of a description of the required computing resources, duration of the job/application usage, valuation (maximum price) and preferred bidding strategy. A provider specifies its supply request in form of a resource description, duration of the provision, valuation (reserve price) and preferred bidding
strategy. Then the BidGenerator instantiates the selected bidding strategy and starts the bidding process.

The bids are submitted to the Trading Management (TM) component, which implements and runs market mechanisms (section 2.1) for technical and economic matching. The TM is a platform, which defines the interfaces and rules for implementing market mechanisms and the conversation protocol. In our case, the bids and offers (e.g. simplified example will be a CDA for CPUs with 2GHz), which are generated by the BidGenerator, are submitted to the TM via well defined interfaces of the selected market mechanism. When there is a match, the BidGenerator receives and informs the consumer or provider to execute its application on the allocated resource.

The Market Information Service (MIS) (Brunner et al. 2008) obtains economic data from the Trading Manager and provides it to the BidGenerator. The architecture has been designed to meet both the economic information requirements and that of scalability and robustness of distributed environment. Aggregation mechanisms are used to reach scalability in number of data and agent requests. Many of the introduced bidding strategies (Section 2.2) like ZIP, GD are exploiting prediction techniques (see Section 3), which require public market information. The MIS retrieves and aggregates public market information from the TM to provide them to the BidGenerator.

A short sample of products are commodity goods which are analysed commonly in markets, however, we seek to extend the model to more realistic goods for computing resources. We introduce a resource tuple with core description entities as baseline for representing and aggregating market information for computing resources.

**Definition 1:** We define a resource tuple \( R = \{CPU, M, S, T_s, T_e, D\} \), where \( CPU \) is the processing power, \( M \) is the memory, \( S \) is the storage, \( T_s \) is the earliest start time, \( T_e \) is the latest end time and \( D \) is the duration of the job or resource reservation.

A demanded resource request can technically be allocated if each of the individual requirements is met e.g., minimum \( CPU \) of 2GHz, memory capacity of 1GB, storage of 10GB, etc. The same principle is used for the market information system, a bidding agent can search for the minimum, maximum or average price of \( R \), where the individual requirements are met. This information will be used by the prediction mechanisms described in Section 3.1.

## 5 EVALUATION METHODOLOGY

To evaluate the integrated architecture in section 4, market mechanisms, bidding strategies and prediction techniques, we propose a common agent-based evaluation scenario and present first results.

### 5.1 Agent-Based Simulation

Table 1 summarizes the bidding strategies and mechanisms to be evaluated in varying job workload settings. The trading object is a computing resource as a commodity good or resource bundle. Within the simulation we aim to measure the consumer \( U_c \) and provider \( U_p \) utilities and overall welfare adopting various bidding strategies (see 2.2 and 3) in three market mechanisms (see 2.1). Furthermore we aim to measure the quality \( Q \) of the forecast, deviation of real \( R(t) \) to forecast value \( F(t - 1) \), applying the prediction techniques (section 3) in novel bidding strategies, Prediction Traders.

In order to compare the strategies and market mechanism, we define 6 settings, Table 2, whereas each setting comprised from 20 until 200 providers and consumers. Settings 1 to 3 are generated with a Poisson arrival process, suggested by (Feitelson 2008), by increasing the mean \( \lambda \), generating 1000, 3000 and 5000 jobs. Settings 4 to 5 are real workloads, taken from (Feitelson 2008).
### Table 1. Overall simulation scenario.

The workloads *LPC EGEE* (Medernach 2005), *LLNL Atlas* and *LLNL Thunder* are real cluster usage workloads, each containing 234,889, 42,725 and 121,039 jobs, occupying settings 4, 5 and 6. The LPC log was chosen as a basis because it contains a large variety of jobs with different run-times, numbers of used CPUs, and varying submit and start times. Finally, the valuation of a job and the reserve price of a resource are taken from a normal distribution.

<table>
<thead>
<tr>
<th>Provider machines</th>
<th>10, 20, 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumers</td>
<td>10, 20, 100</td>
</tr>
<tr>
<td>Job-Valuation</td>
<td>Normal distribution</td>
</tr>
<tr>
<td>Reserve price</td>
<td>Normal distribution</td>
</tr>
<tr>
<td>Generated workloads</td>
<td>Job arrival, Poisson process $P(\lambda)$</td>
</tr>
<tr>
<td></td>
<td>Job duration, Normal distribution</td>
</tr>
<tr>
<td></td>
<td>Setting 1 to 3, 1000 to 5000 jobs</td>
</tr>
<tr>
<td>Real workloads</td>
<td>Setting 4 to 6, LPC EGEE, LLNL Atlas, LLNL Thunder</td>
</tr>
</tbody>
</table>

### Table 2. Simulation settings.

#### 5.2 Simulation Results

First results are shown in Table 3. The table shows the evaluation of consumer strategic behavior using Truth-Telling and Q-Strategy (2.2.3) in a centralized CDA (2.1.1), decentralized DLGM (2.1.2) and in a FIFO (2.1.3) mechanisms. The providers are acting strategically only in CDA, in DLGM and FIFO they are offering their resources for free.

<table>
<thead>
<tr>
<th>Setting</th>
<th>Strategy</th>
<th>$DLGM_\mu$</th>
<th>$DLGM_\zeta$</th>
<th>$CDA_\mu$</th>
<th>$CDA_\zeta$</th>
<th>$FIFO_\mu$</th>
<th>$FIFO_\zeta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Truth-Telling</td>
<td>-110</td>
<td>272</td>
<td>-7,92*10^4</td>
<td>95</td>
<td>-140</td>
<td>8</td>
</tr>
<tr>
<td>1</td>
<td>Q-Strategy</td>
<td>-175</td>
<td>258</td>
<td>-10,33*10^4</td>
<td>94</td>
<td>-135</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>Truth-Telling</td>
<td>-213</td>
<td>285</td>
<td>-11,95*10^4</td>
<td>94</td>
<td>-277</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>Q-Strategy</td>
<td>-392</td>
<td>266</td>
<td>-14,63*10^4</td>
<td>93</td>
<td>-265</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>Truth-Telling</td>
<td>-404</td>
<td>286</td>
<td>-7,89*10^4</td>
<td>87</td>
<td>-549</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>Q-Strategy</td>
<td>-901</td>
<td>265</td>
<td>-23,22*10^4</td>
<td>91</td>
<td>-525</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>Truth-Telling</td>
<td>-1104</td>
<td>647</td>
<td>-9,91*10^4</td>
<td>392</td>
<td>-4532</td>
<td>400</td>
</tr>
<tr>
<td>4</td>
<td>Q-Strategy</td>
<td>-1172</td>
<td>581</td>
<td>-11,04*10^4</td>
<td>314</td>
<td>-4444</td>
<td>400</td>
</tr>
</tbody>
</table>

### Table 3. Consumer strategic behavior in three mechanisms.

Each line in the table represents the evaluation of a selected bidding strategy in a selected market for one setting. The workloads – setting 1-3 – are created with a Poisson distribution with means $\mu =$
{0.1, 0.3, 0.5} generating 751, 1502 and 3004 jobs. Setting 4 is the real cluster workload – LPC-EGEE, taken from a Parallel Workload Archive (Feitelson 2008). The first two columns represent the setting (corresponding to those of Table 1) and the evaluated strategy. The next two columns represent the average utility per job \( \mu = \frac{\sum_{i=1}^{n} v_{i}^j C_{i}^j + \pi_{i}^j}{n} \) as well as the standard deviation \( \xi \) of job budget and actual payment in DLGM and the last four columns represent the same for CDA and FIFO.

The results show that Truth-Telling strategy achieves the highest average utility for all settings in both DLGM and CDA, which follows that bidding truthfully in DLGM (2.1.2) and CDA (2.1.1) can only increase the consumer utility. Understating the truthful valuation in lower bid results in a poorer “job priority” by DLGM and thus job can be displaced by other jobs which have higher priority. In case of CDA, the prices depend on the current demand and supply, bidding a lower price instead of the truth valuation increases the risk of no allocation by the mechanism. However, in CDA, the matching is based only on the price without considering the “priority” of a job as with DLGM, and thus achieves very low utility compared to DLGM. Furthermore, CDA-provider machine agents do not maintain a priority queue of the submitted job bids and by an allocation the job is immediately submitted and executed on the provider machine. A provider submits an offer as soon as he becomes idle. Thus each time the agents are competing by adapting their job bids based on the used strategy. In overall, the DLGM mechanism outperforms FIFO in all four settings using the Truth-Telling bidding strategy and achieves the highest common wealth for all participating consumer agents.

\[ \text{Figure 2. ZIP agents trading in a CDA with market information.} \]

Figure 2 contains results from simulation of that market in which agents with a ZIP strategy are competing so sell resources as commodity goods. It depicts that the simulations run stable over a period of 50000 trading rounds with 2000 buyers and sellers.

## 6 RELATED WORK

The influence of information to market-based resource allocation mechanisms expands over different research areas. In Section 3, we explained related and potential market-based resource allocation methods and bidding strategies for these mechanisms. We summarize explicitly the information retrieval process and the individual prediction models in Section 4. This related work section describes work concerning the overall evaluation process of different prediction and learning mechanisms based on historical information.

Sandholm and Lai (Sandholm 2008) apply different mechanisms to predict the future demand of computational resources. They conclude to deduct the price of the resources from different real workload traces; however, markets with real market are not analyzed. The goal is to predict high peaks and though allow the consumers to avoid these to get lower prices or other benefits. This is especially important as sophisticated strategies change the markets behaviors. Moreover, different market mechanisms result in different peaks and distribution of the allocations. (Cardosa and Chandra 2007) analyze statistical aggregation for resource allocation. This information retrieval aggregates
historical data, which builds the basis for the prediction mechanisms. They provide further breakout of
a commodity good like analyzed in most market-allocation mechanisms into resource bundles, however, economic mechanisms are not considered. A market-based scheduling mechanism for the
allocation of time-specific resources is shown by (MacKie-Mason et al. 2004). It evaluates different
prediction mechanisms to govern the time slot in competitive bidding process. They started with the
combination of price prediction and bidding strategies to conclude that price prediction improves
performance in a decentralized market-based scheduling environment.

Another survey and evaluation of market-based prediction mechanism is (Wellman et al. 2002), which
analysis the prediction mechanisms of all participants of a trader competition. The evaluation compares price prediction mechanisms such as historical averaging, machine learning, and a
competitive economy analysis in regard to their relative prediction accuracy. Wellman et al. conclude
that a combination of historical data with current information produce the best prediction in a
competitive analysis. In our work we will analyze and compare processes as entities in competition.

We can conclude from the presented related work that many mechanisms for the statistical prediction
of prices exist for markets. Furthermore, different learning mechanisms and trading strategies are
introduced, which are shown in previous sections. The related work, stated in this section, began with
the comparison and analysis of these techniques. However, to our knowledge an overall approach combining the different areas of efficient information retrieval, statistical prediction, learning rules,
trading strategies, and different market mechanisms is still an open issue.

7 CONCLUSIONS

In this paper, we proposed and analyzed alternative approaches to improve information flow within
markets. Those approaches are measured with respect to the impact endogenous trading information
has on the allocation quality. More precisely, we introduced promising market mechanisms, bidding
strategies and prediction techniques, which have been evaluated and selected regarding their potential
to optimize the resource allocation. Afterwards, we explained how these methods can be technically
combined, integrated and evaluated within the allocation and information gathering process. For an
easier and complete evaluation process, we developed a prototype for the proposed framework and
showed baseline experiments with the proof-of-concept implementation. Positive integration results
were obtained within an implemented market system for computing resource allocation.

As part of our analysis, we found that related work often proposes mechanisms in isolation for each
topic, like the market mechanism, bidding strategy or information treatment, without regarding the
complete interaction of components. Moreover, the focus of these mechanisms is not on the
optimization of computational resource allocation. With the prototype, we show the integration of the
approaches in a real framework, enabling autonomous provisioning and usage of computing resources.
We identified evaluation scenarios and metrics for the impact of information in order to determine
suitable combinations of market mechanisms, bidding strategies and related prediction techniques.

The future focus is on the evaluation of the proposed market mechanisms, bidding strategies and
related prediction techniques from consumer to provider side. Applying the evaluation model with the
proposed metrics will enable to measure the impact of real systems and identify optimal configurations
of market mechanisms and bidding strategies with related prediction techniques to enable efficient
market-based allocation of computing resources.

References

Journal of Autonomous Agents and Multi-Agent Systems
Bergemann, D., Valimaki, J. (2007). Information in mechanism design. 9th World Congress of the


He, M., Leung, H., Jennings, N.(2003). A fuzzy-logic based bidding strategy for autonomous agents in continuous double auctions.IEEETransactions on Knowledge and DataEngineering15(6)1345-1363


AN ECONOMIC ANALYSIS OF SERVICE-ORIENTED INFRASTRUCTURES FOR RISK/RETURN MANAGEMENT

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0284.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Grid Computing, Portfolio management, Economics of IS, Risk</td>
</tr>
</tbody>
</table>
AN ECONOMIC ANALYSIS OF SERVICE-ORIENTED INFRASTRUCTURES FOR RISK/RETURN MANAGEMENT

Buhl, Hans Ulrich, FIM Research Center Finance & Information Management, University of Augsburg, Universitätsstraße 16, 86159 Augsburg, Germany, hans-ulrich.buhl@wiwi.uni-augsburg.de

Fridgen, Gilbert, FIM Research Center Finance & Information Management, University of Augsburg, Universitätstraße 16, 86159 Augsburg, Germany, gilbert.fridgen@wiwi.uni-augsburg.de

Hackenbroch, Wolfgang, Hypo Real Estate Holding AG, Unsöldstraße 2, 80538 Munich, Germany, wolfgang.hackenbroch@hyporealestate.com

Abstract

Risk/return management has not only evolved as one of the key success factors for enterprises especially in the financial services industry, but is in the times of the financial crisis crucial for the survival of a company. It demands powerful and at the same time flexible computational resources making it an almost ideal application for service-oriented computing concepts. An essential characteristic of service-oriented infrastructures is that computational resources can be accessed on demand and paid per use. Taking the estimation of covariances for a portfolio of risky investment objects as an example, we propose quantification for the economic value of fast risk/return management calculations. Our model analyzes the influence factors on the optimal computing capacity dedicated to these calculations and reveals interesting insights in how far the optimal computing capacity depends on market parameters. Our main result is that more volatile markets require a lower computing capacity as the optimal computing capacity depends positively on changes of the market risk but negatively on the risk itself.

Keywords: Risk/Return Management; Service-Oriented Infrastructures; Grid Computing.
1 INTRODUCTION

Risk/return management is crucial for today’s enterprises in order to strive and even to survive in a market environment that can be characterized by tight competition and global integration of markets. This is more than ever emphasized by the problems arising from the current crisis on the financial markets. Especially in the financial services industry, already strict rules and regulations – that nevertheless have not prevented the crisis – are expected to be much more tightened as generally decided by the G20 meeting in April 2009. Consequently, very sophisticated and resource intensive methods for risk/return quantification and aggregation have to be in place. Innovative approaches of distributed computing like grid computing, cluster computing or service-oriented architectures (SOA) are en vogue in academia as well as in practice, offering potentially suitable infrastructures for the corresponding complex calculations. We will speak of service-oriented infrastructures (SOI) in this context as an instance of (the abstract principle of) SOA where (mostly resource-intensive) distributed services are made available transparently over a grid network. Up to now the intellectual treatment of SOI is usually technically oriented and most often neglecting the necessary economic aspects. Even though these aspects are addressed in general by approaches like for instance “grid economics” or “utility computing”, the reflection on benefits and cost still constitutes a widely unresolved issue for specific application domains. We are therefore striving to narrow the gap between the technical capabilities of service-oriented computing and its economical application in risk/return management.

One basic task in risk/return management is the frequent estimation of the risk exposure associated with a portfolio of investment objects (e.g. securities). Today, enterprises mostly calculate their risk exposure during fixed time intervals like e.g. several days. With the possibly huge amount of computing capacity a SOI based on grid technologies offers (embracing resources of the whole enterprise or even of external resource providers), calculations can be accelerated dramatically (Middlemiss 2004). Nevertheless, economic models quantifying the business value of a more frequent recalculation of the risk exposure are not available yet. Thus, the question arises, what is the optimal amount of computing capacity that should be allocated to risk quantification, considering benefits as well as cost? We will deliver a solution to this problem in the form of an optimization model.

Concerning risk/return management, we restrict our considerations to publications addressing risk forecasting and in particular the estimation of covariance matrices. Huther (2003, pp. 111), or Faisst and Buhl (2005, pp. 408) for example describe the use of covariances for a comprehensive enterprise-wide risk/return management. Other publications are dealing with the question of how covariances can be empirically estimated or forecasted by analyzing historical data. In fact forecasting is “one of the important problems in finance” (Elton and Gruber 1972) and consequently there are a lot of publications already covering the question of how many and which historic quotations should be used to determine a suitable risk forecast. To give an impression about applicable techniques we refer to Engle (1982), Kupiec (1995) and Hull and White (1998). Alexander (1996, pp. 233) provides an overview of the corresponding methods used in volatility and correlation forecasting. Recent approaches to volatility forecasting like conditional autoregressive Value-at-Risk (CAVaR) models are presented in Taylor (2005). Although it is widely known that the corresponding calculations are very resource and time intensive, there is to the best of our knowledge so far no publication dealing with the specific problem of quantifying the economic value of a frequent (re)calculation of risk.

Grid computing can be regarded as an infrastructure technology enabling the virtualization of physical resources. Available definitions for the term grid computing are mostly of descriptive nature and provide little more than certain essential characteristics (see e.g. Foster 2002; Foster et al. 2001, 2002;
Foster and Kesselman 1998). However, various proponents have agreed with Foster and Kesselman (1998) that “a computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive and inexpensive access to high-end computational capabilities”. Recently, an evolution towards SOA can be observed (Longworth 2004). Grid technologies are one possibility to realize a SOA consisting of so called “grid services”. Grid services are based on specific web service standards, like the specifications (Open Grid Services Architecture) and (Web Services Resource Framework). They extend web services insofar as they imply the dynamic, yet for the user transparent, allocation of (physical) resources to services by a grid middleware and therefore are especially suited to fulfill resource intensive tasks. There is an extensive literature on service-oriented computing or grid computing in general (see for instance Berman 2005; Foster and Kesselman 1998; Silva 2006; Singh and Huhns 2004). Some publications even consider the application of these technologies for portfolio management, derivatives pricing or other areas of financial risk management (Brownlees et al. 2006; Crespo et al. 2006; Schumacher et al. 2006). However, these approaches describe how grid technologies can be applied, but do not quantify the resulting business value. In the context of grid computing also resource allocation mechanisms have been widely discussed—most often under the term “grid economics”. We refer to Regev and Nisan (1998), Buyya et al. (2000), Nabrzyski et al. (2003) and Wolski et al. (2004) which provide an overview of this area. They most often dwell either on the question which principles are appropriate to resource management or on technical and architectural issues connected with the development of resource management systems. Accordingly, in most of the existing approaches demand for computational resources is merely an external factor whereas in our approach it is subject to optimization.

In this context an important characteristic of SOI based on grid technologies is the on-demand access to distributed resources. When resources stem from an external provider this concept is often labelled utility computing, meaning that resources can be consumed und priced as easy as for instance electricity or water. Utility computing has been subject to research as well. Bhargava and Sundaresan (2004) analyze pay-as-you-go pricing scenarios where providers guarantee computing capacity, but users do not make a commitment towards actual use. Our paper to some extent continues the ideas of Bhargava and Sundaresan (2004). However, we take the perspective of a service user and present a rationale for decisions on computing capacity in the context of risk/return management.

2 RISK/RETURN MANAGEMENT

The term “risk” is used heterogeneously in general speaking as well as in academic circles. Therefore we feel that it is appropriate to begin with a definition of risk before we describe the various requirements and objectives of risk/return management applications. While in the economic literature risk is often generically explained as the “possibility of missing a planned outcome” we will follow a more finance-related approach. From this point of view we define with Schröck (2001, p. 24) risk as “the deviation of a financial value from the expected value”. A positive deviation is often in general speaking referred to as “chance” while a negative deviation is characterized as “danger”. Because of this two-sided perception of risk, variance or standard deviation of a risky value are suitable and well accepted measures of risk. We will use the standard deviation of historical portfolio returns as the risk measure later in this text. Synonymously we will speak of the volatility of a portfolio and define it as the “annualized standard deviation of percentage change in daily price” (Spremann 2003, pp. 154).

Enterprises are investing capital into investment objects in order to generate cash inflows and subsequently to increase the return of the invested capital. Typically risk-averse management is making risky investments hoping to achieve an excess return over the risk-free rate. There is a general connection between risk and return of an investment object: higher return is systematically associated with higher risk. This connection is theoretically explained by economic models like the CAPM (the “Capital Asset Pricing Model” was originally developed by Sharpe (1964), Lintner (1965) and Mossin (1966) and empirically evaluated later on (an overview of relevant empirical studies can be found e.g. in Copeland and Weston (1988, pp. 212)). Following the argumentation of Wilson (1996, pp. 194) it is
therefore crucial for the survival and success of an enterprise to be able to allocate the available capital to the right combination of investment objects, taking into account their specific contributions to the overall risk and return. Investment objects in this context are not restricted to securities. Almost all business transactions are associated with uncertainty and thus contribute to an enterprise’s overall risk exposure. Thus in the spirit of an enterprise-wide risk/return management all investments an enterprise is engaged in, like credit decisions or even customers or projects can be seen as components of the enterprise’s overall investment portfolio, having a return and a variance (as a measure of risk).

One major goal of risk/return management in this context is the prevention of bankruptcy by restricting potential losses resulting from risky investment objects. The increasing importance of this goal is emphasized by the current crisis on the financial markets. A growing number of rules and regulations require enterprises to hold a part of their available capital to back their risky investments (Jackson et al. 1998, pp. 8). This share of the available capital then makes less or no contribution to the overall earnings. By management decisions these restrictions are broken down along the organizational hierarchies into guidelines on business unit or departmental level. We are assuming in the following text that those guidelines are essentially representing limits for the maximum risk a department, business unit and consequently an enterprise is willing (or able) to take.

In order to evaluate whether an enterprise or department complies with a given risk limit, it is necessary to calculate the current risk exposure frequently. For simplifying means, we concentrate on one fundamental instrument in this paper: The covariance approach. This constitutes a basic principle in finance and forms the foundation for many risk/return management applications ranging from Markowitz portfolio optimization to Value-at-Risk (VaR) calculations. In our context covariances are used for determining the overall risk position of an enterprise taking into account diversification effects that exist between the investment objects. Nevertheless, the proposed methods are also applicable to other risk measures as long as they take dependencies between single investment objects into account.

Following the covariance approach, we can represent risky investment objects by random variables. Typically historical data are used in order to derive a distribution for a random variable and calculate the distribution parameters. It is worth mentioning that considering the portfolio risk and return merely on an aggregated level is not satisfactory because all information is lost about the risk attributable to a single investment object. It is crucial to separate the portfolio and decompose it into data per investment object, i.e. to calculate the covariances. Only then the enterprise can perform economically rational investment decisions on different aggregation levels. With \( \sigma^2_i \) and \( \text{Cov}_{ij} \) denoting variance and covariance of investment objects respectively we can determine the overall risk of a portfolio \( \sigma^2_j \), consisting of \( n \gg 0 \) investment objects (numbered from 1 to \( n \)), as

\[
\sigma^2_j = \sum_{i=1}^{n} \sigma^2_i + \sum_{i=1}^{n} \sum_{j>i} \text{Cov}_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{i} \text{Cov}_{ij}
\]

The so defined matrix of all covariances is called covariance matrix. An important characteristic of covariances is that \( \text{Cov}_{ij} = \text{Cov}_{ji} \). This makes the matrix symmetric and thus not all of its values have to be calculated. The total number of covariance calculations necessary is given by \( n(n+1)/2 \).

3 A VALUATION MODEL FOR FAST RISK QUANTIFICATION

We consider an enterprise which has access to (and is possibly engaged in) a set of risky investment objects as well as to a risk-free investment alternative. It is frequently (re)calculating its overall risk position by estimating the covariance matrix of its portfolio. Our main hypothesis for the valuation of

\[A\] A thorough overview and discussion of the statistical analysis of financial data can be found in Shiryaev (1999) pp. 314.

\[B\] As we analyze the complete investment objects and not parts or quantities of investment objects (e.g. stocks), this formula contains no weights.
benefits is: the faster risk/return management calculations can be executed the higher will be the return of the enterprise because given risk limits can better be exploited.

Since the enterprise is acting in an uncertain and dynamic environment its risk position is changing willingly (by investment decisions) or unwillingly (by “movement” of the markets). Because the estimation of risk cannot be accomplished in real-time the covariances at hand are always significantly outdated. We are in the following recurring to the fact that the enterprise is adjusting its risk position to a value somewhere below a certain threshold thus constituting a “safety margin”. In the regulatory context this is often called “haircut”, like in Basel Committee on Banking Supervision (2004). The financial crisis made clear, that – depending on the assets invested in – a high safety margin is necessary to account for changes of parameters, especially in liquidity. It is doing so by using the capital allocation between the risky investment objects and the risk-free alternative for balancing their overall risk position. It is important to understand that our model is not addressing the evaluation of the efficient set of investment objects or portfolio optimization (both require covariances), but the aggregation and management of the risk position of an enterprise. Whenever covariances are available the safety margin can be adjusted immediately in a way that the resulting (and over time changing) overall risk position of the enterprise does not exceed the given risk limit at any time. Hence, the faster covariances are available, the smaller the safety margin can be. We will use this effect to quantify the benefits of fast covariance estimations depending on the time needed for the completion of one covariance matrix.

3.1 Model Setting and Basic Assumptions

The time interval under consideration consists of equidistant periods such that $t=0$ denotes the beginning of the current period. We shall write for example $\sigma_t$ to indicate the value of a model parameter at the end of period $t$. Correspondingly (dis)investment decisions take effect only at the end of each period. If not mentioned otherwise all variables assume real values, i.e. values $\in \mathbb{R}$.

The enterprise is equipped with a total capital of $K>0$ which is always completely allocated to the risky portfolio and/or the risk-free alternative. We denote the risky portion of $K$ with $x \geq 0$ and furthermore use $x$ for risk adjustment. $T$ indicates the length of the calculation time frame, $T \in \mathbb{N} \setminus \{0\}$ with $\mathbb{N}$ as the set of natural numbers. At the end of each time frame we choose $x$ in a way that the risk limit is “probably” not exceeded during the next time frame. We will formulate more precisely what is meant by “probably” later on. Future returns of the portfolio are modelled as independent random variables. Their probability distribution for each period can be characterized by mean and standard deviation. This implies that the investment objects can be marked to market, i.e. there is a price attached to them. We additionally need a set of assumptions for the deductions following thereafter.

**Assumption 1** The enterprise is generally risk-averse and striving for efficient combinations of investment objects. Investment objects are perfectly divisible and traded on a no-frictions market.\(^4\)

**Assumption 2** The risky part of the enterprise’s capital yields the expected return $\mu$, the risk-free investment pays the time-invariant risk-free interest rate $i$, which is equal to the borrowing rate.\(^5\) We always have $\mu>i>0$.

Assumption 2 is made in the spirit of the model of (Modigliani and Miller 1958) where enterprises and investors can borrow or place money at will for a risk-free rate, but expect a premium for taking the risk associated with the investment in risky assets. With the so defined parameters and $x$ as decision

\(^4\)In the sense of the “Portfolio Selection” theory (Markowitz, 1971) investors are trying to achieve the highest possible return on their investment for a given risk. They are acting under perfect trading conditions, i.e. no arbitrage, no transaction cost, strong information efficiency etc.

\(^5\)The equality of lending and borrowing rate is assumed for sake of simplicity and justifies the case $x>1$, where the enterprise is actually borrowing money for making risky investments.
variable we can determine the overall expected return \( \mu^U(x) \) and risk \( \sigma^U_t(x) \) of the enterprise according to common rules of statistics as

\[
\mu^U_t(x) = x\mu + (1-x)i = i + x(\mu - i) \quad \text{and} \quad \sigma^U_t(x) = x\sigma_t.
\]

(1)

The overall risk of the enterprise is expressed by the portfolio risk (the enterprise is the weighted “sum” of its investment objects) and thus changes over time driven by the varying \( \sigma_t \). Note that due to our focus on changing risk we do not regard a changing \( \mu^U_t(x) \) over time (which would result in an index \( t \) as in \( \sigma^U_t(x) \)). As we see, with ascending \( x \) the overall returns as well as the overall risk of the enterprise are both increasing. On the one side the enterprise certainly strives for the highest possible return, on the other side a limitation exists for \( x \) from the given risk limit.

It is common practice to use some variation of a random walk for the price movement on security and commodity markets. This approach goes ultimately back to Louis Bachelier (1900) who compared the stock market with a “drunkards walk”). Although controversially discussed, it was picked up more than half a century later by Mandelbrot (1963, 1972) and Fama (1965) among others. Following this theory of random walks historical (e.g. daily or weekly) portfolio returns can be used for estimating mean \( \mu \) and standard deviation \( \sigma_t \) of future portfolio returns. It is important to understand that in our model the standard deviation is possibly changing in each period (indicated by its index \( t \)).

**Assumption 3** The initial calculation of covariances starts in \( t=0 \) and is finished after \( T \) periods. Each new covariance calculation begins in the finishing period of the previous covariance calculation.

According to assumption 3, whenever a covariance matrix is completed, the input data used for its calculation are \( T \) periods old. We can immediately determine the portfolio risk by summing up the covariances in the matrix. This can then be used for a risk adjustment decision as well as for portfolio optimization. The moment before the next matrix is finished the input data used for risk calculations are already \( 2T \) periods outdated. Therefore the uncertainty interval that has to be taken into account spans \( 2T \) periods: in the worst case the risk has been going up over \( 2T \) periods before the enterprise realizes that it is exceeding the maximum risk it is willing (or able) to take (see figure 1). Without loss of generality we will concentrate our analysis on the first covariance matrix calculation and the corresponding adjustment decision, therefore focussing on the time interval \([0;2T]\). During this time the portfolio risk is fluctuating in a non-predictable way.

![Figure 1. Period Model and Relevant Time Intervals.](image)

### 3.2 The Risk-at-Risk Approach

We will now dwell on the portfolio risk at time \( t \), denoted as \( \sigma_t \), modelling it as a random variable. This relates to a phenomenon known from the behavior of stock market prices called **heteroscedasticity** (see e.g. Spremann 2003, pp. 152). In analogy to the periodical returns we write

**Assumption 4** The \( \sigma_t \) are normally distributed.

This distribution assumption can (and would in practice) be relaxed by approximating the distribution of the \( \sigma_t \) delivered by the calculated sequence of standard deviations. Arranged in increasing order one
can easily deduce the quantiles needed in our model. Nevertheless, for reasons of simplicity and without significantly changing the general result we assume \( \sigma_t \) to be normally distributed here.

We will again focus on two distribution parameters: Our notation for the (strictly positive) mean will be \( \mu \), for the standard deviation \( \sigma \) (both tagged with an \( \sigma \) indicating the fact that the distribution applies to the portfolio risk), thus \( \sigma_t \sim N(\mu_t; \sigma_t) \) with \( N \) short for the normal distribution.

The distribution parameters in our model can again be estimated using historical data. For example, the standard deviation of the portfolio risk can be taken as an estimate for the expected portfolio risk \( \mu_t \) and thus as the starting point for the random walk of \( \sigma_t \). In order to maximize \( x \) in equation (1) under the given constraints we have to consider the uncertainty interval \([0;2T]\). Because of assumption 4 the standard deviation of the expected portfolio risk after \( 2T \) periods is (as a sum of normally distributed random variables) again normally distributed according to \( N(\mu_t; \sigma_t \sqrt{2T}) \). As a consequence for the overall risk of the enterprise we have \( \sigma^x_t(x) \sim N(\mu_t; \sigma_t \sqrt{2T}) \).

In order to rephrase the fuzzy formulation “the risk limit is probably not exceeded” we will follow an approach comparable to the VaR for quantifying portfolio risk. We speak of a Risk-at-Risk over a holding period and a confidence level \( \alpha \), \( 0 < \alpha < 1 \) and think of it as the standard deviation \( \sigma \) which is exceeded within the holding period only with the (small) probability of \( (1-\alpha) \). With \( \Phi(x) \) denoting the standardized normal distribution function, we know for the distribution of \( \sigma^x_t(x) \) over \( 2T \) periods that

\[
P(\sigma^x_t(x) \leq \sigma) = \Phi\left( \frac{\sigma - x\mu_t}{\sigma_t \sqrt{2T}} \right)
\]

At the same time we require in the spirit of the Risk-at-Risk approach the probability given above to be greater than or equal to the confidence level \( \alpha \), i.e.

\[
P(\sigma^x_t(x) \leq \sigma) \geq \alpha \quad \text{for } t = 2T
\]

In the marginal case both sides of the equation are equal and we can therefore state—with \( q_\alpha \) as the (onesided) \( \alpha \)-quantile of the standardized normal distribution—that

\[
\frac{\overline{\sigma} - x\mu_t}{x\sigma_t \sqrt{2T}} = q_\alpha \iff x = \frac{\overline{\sigma}}{q_\alpha \sigma_t \sqrt{2T} + \mu_t} = q_\alpha
\]

\( x \) gives us the portion of risk-free and risky investment objects in a way that equation (2) holds. We can calculate the overall expected earnings of the enterprise, given this capital allocation, as

\[
B(x) = \mu^U(x) \cdot K = (i + x(\mu - i)) \cdot K
\]

Obviously \( B(x) \) represents only expected, calculatory (and not real) earnings because the returns are not fully cash-flow effective and the earnings themselves subject to interpretation. We will neglect the adjectives “expected” and “calculatory” and speak of earnings or, more generally, of benefits.

By inserting \( x \) from equation (3) into \( \mu^U(x) \) from equation (1) we find for the benefits (with the number of periods needed for the completion of one covariance matrix as the independent variable)

\[
B(T) = \left( i + \frac{\sigma(\mu - i)}{q_\alpha \sigma_t \sqrt{2T} - \mu_t} \right) \cdot K
\]

Considering equation (4) an enterprise could maximize its benefit by minimizing the time \( T \) that is needed to calculate a covariance matrix. Yet there is a trade-off between the benefits and the cost, i.e. cost caused by the infrastructure that is necessary to compute the calculations.
4 SERVICE-ORIENTED INFRASTRUCTURES FOR RISK QUANTIFICATION

From a SOI point of view the risk calculation can be regarded as a service providing its user transparently with up-to-date risk information for the relevant investment universe. In this section we derive the relationship between the computing capacity (i.e. cost) allocated to risk quantification and the time needed for the computation.

4.1 Computing Capacity for Risk Quantification

We denote with $z$ the computing capacity necessary for estimating one covariance matrix in exactly $T$ periods. We use CPUs as a measure for computing capacity and are aware of the fact that this means a one-dimensional view on matters as other determinants of system performance are ignored. We denote with $w$ the workload—measured in CPU hours—for estimating one covariance. Applying a simple moving average technique with a rolling sample of historical data (Elton and Gruber 1972, pp. 409) we get unbiased estimators of the expected value and (co)variances for every point in time.

**Assumption 5** The same workload $w$ is necessary for estimating variances and covariances. Furthermore every covariance is estimated from scratch, i.e. no intermediate results are used.

**Assumption 6** The length of the calculation time frame $T$ depends solely on the time needed for the computation, neglecting e.g. latency or transmission times. Correspondingly the only cost relevant is cost for computation which occurs in the form of a (internal or external) factor price per CPU hour over a given time.

Note that for the calculation of covariance matrices on a SOI it is convenient that the computations can be distributed on several nodes and executed in parallel, as all pairwise covariances can be calculated independently from each other. Thus efficiency losses are considerably low.

We can now deduce the computing capacity $z(T)$ (workload per time) that is required in every period over $T$ periods. We already know that for $n$ investment objects $n(n+1)/2$ covariances have to be calculated. Multiplied with the workload per covariance this determines the total number of CPU hours needed. This in turn—divided by the calculation time frame—leads to the functional relationship

$$z(T) = \frac{n(n+1)w}{2T} \Leftrightarrow T(z) = \frac{n(n+1)w}{2z}$$

(5)

Given $n$ investment objects and a computing capacity of $z$ CPUs, the covariance matrix will be completed after $T(z)$ periods. This constitutes an important parameter of the covariance estimation service since it describes the economic value that should be attributed to the consumption of capacity for covariance estimation. By inserting (5) into equation (4) and neglecting the mean risk compared to the standard deviation of the risk we quantify the benefits as

---

$^6$Variance estimation requires only one historical time series, so its intrinsic workload is smaller than the workload for covariance estimation. Nevertheless, this effect can be neglected since for $n$ variances there are $n(n-1)/2$ covariances in a given covariance matrix. With $n$ sufficiently large the variances have merely no effect on the number of calculations (e.g. with $n=201$, the number of variances is only 1% of the number of covariances to be calculated).

$^7$This could be considered awkward for the simple moving average procedure but is a realistic approach for more sophisticated methods.

$^8$Here as well as for the optimal $T(z^*)$ later in this text the outcome is assumed real-valued. In reality and in order to fit it to the discrete-time period model one has to check the neighboring integer values to obtain the discrete optimum.

$^9$The exact quantification would lead to a strictly increasing and concave benefit function $B$, as well. It would nevertheless be tedious to continue in our analysis with the exact expression. In order to avoid writing overhead we deliberately simplify our objective function by neglecting the mean risk, which is a numerically justifiable approximation in our context.
\[ B(z) = a + 2b\sqrt{z} \quad \text{with} \quad a = iK > 0, \quad b = \frac{\sigma^2(\mu - i)K}{2q_0\sigma_\alpha \sqrt{n(a + 1)w}} > 0, \quad B'(z) = \frac{b}{\sqrt{z}} > 0, \quad B''(z) = -\frac{b}{2}\frac{1}{2z^2} < 0. \] (6)

\( B(z) \) is a strictly increasing and concave function.

4.2 Optimizing Computing Capacity on a Service-Oriented Infrastructure

In this paper we abstract from a specific SOI framework or technological implementation, but only consider the essential characteristics of such an infrastructure: A service user can consume exactly the amount of resources needed and is charged by the service provider on a pay-per-use basis. Thereby, it makes no difference whether resources stem from internal sources like enterprise-owned desktop computers and servers that are connected via a grid network or from an external provider. In the case of allocating the (limited) resources of an internal SOI, the enterprise faces opportunity cost that may be taken as a usage price. For external resources a price per unit of computing capacity is set by the provider. The price can be changing depending on the contract and service level agreement used. For example resources could be more expensive when delivered fail-safe during peak times while covering basic load on a lower service level might be cheaper (Bhargava and Sundaresan 2004, p. 203).

In an external provisioning scenario we assume a straightforward cost function using a factor price \( p \) (measured in e.g. $ per workload over a period). Using equation (6) our cost function \( C(z) \) and our objective function \( Z(z) \) then are defined as

\[ C(z) = pz, \quad Z(z) = B(z) - C(z) = a + 2b\sqrt{z} - pz. \]

Note that—as it is generally the case for the parameters in our model—\( C(z) \) describes the cost per period and \( B(z) \) the benefits per period depending on the capacity used per period for the completion of the covariance matrix over \( T \) periods. As the difference of a strictly concave benefit function and a linear cost function, \( Z(z) \) is again a strictly concave function. The first derivative \( Z'(z) = b/\sqrt{z} - p \) features its only null at \( b^2/p^2 > 0 \). Due to the strict concavity of \( Z(z) \) and with respect to equation (6) the only maximum of \( Z(z) \) is at

\[ z^* = \frac{b^2}{p^2} = \frac{\sigma^2(\mu - i)^2K^2}{4q_0^2\sigma_\alpha^2n(n + 1)wp^2}. \] (7)

On a short-term, iterative basis this result can be used to allocate resources of an external provider to risk/return management services. It is possible to examine in detail how input parameters affect \( z^* \). For example the more capital the enterprise has to its disposal the more (in absolute terms) it will invest into risky investment objects. Higher risk exposure in turn increases the importance of risk/return management which is correctly reflected by the positive sign of \( \partial Z/\partial \alpha \). The same argumentation holds when the enterprise faces a higher risk limit \( \sigma \). In this case it should allocate more capacity to risk/return management applications, which is consistently leading to an increasing \( z^* \). Eventually when the risk premium \( (\mu - i) \) rises (due to higher \( \mu \) and/or lower \( i \)) investing into risky objects becomes more attractive, resulting in a larger share of risky capital. In order to manage the consequently more voluminous portfolio our model suggests that additional capacity should be allocated. In the denominator of equation (7) we have the parameter \( w \) determining the CPU hours needed for one covariance estimation. Increasing \( w \) generates higher cost. This in turn leads to less capacity allocation in the optimal case, reflecting the known trade off between accuracy and speed of risk/return management calculations. The behavior of \( z^* \) depending on the confidence level \( \alpha \) is quite surprising. Theoretically \( q_\alpha \) could grow infinitely (with increasing confidence level \( \alpha \)) leading to an infinitesimal small \( z^* \). This is the case because in our model the only way to account for a higher confidence level is a larger safety margin. This causes diminishing benefits and therefore a decreasing \( z^* \) in the optimum. Another interesting and to some extent counter-intuitive result is produced in combination with the parameter \( \sigma_\alpha \). One would possibly expect that with increasing volatility of the portfolio risk the optimal capacity allocation increases which is actually not true. Basically, for a given risk limit \( \sigma \) a higher volatility of the portfolio risk can be leveraged by faster risk calculations (so that the enterprise...
can still get close to the risk limit without hazardously exceeding it during the uncertainty interval). Going back to Bachelier’s proposition the benefits of fast covariance calculations are of order $\frac{1}{\sqrt{T}}$. On the other side, the cost are depending on $T$ in a reciprocal ($1/T$) fashion. As a consequence for higher volatility of the portfolio risk the cost are increasing more quickly than the benefits, leading ultimately to an increasing $T^*$ and decreasing $z^*$, respectively.\(^{10}\)

Figure 2 shows a numerical example of the described optimization. All values are per hour and were in parts estimated from intraday data of the German stock index DAX. The example pictures a company with capital of $K=50$ million $ and $n=200,000$ investment objects. The company tries to hold a risk limit of $\sigma=0.1\%$ with a confidence level $\alpha=99\%$. It can realize a risk free return $i=0.0008\%$ and a risky expected return $\mu=0.001\%$. The workload per covariance is assumed $w=2.5\times10^{-10}$ CPU hours, the price per CPU $p=2\$. It features an optimum at 72 CPUs.

Besides this straightforward linear cost function, especially in an internal provisioning scenario, more complex cost structures may apply. For instance, in a scenario where a high number of services are competing for a limited number of resources, opportunity cost may not increase proportionally, because ever more services with increasing opportunity cost are suppressed. After all, it becomes clear, that a detailed analysis of cost structures along with the knowledge of the concrete utility function of a service not only delivers the opportunity to allocate resources on-demand. The analysis also supports design decisions concerning for instance the questions whether external or internal provision or a mixture of both is preferable, whether or not the service should be deployed doing risk calculations in the background and whether an existing SOI is sufficient or should be enlarged due to an already high utilization rate resulting in high opportunity cost.

5 LIMITATIONS OF THE MODEL AND CONCLUSION

In this paper we demonstrated how the economic value that can be derived from risk/return management calculations can be measured considering an enterprise that has to decide on the amount of capital it wants to reserve to cover potential losses resulting from a risky investment portfolio. Several assumptions (e.g. regarding the distribution of the expected portfolio risk) were necessary to achieve an analytical solution.

Using the covariance approach as an example we moreover developed an optimization model that delivers the optimal amount of computing capacity that should be allocated to risk calculations at a time. In doing so we restricted our analysis to one well-defined risk/return management problem. Although covariances are fundamental and widely used in financial applications we thereby covered only one element of numerous risk/return management methods and algorithms. Other approaches and applications for SOI concepts (like for instance Monte-Carlo simulations which also have a very high parallelization potential) have to be examined as well. In fact, most of the basic principles introduced in this paper can be adapted to other scenarios in more sophisticated and complex surroundings.

Putting it all together a SOI is especially advantageous when market parameters determining the benefits of risk calculations are highly volatile as could be observed during the crisis since July 2007.

\(^{10}\)Such a situation could be observed e.g. during a “regime switch” between two volatility clusters where a time interval with relatively low volatility switches into an interval with higher volatility or vice versa.
resulting in varying demand for computing capacity. With a SOI, resources can be reallocated at any time to reach an economic optimum. As discussed, not only benefits but also (opportunity) cost may vary depending on the total demand for capacity. For example, during “quiet times” risk calculations may be computed more frequently generating added value out of readily available excess capacity even if benefits are comparably small. One caveat not mentioned in this paper is information security. As information on investment objects may be sensitive business data, spreading the calculations over the company or even over service providers may not be desired. Implementing a system as described would therefore require additional security mechanisms and persuasion of the management.

After all, this paper is a contribution to understand the application of service-oriented infrastructures in the specific domain of risk/return management. Although a validation of our findings based on real-world data is still subject to further research, in our point of view, such a systematic and economic analysis is a requirement as a first step for the further development of the new concepts like service-oriented computing or utility computing.

References


Biased Project Status Reports: A Survey of IS Professionals

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0132.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>IT Project Management, Information Systems Development (ISD), IS professionals, Partial Least Squares</td>
</tr>
</tbody>
</table>
BIASED PROJECT STATUS REPORTS: A SURVEY OF IS PROFESSIONALS

Thompson, Ronald, Wake Forest University, Babcock School of Management, 7659 Reynolda Station, Winston Salem NC 27109, USA, ron.thompson@mba.wfu.edu

Iacovou, Charalambos, Wake Forest University, Babcock School of Management, 7659 Reynolda Station, Winston Salem NC 27109, USA, charles.iacovou@mba.wfu.edu

Smith, H. Jeff. Miami University, Farmer School of Business, 311 Upham Hall, Oxford OH 45056, USA, jeff.smith@muohio.edu

Abstract

This paper summarizes an empirical investigation that explored biased project reporting by Information Systems (IS) professionals. The study is based on a survey of 91 professionals who were involved with system implementations in various governmental agencies. Our investigation assessed the impact of project importance, control, structure, and size on biasing behaviors. To formulate the research hypotheses for our study, we adopted a Message Exchange Perspective. The results reveal that IS professionals are more likely to bias their project status communications when working in projects that are (1) large, (2) important, and (3) lack controls. The practical and research implications of our findings are discussed.

Keywords: Information systems development, project reporting, biased reporting.
1 INTRODUCTION

Recent empirical evidence has highlighted an important impediment to the effective management of information system (IS) development initiatives: inaccurate project status reports (Iacovou, 1999, Snow and Keil, 2002b, Snow and Keil, 2002a, Snow et al., 2007, Smith and Keil, 2003, Keil et al., 2004, Keil et al., 2007). Project status reporting refers to the communication of information about the status of a project to interested decision makers, such as project managers, sponsors, and auditors. Typically, in such updates, the reporter compares the state of the project versus expectations in a project plan, and identifies issues that may cause problems in its completion. Without accurate status reports, project executives are unable to direct their projects effectively, allocate resources efficiently and respond to problems in a prompt fashion (Thompson et al., 2007).

The research stream that investigated the presence of inaccurate status reports in IS initiatives identified two sources of misreporting. The first source is reporting error and it refers to unintentional inaccuracies and omissions that the providers of status updates make due to cognitive and information processing limitations. The second source is reporting bias and refers to the deliberate distortion of project status updates by such reporters (Snow and Keil, 2002a). While we recognize the importance of both sources of misreporting, the focus of our investigation is on bias. Specifically, we focus on the set of behaviors that an individual pursues in order to convey an impression of the project status to his/her supervisor that it is different from the one that is actually perceived by him/her.

Biased status reports in IS projects seems to be a significant issue. In a survey of 56 project managers, Snow et al. (2007) found that bias was employed in 60% of the reports. And in surveys of 485 IS project managers and 210 project participants, Thompson et al. (2007) found that status reports are frequently “less than perfect” and that biased reporting is associated with poor project outcomes. Given the occurrence and negative impacts of biased reporting, our investigation aims to identify the types of projects that are prone to such biasing. The research question that frames our work is: how do project characteristics influence the propensity of project participants to bias their status reports? To address this question, we adopt the Message Exchange Theory (MET) to provide a conceptual foundation for our investigation. To the best of our knowledge, this is the first empirical attempt to identify project traits that affect biasing in project reporting.

The rest of the paper summarizes our investigation. In the next section, we review prior organizational communications and IS literature to summarize MET and relevant empirical findings. This is followed by a summary of our research model and the hypotheses that support it. The research methodology that we employed to carry out the investigation is described next. We conclude the paper by discussing the study’s findings, limitations and implications.

2 A MESSAGE EXCHANGE PERSPECTIVE OF BIASED REPORTING

The Message Exchange theory (MET) considers communications as message exchanges within dyads of organizational actors; for an overview of the theory, see Stohl and Redding (1987). A major research stream in MET deals with the study of deliberate misreporting in hierarchical dyads (Read, 1962, Fulk and Mani, 1986, Mellinger, 1958, Jablin, 1979, Roberts and O'Reilly, 1974, Athanassiades, 1973, Chow et al., 2000). Given the consistency between the scope of the MET theory and the focus of our work, we decided to adopt it as the conceptual foundation for our investigation.

The MET proposes that messages can serve various functions in organizational settings. For example, they can be used to provide directions, to establish and nurture relationships, and so on. A predominant purpose of such messages is to serve to attain specific goals (Fulk and Mani, 1986, Stohl and Redding, 1987, Jablin and Sussman, 1983). Thus, while full disclosure may be desirable from the perspective of facilitating project management, the MET posits that such disclosures may be
inconsistent with the needs or goals (such as a desire to further their own careers, to secure resources for their work tasks, etc.) of the message senders. Depending on the situation, reporters may prioritize their needs-goals higher than the need to provide accurate and complete information to others (Read, 1962, Tesser and Rosen, 1975, Pfeffer, 2004, Grover, 2005). Consequently, biased reporting will occur in the transmission of reports in organizational settings (Read, 1962, Grover, 2005, Fulk and Mani, 1986, O'Reilly, 1978, Roberts and O'Reilly, 1974, Keil and Robey, 2001).

A variety of communication techniques can be pursued to bias reports, such as the alteration or withholding of relevant status information (Keil and Robey, 2001, Smith and Keil, 2003, Snow and Keil, 2002b). In our context, alteration refers to the misstating of the condition of project activities and/or the magnitude and causes of problems (or achievements). Withholding refers to the deliberate omission of relevant facts or issues that are related to the state of project activities (Fulk and Mani, 1986, O'Reilly, 1978, Roberts and O'Reilly, 1974, Bavelas et al., 1990, Athanassiades, 1973).

A recent study (Snow et al., 2007) identified several motivations that might lead reporters to offer biased status updates in IS projects. The study found that biased reporting could arise both from self-serving motivations (to further the reporter's own goals) and project-supporting motivations (to increase the likelihood of success for the project). According to the study, reasons cited for biasing included: 1) a fear of communicating bad news (because of a risk that executives might “shoot the messenger”); 2) a desire to make the reporter look good or to avoid looking bad; 3) a belief that project problems could be overcome in the end; and 4) a desire to avoid letting others down. Other reasons for biasing included the reporters’ desire to secure slack resources for their project, concerns about the team’s ability to meet the project’s goals, or a hope to be perceived as a “hero” who turned around a troubled project.

While the MET recognizes that biased reporting may provide perceived benefits for the reporter or the project, it is important to acknowledge that such biased disclosures are risky (for the reporters) and are likely to result in negative consequences if they are detected. Therefore, biasing is more likely to occur when (1) the expected benefits of biased reporting are high, and (2) the risk of detection and expected severity of reprimand are low. This assumption is supported by the findings of prior research, which illustrated that selective reporting is more likely under conditions of information asymmetry (Baiman, 1990, Keil et al., 2000). Consistent with this prior research, we expect that reporters in IS projects will recognize the risky nature of biasing, as the recipients of their reports are likely to view biasing unfavorably (even when well-motivated) as it hinders their ability to monitor and lead the project effectively. Consequently, we expect that project conditions that enhance the perceived benefits of biasing (to reporters or projects) and minimize its risks will increase the likelihood of misreporting in status reports.

3 RESEARCH HYPOTHESES

Based on our review of the IS project management literature, we identified four project factors that are likely to be significant in terms of their influence on biasing in project reporting. These factors are: project importance, structure, control, and size. We decided to focus our investigation on these four factors because (1) they appear to be salient to the arguments of the MET-based perspective and (2) they have been shown to impact both project outcomes and communications (Barki et al., 2001, Henderson and Lee, 1992, McFarlan, 1981, Francalanci, 2001) suggesting that they may be affecting project results through their influence on reporting. By investigating the influence of these factors, we are hoping to identify the types of projects that are prone to biasing. To identify the role that these factors play with respect to biased reporting, we use MET-based arguments to derive the hypotheses that form our research model (see Figure 1). Each hypothesis is described next.
Project Importance: Importance refers to the significance that the organization places on the system under development (Barki et al., 2001). It is determined by its strategic potential and the operational value of the system. Importance is likely to increase the likelihood of biasing because it enhances the potential benefits that can be accrued to reporters who engage in effectual misreporting. When faced with bad news in major undertakings, reporters will be more likely to hide problems in order to avoid the negative political repercussions that may be delivered to those associated with the project or those that deliver negative assessments (Iacovou, 1999). Likewise, the political benefits to those seen as responsible for the success of a significant project are likely to be more consequential (Keil, 1995). In sum, we expect that biasing is likely to yield higher rewards in important projects. Consistent with this assertion, Bean (2001) found empirical support for the relationship between project importance and distortion. Thus, we anticipate that participants in high importance projects will bias their reports more frequently than those in low importance ones (Hypothesis 1).

Project Structure: Structure refers to the level of preciseness and stability in the specifications of the desired project deliverables (McFarlan, 1981, Nidumolu, 1995). Low structure projects experience a lack of specificity in the system requirements and frequent changes to them over the course of the project; high structure projects, on the other hand, have clearer and more stable requirements (Nidumolu, 1996, Lee and Xia, 2002). Biasing is more likely to occur in low structure projects because ambiguity and dynamism make it more difficult for report recipients to detect it, which in turn makes it a less risky behavior. Further, the unclear specifications of a low structure project are likely to intensify the difficulty of project management, making it more difficult for supervisors to monitor the progress of project work (Campbell, 1958). When faced with conditions of unstructured requirements and dynamism, even experienced managers must rely on the specialized expertise of team members to gauge progress and direct the project, which makes biasing less likely to be detected and scrutinized. Thus, we assert that participants in low structure projects will bias their reports more frequently than those in high structure projects (Hypothesis 2).

Project Control: Control refers to “the set of mechanisms designed to motivate individuals to work in such a way that desired objectives are achieved” (Kirsch et al., 2002). The introduction of structured
controls in a project is likely to decrease the incidents of biasing as it will make their potential
detection more likely (and thus will make biasing more risky). Misrepresentation and omission of
project facts from status reports will be more difficult to cover when specific metrics and templates
must be used by participants to communicate the status of their activities (as dictated by control
methodologies). Because of this, we anticipate that project participants would be less likely to engage
in misrepresentations when their work is closely controlled. Pinto and Mantel (1990)’s findings lend
support to this argument. They found that the use of schedules, plans, and other control mechanisms
was associated with better communication and improved trouble-shooting. Research showing that
the use of control tools in IS projects positively impacts their performance (Henderson and Lee, 1992) is
also consistent with this assertion. We therefore propose that participants in projects with lax controls
will bias their reports more frequently than those in projects with stricter controls (Hypothesis 3).

In addition to their direct effects, the above two factors (project structure and control) are likely to
have an interaction effect on biasing. Specifically, we argue that the effect of structure will be
moderated by the impact of control. In other words, for low structure projects, we assert that the use
of controls will be important in reducing biasing; the use of similar controls will not be as important
(for biasing detection) in high structure projects --as the specific nature of the project requirements
will deter the likelihood of biased reports. The interaction effect between control and structure is
supported by prior empirical work (Nidumolu, 1995) which demonstrated that dimensions of these two
constructs affect the residual performance risk in an IS project and that their “fit” affects project
outcomes. Thus, we hypothesize that the effect of project structure on biasing will be moderated by
project control (Hypothesis 4).

**Project Size:** Size refers to the magnitude of the resources that are needed to complete the project.
Past research illustrates that the level of resources is affected by the complexity of the development
initiative (Baccarini, 1996, Holttta-Otto and Magee, 2006, Francalanci, 2001). In general, biasing can
be both more beneficial and less risky in large projects. Using biased reports to avoid delivering bad
news can be more desirable because in projects that consume high levels of organizational resources
the risk of potential scape-goating (i.e., shooting the messenger) can be consequential. Likewise, other
benefits of biasing (such the ability to be seen as a hero who rescued a project) are likely to be more
important in larger, more expensive undertakings. In terms of detection risk, the ability of report
recipients to detect inaccurate status updates in large projects is diminished as the complexity (i.e. the
diversity and volume of work tasks) increases (Holttta-Otto and Magee, 2006). Given that biasing is
likely to yield higher rewards in larger projects and may be more difficult to detect, we anticipate that
participants in large projects will bias their reports more frequently than those in small projects
(Hypothesis 5).

While the focus of our work is on biasing behaviors, we recognize that prior work in IS project
communications (Thompson et al., 2007) has focused on a related, yet distinct construct: quality of
project reporting. **Quality of project reporting** refers to the extent that project updates are complete,
accurate, timely and informative to their recipients. Although biasing behaviors have a degrading
effect on the quality of reporting, they are not the only factors that influence the quality of status
updates that are produced in a project (Muller, 2003). Such other factors include the reporting
experience, communication skills and other traits of the reporter, limitations of the report transmission
media (face-to-face meetings versus email updates, for example), the frequency of reporting, and so
on. While it is not our intent to develop an exhaustive model that attempts to explain all the
antecedents of reporting quality (for a comprehensive review of misreporting antecedents see Smith
and Keil, 2003), we incorporate one final hypothesis in our model to depict the distinction between
biasing behaviors and reporting quality: the frequency of biasing will be negatively associated with
reporting quality in a project (Hypothesis 6)

In the next section, we describe the sample and the methodology that we utilized to empirically assess
the validity of our hypotheses.
4 RESEARCH METHODS

Sample: The sample for this study consisted of people who were team members in state governmental projects in an Eastern state in the U.S.A. The data collection for this study took place as part of a larger effort to study biased project reporting involving interviews and surveys with project personnel, auditors and other involved executives. We selected a diverse subset of projects based on several dimensions (project duration, vendor involvement, etc.) from a listing of all major on-going projects that were being pursued by state agencies. The data that was used to assess our hypotheses were collected using survey questionnaires that were sent to all non-executive IS participants in each of the selected projects. 203 surveys were mailed, and two reminders were emailed to each subject. Ultimately 130 participants responded, yielding a gross response rate of 64%. Of these, 39 cases were removed because of missing data. We embraced an intentionally conservative algorithm in deciding which cases should be removed: a case was removed if the respondent did not complete two or more of the items for any construct. This left 91 usable responses, with a net response rate of 45%.

Scope: The respondents were instructed to answer the questions within the context of a specific IS project: the project on which the participant was currently spending most of his or her time. Moreover, they were instructed to focus on their status reporting to their project manager when responding to questions that were related to biasing and reporting quality. Thus, our data collection focused on the "IS professional to the project manager” reporting dyad.

Measures: Where possible, we used previously developed and validated instruments to measure the constructs of interest. To measure frequency of biasing we used five measures (see Table 1 below): two were adapted from Fulk and Mani (1986), one was taken from Roberts and O’Reilly (1974), and the remaining two were created by the authors.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Scales (seven-point Likert)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please indicate your level of agreement with the following statement:</td>
<td>Strongly disagree... Strongly agree</td>
</tr>
<tr>
<td>&quot;There are significant forces in this project that cause me to distort</td>
<td></td>
</tr>
<tr>
<td>information in some of my communication to the project manager”</td>
<td></td>
</tr>
<tr>
<td>About how often during a typical work week do you withhold information</td>
<td>Virtually never... Very frequently</td>
</tr>
<tr>
<td>from the project manager that might be useful to him/her?</td>
<td></td>
</tr>
<tr>
<td>How frequently do you find it necessary to omit particular project status</td>
<td>Virtually never... Very frequently</td>
</tr>
<tr>
<td>facts from the information you pass on to the project manager?</td>
<td></td>
</tr>
<tr>
<td>Of the total amount of project information you receive, how much of it</td>
<td>Virtually none... A great deal</td>
</tr>
<tr>
<td>must be actively changed in some way before you pass it on to the project</td>
<td></td>
</tr>
<tr>
<td>manager?</td>
<td></td>
</tr>
<tr>
<td>How frequently do you find it necessary to alter the contents of your</td>
<td>Virtually never... Very frequently</td>
</tr>
<tr>
<td>progress reports (either verbal or written) to fit the project manager’s</td>
<td></td>
</tr>
<tr>
<td>expectations?</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Measures of Bias in Project Reporting

To measure reporting quality, we used Mohr and Spekman’s (1994) communication quality scale, which measured communication timeliness, accuracy, adequacy, completeness and credibility. (For space reasons, only the measures for bias are included in this paper, but the others are available from the authors upon request). For project importance, we used the 8-item scale developed by Barki et al. (2001). Following their approach, we summed the responses from the 8 items to create a summated index for importance. For project size, we asked the respondents to report the total budget for the
project. Project structure was modeled as a higher-order construct, with two separate but related dimensions; requirements instability and requirements diversity. Taken from Nidumolu (1995), four items were used to measure requirements instability, and three for requirements diversity. We also modeled project control as a higher-order construct, with internal integration and formal planning as the two dimensions. Four items measuring internal integration and three measuring formal planning were taken from Barki et al. (2001). To assess the impact of biasing on reporting quality in a more accurate fashion, we included a control variable in our instrument (frequency of reporting). To measure this control factor, we created three new items (these were modeled after a frequency of reporting measure from Roberts and O’Reilly, 1974 which we felt lacked face validity within this research context). All new items in our instrument were validated using a card-sorting technique as described in Moore and Benbasat (1991).

For data analysis, we utilized the Partial Least Squares (PLS) technique: specifically, PLS-Graph version 3.0 (Chin and Frye, 2001). Table 2 shows the loadings of multi-item indicators on the constructs they are intended to measure, as well as the cross-loadings on other constructs.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Indicator</th>
<th>RB</th>
<th>RQ</th>
<th>RI</th>
<th>RD</th>
<th>II</th>
<th>FP</th>
<th>FR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting Bias</td>
<td>RB1</td>
<td>.57</td>
<td>-28</td>
<td>.11</td>
<td>.24</td>
<td>-.16</td>
<td>-.08</td>
<td>-.01</td>
</tr>
<tr>
<td></td>
<td>RB2</td>
<td>.84</td>
<td>-.45</td>
<td>.09</td>
<td>.21</td>
<td>-.20</td>
<td>.00</td>
<td>-.18</td>
</tr>
<tr>
<td></td>
<td>RB3</td>
<td>.80</td>
<td>-.38</td>
<td>.04</td>
<td>.10</td>
<td>-.24</td>
<td>-.34</td>
<td>-.33</td>
</tr>
<tr>
<td></td>
<td>RB4</td>
<td>.89</td>
<td>-.42</td>
<td>.04</td>
<td>.12</td>
<td>-.13</td>
<td>-.13</td>
<td>-.16</td>
</tr>
<tr>
<td></td>
<td>RB5</td>
<td>.71</td>
<td>-.24</td>
<td>.10</td>
<td>.12</td>
<td>-.12</td>
<td>-.07</td>
<td>-.15</td>
</tr>
<tr>
<td>Reporting Quality</td>
<td>RQ1</td>
<td>-.48</td>
<td>.81</td>
<td>.05</td>
<td>.02</td>
<td>.24</td>
<td>.24</td>
<td>.28</td>
</tr>
<tr>
<td></td>
<td>RQ2</td>
<td>-.40</td>
<td>.91</td>
<td>-.03</td>
<td>-.10</td>
<td>.13</td>
<td>.13</td>
<td>.29</td>
</tr>
<tr>
<td></td>
<td>RQ3</td>
<td>-.41</td>
<td>.92</td>
<td>-.02</td>
<td>-.06</td>
<td>.08</td>
<td>.22</td>
<td>.25</td>
</tr>
<tr>
<td></td>
<td>RQ4</td>
<td>-.34</td>
<td>.93</td>
<td>.00</td>
<td>-.03</td>
<td>.07</td>
<td>.12</td>
<td>.18</td>
</tr>
<tr>
<td></td>
<td>RQ5</td>
<td>-.41</td>
<td>.84</td>
<td>.01</td>
<td>-.03</td>
<td>.14</td>
<td>.09</td>
<td>.23</td>
</tr>
<tr>
<td>Requirements Instability</td>
<td>RI1</td>
<td>.02</td>
<td>.05</td>
<td>.78</td>
<td>.58</td>
<td>.01</td>
<td>-.04</td>
<td>.13</td>
</tr>
<tr>
<td></td>
<td>RI2</td>
<td>.05</td>
<td>.01</td>
<td>.90</td>
<td>.65</td>
<td>-.09</td>
<td>-.18</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>RI3</td>
<td>.20</td>
<td>-.10</td>
<td>.82</td>
<td>.57</td>
<td>-.16</td>
<td>-.16</td>
<td>-.05</td>
</tr>
<tr>
<td></td>
<td>RI4</td>
<td>.05</td>
<td>.05</td>
<td>.77</td>
<td>.64</td>
<td>-.23</td>
<td>-.05</td>
<td>.03</td>
</tr>
<tr>
<td>Requirements Diversity</td>
<td>RD1</td>
<td>.15</td>
<td>-.04</td>
<td>.70</td>
<td>.86</td>
<td>-.20</td>
<td>-.13</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>RD2</td>
<td>.11</td>
<td>.04</td>
<td>.60</td>
<td>.83</td>
<td>-.05</td>
<td>.03</td>
<td>.18</td>
</tr>
<tr>
<td></td>
<td>RD3</td>
<td>.27</td>
<td>-.11</td>
<td>.57</td>
<td>.84</td>
<td>-.23</td>
<td>-.06</td>
<td>.00</td>
</tr>
<tr>
<td>Internal Integration</td>
<td>HI1</td>
<td>-.13</td>
<td>.06</td>
<td>-.14</td>
<td>-.23</td>
<td>.83</td>
<td>.16</td>
<td>.28</td>
</tr>
<tr>
<td></td>
<td>HI2</td>
<td>-.24</td>
<td>.10</td>
<td>-.12</td>
<td>-.22</td>
<td>.68</td>
<td>.34</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>HI3</td>
<td>-.17</td>
<td>.12</td>
<td>.01</td>
<td>.03</td>
<td>.69</td>
<td>.49</td>
<td>.43</td>
</tr>
<tr>
<td></td>
<td>HI4</td>
<td>-.16</td>
<td>.18</td>
<td>-.20</td>
<td>-.19</td>
<td>.89</td>
<td>.29</td>
<td>.46</td>
</tr>
<tr>
<td>Formal Planning</td>
<td>FP1</td>
<td>-.17</td>
<td>.17</td>
<td>-.29</td>
<td>-.24</td>
<td>.45</td>
<td>.80</td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td>FP2</td>
<td>-.07</td>
<td>.20</td>
<td>.05</td>
<td>.11</td>
<td>.25</td>
<td>.80</td>
<td>.41</td>
</tr>
<tr>
<td></td>
<td>FP3</td>
<td>-.13</td>
<td>.08</td>
<td>-.03</td>
<td>-.01</td>
<td>.29</td>
<td>.84</td>
<td>.27</td>
</tr>
<tr>
<td>Frequency of Reporting</td>
<td>FR1</td>
<td>-.14</td>
<td>.17</td>
<td>.04</td>
<td>.09</td>
<td>.28</td>
<td>.36</td>
<td>.83</td>
</tr>
<tr>
<td></td>
<td>FR2</td>
<td>-.14</td>
<td>.17</td>
<td>-.03</td>
<td>.07</td>
<td>.27</td>
<td>.33</td>
<td>.75</td>
</tr>
<tr>
<td></td>
<td>FR3</td>
<td>-.24</td>
<td>.24</td>
<td>.02</td>
<td>.09</td>
<td>.35</td>
<td>.29</td>
<td>.88</td>
</tr>
</tbody>
</table>

Table 2: Indicator Loadings and Cross-Loadings

For adequate item reliability, ideally the item loadings should be higher than .70, although slightly lower loadings for individual items are usually acceptable providing other items measuring the construct are greater than .70. The only item that appears to be potentially problematic is the first measure of reporting bias, RB1, which displayed a loading of .57. After reviewing this item from a face validity basis, we decided to retain it in an effort to capture as much of the meaning of the reporting bias construct as possible. (As an aside, we tried re-running the model without this item and...
observed no significant differences in the results obtained).

To assess scale reliability and internal consistency, we computed the Composite Reliability (CR) score (similar to Cronbach’s alpha) and the average variance extracted (AVE). For adequate reliability, the CR score should be greater than .70 (or .80, for more mature streams of research). All of our CR scores exceeded .80. In addition, the AVE scores ranged from .59 to .78, comfortably exceeding the recommended level of .50. The CR and AVE scores are shown in Table 3.

For discriminant validity, two tests were performed. First, the cross-loadings of the items were examined, to ensure that: 1) each item loads more highly on its own construct than on any other construct, and 2) there are no items that load more highly on a construct than the items intended to measure that construct. All measures passed this test. Second, we compared the square root of the AVE for each construct to ensure it was greater than the correlations between that construct and all other constructs (see Table 3). All measurement scales passed this test as well.

<table>
<thead>
<tr>
<th>Construct</th>
<th>CR</th>
<th>AVE</th>
<th>RB</th>
<th>RQ</th>
<th>RI</th>
<th>RD</th>
<th>II</th>
<th>FP</th>
</tr>
</thead>
<tbody>
<tr>
<td>RB</td>
<td>.88</td>
<td>.59</td>
<td>.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RQ</td>
<td>.95</td>
<td>.78</td>
<td>-.47</td>
<td>.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RI</td>
<td>.89</td>
<td>.67</td>
<td>.10</td>
<td>.00</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>.88</td>
<td>.71</td>
<td>.21</td>
<td>-.05</td>
<td>.74</td>
<td>.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>.86</td>
<td>.61</td>
<td>-.22</td>
<td>.15</td>
<td>-.14</td>
<td>-.19</td>
<td>.78</td>
<td></td>
</tr>
<tr>
<td>FP</td>
<td>.85</td>
<td>.66</td>
<td>-.16</td>
<td>.19</td>
<td>-.13</td>
<td>-.07</td>
<td>.41</td>
<td>.81</td>
</tr>
</tbody>
</table>

Note: CR – composite reliability score; AVE – average variance extracted; RB – report biasing; RQ – report quality; RI – requirements instability; RD – requirements diversity; II – internal integration; FP – formal planning. For adequate discriminant validity, the diagonal elements (i.e. square root of AVE) should be larger than the off-diagonal ones.

Table 3: Tests of Reliability and Discriminant Validity.

Satisfied that the measures were adequate, we then examined the structural model results. These are shown in Figure 2. Note that we have expanded this model from the earlier research model, to show the higher-order constructs (project uncertainty and project control) with their separate dimensions, as well as the interaction term (Uncertainty × Control) that was added to test for a moderating influence of control on uncertainty (see hypothesis 4). Finally, in addition to the control variable of reporting frequency, we also added the influence of project size on importance as a control (as those two factors are likely to be related). Table 4 summarizes the results with respect to our tests of hypotheses.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Relationship</th>
<th>Path Estimate</th>
<th>t-statistic</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Project Importance → Biasing</td>
<td>.31</td>
<td>3.18</td>
<td>Yes</td>
</tr>
<tr>
<td>H2</td>
<td>Project Structure → Biasing</td>
<td>.07</td>
<td>0.80</td>
<td>No</td>
</tr>
<tr>
<td>H3</td>
<td>Project Control → Biasing</td>
<td>-.35</td>
<td>2.80</td>
<td>Yes</td>
</tr>
<tr>
<td>H4</td>
<td>Control moderates Structure → Biasing</td>
<td>.09</td>
<td>0.77</td>
<td>No</td>
</tr>
<tr>
<td>H5</td>
<td>Project Size → Biasing</td>
<td>.22</td>
<td>2.52</td>
<td>Yes</td>
</tr>
<tr>
<td>H6</td>
<td>Report Biasing → Reporting Quality</td>
<td>-.43</td>
<td>3.56</td>
<td>Yes</td>
</tr>
<tr>
<td>Control</td>
<td>Reporting Frequency → Quality</td>
<td>.19</td>
<td>1.88</td>
<td>p &lt; .10</td>
</tr>
<tr>
<td></td>
<td>Project Size → Importance</td>
<td>.27</td>
<td>2.95</td>
<td>p &lt; .05</td>
</tr>
</tbody>
</table>

Table 4: Results of Tests of Hypotheses
DISCUSSION OF FINDINGS

We note from Figure 2 that the amount of variance explained ($R^2$) in biasing was .24 indicating that a substantive portion of the variation in biasing behaviors can be explained by the project factors that were identified as antecedents in our model. Moreover, the variance explained for reporting quality was .25, and the path estimate from biasing to reporting quality was -.43, suggesting that the frequency of report biasing does have a substantive, negative impact on report quality. This confirms that understanding the factors that influence report biasing is indeed important.

With respect to the effects of the antecedents, three of the four factors hypothesized to influence biasing (project importance, control and size), did in fact do so. As hypothesized, it appears that participants in more important projects are more likely to engage in biasing behaviors. Similarly, in larger projects the participants are more likely to bias their reports. In contrast, the more perceived control over the project, the less frequently biasing will occur.

As our results indicate, the degree of project structure did not have an influence on the frequency of report biasing. We had hypothesized that the more uncertainty that existed in project requirements (both in terms of the diversity and variability of requirements), the more that biasing would occur. Our rationale was that low structure projects would have more ambiguity and dynamism, making biased reporting more difficult to detect, and hence participants would see it as a less risky behavior and tend to bias more. Our results did not support this hypothesized relationship. One possible explanation is that the degree of structure was average, not low, in the projects that were included in this sample. An examination of the descriptive statistics for the seven indicators measuring requirements diversity and instability revealed that the means ranged from 3.5 to 4.5 (on a 7-point scale), suggesting that this might have had a dampening effect. One reason for the presence of a higher level of structure (across all projects in our sample) is that the state that was the focus of our data collection effort required all of its major projects to undergo a certification review before receiving funding. Moreover, the state required periodic reviews for such projects. Thus, it is possible...
that these reviews compelled agencies to structure their projects (and reduce variability in requirements) in a way that is not typical in other environments. We also recognize the limitations of using interval scales when testing for interaction effects (hypothesis 4). We tried running the PLS model without the interaction effect, and found adding it only increased the variance explained by a nominal amount, and hence we conclude there was no evidence of an interaction.

6 LIMITATIONS AND IMPLICATIONS

Before discussing the practical and research implication of our findings, there are several limitations with this study which need to be acknowledged. First, our data was cross-sectional in nature, which does not allow us to truly test for causality. While we can propose directionality in relationships based on previous research, our data and statistical analyses can only detect correlations and not causation. Nevertheless, it is important to recognize that the associations that we have identified among the constructs of interest are consistent with the causal relations in our chosen theory. Second, our sample was taken from participants involved with projects at state governmental agencies. As such, we cannot generalize the results outside of that context without caution. Third, while we took several steps to enhance and assess the validity of our measures, a couple of measurement issues exist. Specifically, project size was measured with a single indicator and new items were used in measuring three of our constructs.

Despite the above limitations, we believe the results from this study provide interesting implications for both practitioners and researchers. For practitioners, our work has confirmed empirically that biased reporting exists in IS projects and that such biasing is likely to be affected by three dimensions of the project itself: its size, importance and use of controls. Our findings indicate that risk of biased reporting gets amplified as the size and importance of the undertaking increases. Thus, managers who are tasked with the oversight of large and/or important projects will be well advised to pursue tactics that could reduce the potential benefit of biasing and increase its probability of detection. Such tactics include regular communication audits, the use of multiple sources to triangulate the exact state of project tasks without relying on a single source, and the nurturing of an open project communication culture that minimizes the need for face saving behaviors. As the pursuit of such tactics is likely to require substantial resources and time, we advise project executives to focus on high biasing-risk projects when identifying initiatives that could benefit from the use of such measures. Moreover, it is important for such executives to recognize that the potential for increased biased reporting can be countered through the use of normal project control activities, such as internal integration efforts (e.g. regular project meetings) as well as formal planning techniques (such as project management methodologies and software). Our empirical findings have confirmed the suppression effects of such controls on biasing.

Finally, our study’s findings yield several opportunities for further research. First, it would be interesting to see whether the lack of an influence of project structure on biasing would hold in a different sample that included more low-structure projects. Perhaps there is a necessary threshold with respect to project structure that has to be reached before it begins to influence biasing behavior. Second, our results were obtained from projects undertaken by governmental agencies. While we anticipate that the findings would hold across projects drawn from the for-profit sector, it would be useful to replicate the study in a different (non-governmental) context to confirm this assumption. Third, our study’s sample consisted of IS professionals acting as project team members reporting to their project manager. It could be fruitful to see if similar reporting behaviors exist in other dyads that are present in project settings (when a project manager reports to an auditor, for example). Fourth, now that we have established that project characteristics do influence the propensity of reporters to bias their status updates, it would be useful to identify additional project traits that could complement the set of the four that we have identified. This will yield a more comprehensive perspective on the effects that project environments have on biasing.
In conclusion, our research findings have confirmed that biasing is indeed present in IS projects and has identified three project characteristics that influence it. We hope that the results of our study will guide practitioners in improving their project monitoring and oversight practices and will provide fruitful exploration avenues to researchers who are interested in helping such executives. Only with such additional scholarly explorations will we be able to gain a complete understanding of the biasing phenomenon and find ways to reduce its undesirable effects.

References


IS PROJECT EVALUATION METHODOLOGY - SCIENCE OR ART?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0426.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Actor-network theory, Evaluation, IS Investments, Case Study</td>
</tr>
</tbody>
</table>
Abstract

Pre-investment evaluation of information system project proposals persists in being problematic and highly risky in practice. Numerous evaluation approaches and methodologies, offered in the literature, have not contributed to major improvements in practice. As the literature review shows, these methodologies have assumed an ideal of objective and scientific evaluation and taken the view of methodology as science. In this paper we aim to revisit the notion of IS evaluation methodology in practice and specifically explore the methodology as both the ‘science’ and the ‘art’. We achieve these aims by conducting an Actor Network Theory (ANT) study of IS evaluation in a large financial services company. The ANT study allowed us to investigate the methodology as it is enacted in IS project evaluations in practice without making any a priori decision of what methodology is or should do. By defining a series of processes, inscription aids and mandated checkpoints we found that the evaluation methodology engenders the evaluation process as ‘science’; and by allowing a degree of freedom in conducting the evaluation processes, the methodology is also enabling the evaluation processes to emerge as ‘art’ and by doing so stimulating creativity and innovation. Thus the ANT approach assisted in our discovery of the dual nature of methodology as simultaneously science and art.

Keywords: Pre-investment IS project evaluation process, IS project proposal evaluation methodology, Actor-network Theory, socio-technical approach
1 INTRODUCTION

The increasing costs of investments in information systems (IS) and persistently high failure rate of IS projects in practice indicate that IS project evaluation is of critical importance to organizations. The evaluation of IS projects before the decision is made to fund it – referred to as pre-investment evaluation – is potentially critical to increase both the success rate of IS projects and to achieve expected business value. Making the right IS investments is, and will continue to be, a critical component of organisational planning for the future (Lin & Pervan 2001b, Murphy & Simon 2002).

To improve the practice of pre-investment evaluation of IS project proposals, IS researchers have over the years proposed numerous evaluation methodologies with the aim to advance and assist ‘rigorous’ and ‘scientific’ assessments (Berghout 2001). The focus of scientific evaluation of IS projects has been on the tangible aspects of IS costs and benefits, thus privileging the technology and its measurable impact on performance. The assumption has been that the more evaluation methodology is scientific, that is objective, rigorous and based on precise measurements and calculations, the more accurate and certain the evaluation result will be. The failure of evaluation of IS is attributed to the lack of appropriate adoption of the scientific rigour of the methodology.

The IS evaluation literature has been largely dominated by an objectivist view of the IS evaluation process, technological determinism and the perception of IS evaluation methodology as a ‘science’. However, there are indications which suggest that organizations generally ignore these IS evaluation methodologies, because they are perceived to be too complex, difficult to understand and unreliable (Al-Yaseen & Eldabi 2004). While we have seen a growing number of ‘scientific’ methodologies in the literature (Nagm & Kautz 2007), we have also seen reports about the lack of methodology adoption in the practice of IS evaluation (Nagm & Kautz 2008). Is something wrong with practice or with evaluation methodologies?

We propose that IS evaluation in practice is not well understood by proposers and proponents of “scientific” evaluation methodologies of IS. In this paper we argue that IS evaluation in practice is not only a science but is also an art form which involves creativity, improvisation, imagination and community building. More precisely we aim to demonstrate that the IS evaluation methodology in practice has a dual nature: that it is simultaneously science and art. We achieve this by conducting an Actor Network Theory (ANT) (Latour 2005, Law 2002) study of IS evaluation in a large financial services company. The ANT study enables us to investigate the nature of methodology as it is enacted in IS project evaluations in practice without making any a priori assumptions about what a methodology is or should be.

Before we present our results we first discuss the IS evaluation literature and raise key problems and themes. We proceed by presenting the ANT approach to the case study. We then present an ANT account of an IS project evaluation methodology in an exemplary case company, based on which we reveal and discuss the dual nature of methodology as simultaneously science and art.

2 LITERATURE REVIEW

The IS evaluation literature is divided into two distinct areas, pre-investment evaluation and post-implementation evaluation. The focus of pre-investment evaluation is on the justification of IS investments before being initiated (Murphy & Simon 2001). It is also known as ‘predictive evaluation’, emphasizing the speculative nature of the estimation of an IS worth and impact on the organization in future that depends on the evaluator’s judgement (Remenyi & Sherwood-Smith 1999). This differs from post-implementation evaluation which aims to evaluate IS projects after the implementation has occurred. In this paper, we focus on pre-investment evaluation of IS projects.

The predominate stream of research in the IS evaluation literature has been on the development of ‘scientific’ evaluation methods and their use in practice. There have been numerous attempts at counting and investigating these methods, Berghout (2001) for instance identified the existence of
Over 65 evaluation methods, most centred around ‘scientific’ appraisal techniques like discounted cash flow analysis (DCF), net present value (NPV), internal rate of return (IRR), cost benefit analysis (CBA) and payback period. These methods often described as being formal, summative or functionalist. They offer some advantages for IS investments that are expected to produce cash flows or some other tangible financial benefit, but disadvantage IS investments which may add value to the organization in a less tangible way. They exemplify a scientific ideal to prescribe a rational, formal, quantitative and exact evaluation process. These represent as Stockdale, Standing & Love (2006) argue the hallmark of the scientific paradigm.

These ‘hard evaluation techniques’ (Ballantine, Galliers & Stray 1994) or ‘rigorous scientific methods’ (Smithson & Hirschheim 1998) have been often questioned and criticized:

- For their assumption that value and impact of IS can be reduced to numbers which can be calculated and counted (Williams & Williams 2004).
- They ignore intangible/qualitative aspects of IS investments (Hirschheim & Smithson 1999)
- They ignore the wider context within with the evaluation takes place (McBride & Fidler 2001)
- By ignoring organisational issues, they are difficult to apply (Ballantine & Stray 1998)
- They are susceptible and sensitive to manipulation (Hirschheim & Smithson 1999)
- They stifle innovative development by focussing on formal-rational approval processes (Howcroft & McDonald 2004)
- They obscure subjective aspects of IS evaluation (McBride & Fidler 2001)

More recently there has been a growing awareness among researchers that the scientific rational view of evaluation has to be expanded, perhaps being replaced by a perception of evaluation as a social and political phenomenon (Berghout, Nijland & Grant 2005). A number of alternative or complimentary methods have been proposed that include more ‘subjective’ aspects of evaluation (Klecun & Cornford 2005). The emergence of such methods has no doubt been inspired by a new way of thinking in the IS discipline. For instance May (1997) states:

“This new world of IT-enabled value creation contains none of the cold sterility of scientifically precise, formula-driven absolutes. Our central certainty of value has given way to a series of negotiated, mutually-agreed-on business objectives” (pg. 92)

However despite the criticism we still don’t have a clear understanding of IS evaluation methodology that is more than science. If IS evaluation is not a completely scientific enterprise, the question is: What is it? If an IS evaluation methodology is criticised for being overly "scientific", what is missing?

We sought to answer these questions by drawing on an empirical study of IS evaluation processes in practice. Without assuming any view or position from the literature we aimed to find out what is a nature of IS evaluation methodology in practice. We therefore examined the practices of IS evaluation in a company that has extensive experience in assessing, developing and deploying IS, and a track record of successful IS adoption and implementation. It also has well established processes for developing and evaluating IS proposals, including the use of a range of evaluation techniques. Throughout the empirical study, lasting 16 months we encountered many actors, such as business managers, IS managers, project managers; a plethora of documents including an IS evaluation methodology, IS project ideas, project proposals, IS business case, etc. We came to realise that all of them, human actors and various documents and technologies (that can be seen as non-human actors or ‘actants’) play a role and exert agency due to their mutual interactions and influences. This led us to adopt actor-network theory (ANT) to enable a more comprehensive understanding of IS project proposals and their evaluation methodologies and processes (Latour 2005, Callon 1986, Law 2004).
METHODOLOGY: AN ANT CASE STUDY

ANT is an approach to sociomaterial analysis developed in the sociology of science and technology (Callon 1986, Latour 1986, 2005, Law 1999, 2004) which found a fertile soil in IS research (Walsham 1997, McLean & Hassard 2004). ANT considers entities such as IS, organisations, methodologies, and users, as actors or actants that have agency by virtue of their relations with other actors/actants in heterogeneous networks (Law, 2004). ANT is grounded in an ontology of relationality and treats actors as enacted and relational effects. The world is seen as being comprised of enumerable heterogeneous actors engaged in a continuous process of mutual interconnecting to form or diminish actor-networks. Actor-networks thus emerge, grow, sometimes stabilize and black-box and sometimes destabilize and dissolve. The emergence of actor-networks result from enrolment of other actors (allies) and alignment of their interests with that of the network. Generally, ANT enables analysis of the conditions, constraints and modification of agency within actor-networks that intertwine the humans, culture, language, artefacts and technology (and many other things).

In particular two concepts drawn from ANT assist in analysing IS evaluation practices. These concepts are, ‘translation’ and ‘inscription’. Translation is the process by which an actor creates a body of allies by enlisting other actors to align with their actor network (Walsham 1997, Latour 2005, pg. 218). Inscription describes the process by which designers inscribe their interests into technical artefacts such as information technologies or into other actants such as policy documents and methodologies. Inscription can be said to define the roles and use of these artifacts and has been described as “frozen organizational discourse” (Walsham 1997 pg. 469).

Furthermore, ANT presents challenges for researchers aiming to investigate emergence of heterogeneous actor-networks. How to adopt ANT to conduct empirical studies remains open to researchers’ imagination and is not prescribed by ANT’s proponents and followers. For instance in their particular kinds of ethnographic studies – Latour’s investigation of a “laboratory life” (1979) and later on a failed technology project called Aramis (1986), Law’s aircraft stories (2002) and Mol’s treatment of atherosclerosis in a hospital (2002) – they followed their objects and subjects, recorded events and collected other material evidence in many different ways. The major feature in their work is a detailed description of a story as a basis for examining and theorizing novel and often complex concepts and questions. Latour himself describes ANT as a “very crude method to learn from actors without imposing on them an a priori definition of their world building capacity” (1999, pg. 20). We learn from him and other ANTs to ‘follow the actors’, let them tell their own stories, use their own vocabularies and unfold their own meanings.

Our study started without any pre-conceptions about a theoretical foundation or the nature of evaluation processes. It was motivated and initially driven solely by the research questions. It focused initially on the human actors and what they do, how they go about proposing ideas for new projects and how the ideas grow into official IS Project Proposals. In the process we were particularly interested in the evaluation processes and how methodology acts as a non-human actor. It turned, almost intuitively, into a ‘journey’ of following the actors, not only humans but also objects, documents and other devices employed during projects’ evaluations.

We conducted our study in a large multi-national financial services company referred to as ALFA Group. The Australasian operation – which we shall call ALFA Bank – has an impressive history spanning approximately 150 years and with an investment portfolio in excess of $1 trillion dollars. One of its divisions, ALFA Invest, was the prime focus of the study. This organisation was selected because the company is known to have a well established practice of IS project proposal evaluations; secondly it has a track record of successful IS deployment and implementation; thirdly, it has recently overhauled its evaluation methodology worldwide and finally because it was quite receptive to our request to study these practices in depth; hence this presented a unique opportunity to study the adoption of the methodology in depth.

A financial crisis in 2004 forced ALFA Group to reassess the way in which it evaluated IS project proposals. In order to manage the $2 billion commitment which ALFA Group made to rebuild ALFA
Bank, ALFA Group decided that it needed more discipline around the way proposals for the $2 billion spend would be evaluated. To facilitate this, a new Evaluation Methodology was introduced with two core principles at its heart – doing the ‘right projects’ and doing the ‘projects right’. From 2005 to 2007, over $3 billion worth of projects have gone through the new methodology which has helped ALFA achieve a way to produce consistent and comparable IS Project Proposals.

Data collection includes a) interviews with 36 senior executives and managers of ALFA Invest division as well as some from ALFA Bank (the parent company) over a 16 month period from July 2006 to October 2007 (listed in Table 1), and b) company documents related to IS project evaluation methodology including examples of recently approved IS project proposals.

Table 1  Interviews conducted in ALFA Invest and ALFA Bank

<table>
<thead>
<tr>
<th>Roles</th>
<th>Human actors</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS Management (13 people)</td>
<td>Head of IS Group&lt;br&gt;Head of Business Demand (two people)&lt;br&gt;Head of IS Architecture&lt;br&gt;Head of Application Development Management&lt;br&gt;Head of IS Development&lt;br&gt;Head of IS Business Support&lt;br&gt;IS/Business Relationship Partners (six people)</td>
</tr>
<tr>
<td>Business Management (8 people)</td>
<td>CEO and Chairman of ALFA Group&lt;br&gt;CEO of ALFA Bank&lt;br&gt;Head of Strategy&lt;br&gt;Chief Operating Officer&lt;br&gt;General Manager – Business Unit&lt;br&gt;Head of Financial Planning&lt;br&gt;Head of Central Business Operations&lt;br&gt;Head of Business Development</td>
</tr>
<tr>
<td>Projects &amp; Project Management (15 people)</td>
<td>Senior Project Analyst (five people)&lt;br&gt;Senior Project Manager (four people)&lt;br&gt;Project Director (two people)&lt;br&gt;Head of Project Methodologies&lt;br&gt;Head of Projects&lt;br&gt;Head of Portfolio Management&lt;br&gt;Head of Regional Projects Board</td>
</tr>
</tbody>
</table>

Our investigation started with an interview with a Senior Project Analyst who explained how projects are viewed by people in the company, as well as how they are evaluated within the broader business context. From that point the inquiry detoured into several directions following the Senior Project Analyst’s suggestions including Head of Regional Projects Board, and other Senior Business and IS Analysts. We started with semi-structured interviews guided by an interview schedule but soon departed from it and adopted unstructured interviews that proved more suitable to addressing emerging issues. The length of each interview on average was approximately one hour, but in some cases interviews spanned two hours over two separate sessions. Through these interviews we encountered non-human actors, IS project documents, evaluation methodology and strategy.

The analysis of empirical data started early on as we encountered new actors (humans and non-humans) and cannot be clearly separated from data collection. Namely, following the actors and their relations with objects and other actors prompted the chain of interviews and collection of documents. The analysis of the interviews and documents in turn led to seeking explanation of activities, events and outcomes by interviewing new actors and seeking new documents. These interviews revealed how different business realities and interests are negotiated and inscribed in the production of IS project proposals. After data collection the analysis became more refined focusing on stages in the production of IS project proposals, evaluation methodology, the roles of specific actors or actants and the ways they enrolled and acted upon the production and evaluation of the proposals while creating actor-networks. Snapshots of networks were presented in a graphical form as ANT maps. Our ANT inquiry included coding the interviews and documents using Nvivo. The coded texts were then extracted in a
systematized form, assisting our investigation of actions and their relations in the networks. These also helped writing the story of IS projects evaluation processes and the nature and role of methodology.

4 IS PROJECT EVALUATION METHODOLOGY: DEFINING THE ‘SCIENCE’ AND ENABLING THE ‘ART’

At the pre-investment phase of a project, each IS project proposal has to go through five stages of the evaluation methodology, the Idea Stage, Concept Development Stage, Business Case Stage, Proposal Submission Stage, and Proposal Assessment Stage. In the process an IS project proposal is developed and continually evaluated. Initially, an idea for an IS grows and increasingly becomes more defined and concrete to the people involved as it moves through the stages defined by the methodology. From the first idea stage, which involves informal discussions and the fermentation of ideas, the IS idea converges into a concept through the Concept Development stage before forming a coherent IS Business Case drafted during the Business Case stage. This business case is then transformed into an IS Investment Submission document before being further refined and shaped in a consolidated form that will be put forward to the Regional Investment Committee for deliberation during the final proposal assessment stage.

For example the business case stage as described by Figure 1 involves a number of human actors (like the Business Sponsor, Finance Partner, Critical Business Stakeholders and more) who Business Case Document. As these actors directly determine the content of the Business Case Document, we presented them in the core of the actor-network. The core describes the actors and relations at the centre of an activity. Other actors such as Business Plan, ALFA Group Management and the Methodology are presented in the periphery which indicates that they are acting from a distance.

Figure 1 - IS project evaluation methodology acting from the periphery of the actor-network
The methodology in this case acts in two major ways: first, the methodology can be seen to act as a scientific method by prescribing a set of systematic processes, inscription forms, norms, rules and checkpoints that need to be passed. The evaluation methodology as science is systematic, formal, prescriptive and rigorous. It ensures standardization and comparisons across projects, equitable treatment of all projects, and involvement of relevant stakeholders in each project evaluation. However, it is at the same time enabling and encouraging the “art” of evaluation to emerge. The art in this context refers to the “expression or application of creative skill and imagination” (OED 1989), innovation, improvisation and informal relations. In the section below, we explore these two ways in which the methodology acts – as science and as art – and in doing so, reconceptualise the notion of an evaluation methodology.

4.1 Evaluation methodology acts by defining the ‘science’

By defining a number of elements, the methodology acts to inscribe the process interests as well as control (checkpoints), standardisation (inscription norms), consistency (inscription forms), and policing (inscription rules) interests of ALFA Group management.

The methodology specifies a number of processes that should be carried out in each stage; it also defines key ‘obligatory’ human actors that need to be engaged for certain processes. For example, the methodology stipulates that the Business Sponsor needs to approve and sign off all major documents (for instance the IS idea document) before they can be submitted to the Local Investment Committee for deliberation as Figure 2 below illustrates.

While there are some processes that must be performed as defined by the methodology, others are less prescriptive. For example in producing the IS Concept document the methodology defines a linear sequence from developing the concept, to developing the high-level architecture to finalising and reviewing the concept document; however, in practice it is possible for all three processes to be carried out simultaneously as Figure 3 demonstrates.
The methodology does describe some actors and processes but leaves others flexible. In such a way the methodology is an actor that shapes the IS Project Proposal emergence sufficiently strictly to meet management needs, but with the flexibility to ensure enrolment of actors and the growth of the actor-network. The power of the methodology is in part the ‘science’ it imposes (in varying degrees) and in part the ‘art’ it enables in creating relations and building networks.

The methodology does not only define a series of processes to follow, but also defines a set of inscription devices to complete (see Figure 1), in order to create an audit trail and evidence of actions in all stages. The Head of Project Methodologies in ALFA Bank states: ‘the methodology we have built, is saying what you have to do, [and] also provides the tools and templates to help project managers actually do it and to deliver [projects]’.

While the methodology does not define how these inscription devices can or should be put together, it does define a series of inscription forms, norms, and rules that help create inscriptions that strategies and business plans are enacted. It also ensures that different IS project proposals can be compared to one another. These inscription aids also act to ensure that people are serious when they develop IS Project Proposals. There is a audit trail that is left behind. To enforce some control especially around estimations and value articulation, ALFA Group have embedded inscription rules into key inscription devices (financial analysis spreadsheet for example) that cannot be manipulated and must be reviewed by GM Finance.

The methodology also acts by clearly defining a number of checkpoints that each IS Project Proposal must pass through from the Idea Stage right through to the Proposal Assessment Stage. The Head of Project Methodologies in ALFA Bank states:

‘At the end of these stages you have what is called a ‘stage gate’ where you have to pass the gate to get into the next stage. So there is a check mark by certain people at the end of each of those stages to see if the project should then progress to the next stage’

These checkpoints, called ‘stage gates’ in ALFA, not only act out management’s control interests but they promote a practice of continuous evaluations which are both formal and informal. Importantly for each checkpoint, a ‘gatekeeper’ has been made obligatory by management through the methodology ensuring that progress of an IS Project Proposal from one stage to the next is regulated and policed by someone who does not stand to benefit or suffer from the progression of an IS Project Proposal. By defining a number of checkpoints, the methodology acts by helping to identify and eliminate ‘bad’ proposals; those that do not provide justifiable benefits, are not feasible, are not supported by key actors or are misaligned with the strategic priorities of the business. The current methodology has been instrumental in helping to ‘stop wasting a lot of the organization’s money on projects that never go anywhere’ (who said this?). In fact, statistics in ALFA suggest that the methodology has helped stop 30% of all projects before they were approved. These projects would have been allowed to go through...
with the previous methodology, and this demonstrates how the methodology acts to increase instances of project success by helping to detect and eliminate bad projects (those with a weak financial case, unsupported by key people or unaligned to agreed strategic priorities). It also acts by amplifying beneficial, valuable or strategically aligned IS Project Proposals.

4.2 Evaluation methodology acts by enabling the ‘art’

By not being overly prescriptive, the methodology also acts by providing the creative freedom for actors to explore the ‘art’ in IS project evaluation – the art of translation. The art of engaging with others, the art of finding the right stakeholder or forum to present ideas, the art of persuading and influencing and enrolling others, the art of shaping, synthesising, sizing and crafting the IS proposals, the art of rallying support in meeting rooms and corridors, the art of story telling, the art of adapting (undefined, and un-prescribed) paths to develop an IS Project Proposal -- what the people in ALFA call a journey. The ‘journey’ refers to how people enact the processes, inscription aids and control points to enrol actors, to build allies and ultimately to make a compelling case for the selection of the proposed IS over others which are all competing for funds from the same limited pool. The methodology acts by not explicitly defining the paths people take to travel between processes and by not strictly defining how they carry out the processes. This enables people to approach each project different and attend to unique nuances that arise. By acting to allow both the ‘science’ and the ‘art’ in IS project evaluation the methodology is both a powerful and essential non-human actor in the heterogeneous actor-networks that help steer the emergence of the IS Project Proposal.

Allowing individual people to adopt their own approach to developing an IS Project Proposal, ensures that the culture of each local business unit can influence ‘how’ IS Project Proposals are initiated and assessed. It also ensures that people are free to find their own path to develop the IS Project Proposal recognizing not all projects are the same and that prescribing a strictly pre-defined path would not work in practice. Thus, because the IS Project Proposal involves many different people (business and IS actors) each bring to the journey their own understanding of process, and their own preferences for the path to take. This means potentially multiple different paths exist which can be taken by different people simultaneously. Many people in ALFA have described it as a ‘journey’ because the path is emergent, unknown and always different. For example, during the IS concept development process in the methodology, everyone has a different view of how the concept should be developed, what approach should be taken, which processes and tools should be involved. The methodology does not define how this should be done, or even how to clarify differences. That is why the people selected as IS project champions and IS/business relationship partners must have gone through development and assessment of many other IS project proposals. Their experiences, relationships and trust thus created enable identification and enrolment of the right people to the project the actor-network. Through organizing forms (meetings and forums) actors come together, raise issue, question, and debate the approach to take to develop the concept. They do this not in a predetermined manner, but they will do in the words of the Head of Central Business Operations ‘whatever works’.

Finding their own way to travel through the process defined in the methodology does not mean that anything goes or that projects are necessarily chaotic. As long as they do not break the inscription rules - rules that bring some order to the creative process, and which are defined in the methodology, actors can pursue different ways to debate ideas, clarify ambiguity, minimise conflict, and resolve misunderstanding. Similarly, they use different ways to inscribe their interests and visions in inscription devices. In this way, the methodology acts by being tolerant of multiple, messy and non-coherent realities that gradually cohere through the journey – helping to deal with both conflict and ambiguity.

Even though processes can be defined and standardised in a methodology, what matters is not the perfection of these defined processes or the prescribed elements but rather what the methodology does not prescribe – how people take others on the journey of translation. The journey adapts to the individual and specific needs of the project and is not determined by a ‘one size fits all’ methodology. This is like ‘stepping into the unknown’ as the Head of Central Business Operations explains, there is not necessarily ‘a clear way of getting to anywhere’. This is because the ‘best path needs to be found’.
‘Sometimes -- he continues -- you have to take the long way and sometimes you have to take the short way’. Similar views are shared around ALFA. As one IS/Business Relationship Partner says; ‘every initiative that goes through that process [IS Project Proposal development and assessment] has to find its own way to get there’. In finding the path as the Head of Central Business Operations states, ‘you evolve and adapt and think “what do I need to do to get through”’ and concludes by saying ‘not all roads are going to lead us to there but eventually with the right convergence you’re going to get there’. This makes the practice of IS project evaluation less of a ‘science’ and more of an ‘art’. As a Senior Project Manager comments, ‘there is some art within each of those steps of the process’.

IS Project Proposal evaluation being part ‘art’ does not mean that the journey is always smooth. Conflict does arise in many ways and forms as is evident in the ALFA case. Conflict between human actors (actor-network versus actor-network) for example who are competing for limited funds to demonstrate they have contributed to the organization through projects (that add strategic value) and wish to improve their bonuses for the year. Conflict also comes about within the actor-network as personalities and viewpoints clash, or with people who have hidden agendas or intend to cause disruption to a proposed IS. Ambiguity also surfaces around the initial idea, around how it should be conceptualised, around the estimations of costs and benefits. Ambiguity exists around the new business reality enacted by the proposed IS and how this reality is aligned to strategic priorities and the existing portfolio of projects. However, a prescribed rulebook on how to deal with conflict and ambiguity does not exist in this case, perhaps because conflict and ambiguity cannot be dealt with ‘scientifically’. Actors do find ways to deal with both elements (conflict and ambiguity) through the ‘journey’ itself. In dealing with both conflict on the one hand and ambiguity on the other through the journey, the actor-network stabilises resulting in IS Project Proposal that sufficiently inscribes the interests of the actors involved.

The evaluation methodology presented in the ALFA case, thus acts in these ways; first it defines the ‘science’ through a set of processes to follow, and people that need to be involved, and by defining a series of inscription aids (inscription forms, inscription norms and inscription rules), as well as defining a series of mandated checkpoints. Second, it enables IS Project Proposals to be described by comparable and unambiguous sets of measures. Third, it enables the ‘art’ of translation to surface, and thus encourage innovation and creativity in every IS project journey, allowing enrolment of relevant actors and inscription of their interests.

5 CONCLUSION

The paper makes two contributions. First the paper demonstrates that the IS Evaluation Methodology in practice has a dual nature, that it is simultaneously a science and an art. It redefined the nature of IS project evaluation methodology by revealing how the science and the art of methodology are intertwined and mutually dependent. Second, the paper contributes to better understanding how IS evaluation methodology encourages IS project proposals that are likely to be successful and identifies and discourages those that are likely to fail.

The dual nature of IS evaluation methodology has important implications for practice. The methodology-as-science acts though a series of processes, inscription devices, enrolment of key people, and mandatory checkpoints. It thereby ensures that systematic steps are taken and responsible actors engaged in the process. It also ensures that different IS project proposals by various departments are presented in a comparable form, using similar measures. However, this is just one side of the story. The methodology-as-art acts by intentionally introducing flexibility in the processes and procedures, leaving freedom for participants to adjust the processes and find the appropriate paths to develop and evaluate proposals. The art of the methodology is necessary to deal with realities that are in flux, not well defined and often non-coherent. This dual nature of the methodology as a science and as an art describes what actors in practice call the “journey”.

The notion of the dual nature of the IS evaluation methodology addresses many criticisms of the dominant view of methodology in the literature: that qualitative and organisational aspects are ignored (Hirschhem and Smithson 1999), that it stifles innovation (Howcroft & McDonald 2004), obscures
subjective aspects of evaluation (McBride and Fidler 2001), and that the wider context within which evaluation place is ignored (McBride and Fidler 2001).

Understanding the IS evaluation methodology as a science and an art is important because it reveals the production of successful and unsuccessful projects. As an art the methodology leaves space for the actors to negotiate and enact their meanings, and translate interests and objectives in the IS project proposals. These include negotiation and checking that the projects are aligned with the company strategic priorities. The trails of these negotiation processes are recorded in various documents (an IS idea, IS concept or a business case document) as inscription devices prescribed by the methodology as science. In such a way the discursively produced IS ideas, concepts and business cases are materially inscribed in the respective documents. This is a basis for “detecting” weak or bad IS project proposals, and for their elimination before reaching the final stage of project selection. The implications for a company are significant: it identifies and eliminates bad project early on and thus saves time, money and effort in the organisation. The methodology also acts by “amplifying” beneficial or valuable IS project proposals from the point of view of different participants.

Our findings also show that the practice of IS project evaluation – for which various evaluation methods propose a range of rigorous and well structured processes, models and calculation techniques – is rarely rigorous, is often messy, and seemingly unsystematic. Are these signs of poor evaluation methodology? – as is assumed in most IS evaluation literature. Should (could) the practice be improved by the adoption of, and stringent adherence to, a more rigorous, formal and exact methodology as most of the literature argues? Our case company ALFA Invest learned hard way the answers to these questions – through their own mistakes. Lessons from this company indicate that this is a limited view of practice and evaluation methodology that neglects the embedded art forms of project proposal development and evaluation. It is now the researchers’ turn to find more evidence from practice and develop further the theoretical and practical understanding of IS project proposal evaluation methodology.

References


AN EXPLORATORY EVALUATION OF THREE IS PROJECT PERFORMANCE MEASUREMENT METHODS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0697.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Design Science, IS Performance, IT performance management, IT Project Management</td>
</tr>
</tbody>
</table>
AN EXPLORATORY EVALUATION OF THREE I.S. PROJECT PERFORMANCE MEASUREMENT METHODS

Barclay, Corlane, University of the West Indies, Departments of Computing & Management Studies, Mona, Kingston, Jamaica, CLBarclay@gmail.com

Osei-Bryson, Kweku-Muata, Virginia Commonwealth University, School of Business, Richmond, Virginia, USA, KMOsei@vcu.edu

Abstract

Information systems projects play an important strategic role in organisations and are key drivers to the delivery of change. Given this prominence it is essential to find measurement methods that effectively analyse and communicate the performance to its stakeholders. Further, to assure contribution to both research and practice it is essential to verify the utility of the artefacts (i.e. methods) developed to help validate or justify that the solutions are suitable for practice, and fit the needs and contexts for which it is created.

Grounded in the design science paradigm, this paper reports an exploratory evaluation of the perception of certain qualities of two recently developed measurement methods (The Project Performance Scorecard and Project Objectives Measurement Model) against the traditional Triple Constraint method. An analytic scenario-based survey of fifty-one (51) participants, comprising of three (3) sets of independent sample of seventeen (17) respondents each was used. The study analysed dimensions of task performance, ease of use, perceived usefulness, perceived semantic qualities and user satisfaction from the perspective of the participants. The preliminary study revealed encouraging results for the new methods and the general design process which can help guide current use and further refinements. The limitations of the study and future research directions are discussed.

Keywords: IS projects; measurement methods; project evaluation; design science.
1 INTRODUCTION

Recent developments have seen increased proposals on suitable perspectives or approaches for the evaluation of information systems (IS) project performance. For many years, the de facto standard has been the standard conformance to time, cost and specifications, or the triple constraint method (TCM) (Atkinson, 1999; White & Fortune, 2002). However, some researchers have questioned the suitability and completeness of this approach to effectively analyse the contribution of IS projects to the organisations and its stakeholders (Atkinson, 1999; Cohen & Graham, 2001). In response, several alternative approaches have been developed (Atkinson, 1999; Stewart, 2008) to help address this gap. However, an analysis of the project evaluation literature reveals that diverse empirical investigations have also become a top priority.

Against this background, the research adopts the principles of design science (Hevner et al., 2004) and evaluates two recently developed measurement methods, the Project Performance Scorecard (PPS) (Barclay, 2008) and Project Objectives Measurement Model (POMM) (Barclay & Osei-Bryson, 2008) against the standard approach, the TCM. Hevner et al (2004) proposed that the utility, quality, and efficacy of a design artefact must be rigorously demonstrated via well executed evaluation techniques, and can be evaluated in terms of functionality, completeness, consistency, accuracy, performance, reliability, usability, fit with the organization, and other relevant quality attributes. They further proposed several evaluation strategies including: observational (e.g. case studies and field studies); descriptive (e.g. scenario construction); analytical (e.g. static analysis, architecture analysis); experimental (through controlled experiments and simulation); testing (through functional or black box and structural or white box testing). As part of the design process, observational case studies were first used to help validate and justify the two artefacts/methods in previous studies (see Barclay 2008; Barclay & Osei-Bryson 2008). This approach was used as it allows for an in-dept study of the given artefact in an organizational context that would in turn provide feedback that may be used for the appropriate improvement of the artefact (Hevner et al., 2004). In this current exploratory study, the evaluation process is extended through the application of a mixed-method approach in the implementation of both the descriptive and analytical techniques. This involves the development of a real-life project scenario and the evaluation of the structures of each method using static analysis (e.g. complexity or performance) to identify attributes based on the perceptions of business and project practitioners. Static analysis is commonly used in software development to analyze the components and resources without running the programme (Chess & McGraw, 2004). While the proposed methods are not software components, the technique is suitable as practitioners are given the documented information (i.e. the components) to analyze each method in terms of elements such as usability, perceived semantic quality and completeness. To support this process, the conceptual model proposed by Maes & Poels (2006) is used as a basis to develop the research measurement instrument. This instrument was chosen because it addresses the limitation of procedures to help analyse the static qualities of IS artefacts (e.g. models or methods), and is based on seminal IS Success models (DeLone & McLean, 1992, 2003; Seddon, 1997), and are therefore consistent with the goals of this research.

The research provides contributions to both research and practice. The study extends the knowledge base in the IS project evaluation field through the analysis of the empirical observation of users’ interaction with three (3) distinct measurement approaches. The exploratory findings also have positive implications for both newer methods, and provide an opportunity for them to be implemented in practice to assist in the analysis of the contribution and performance of IS projects.
2 OVERVIEW OF IS PROJECT EVALUATION METHODS

There are various forms of evaluation methods that have been applied to analyse the performance of projects in different contexts. The traditional approach (TCM) has been the most popular technique used to help determine the success or performance of many projects in industries, and is largely endorsed by project management bodies (PMI, 2004). With the growing demand to find techniques that better fit the complexities of contemporary projects, research has begun to respond to this call. Some studies have reviewed factors impacting performance while others have developed strategies to assess performance. A review of key project evaluation literature reinforced that the primary responses to examining the performance of projects, outside of the TCM, have been to develop alternative success criteria to assess these projects (Atkinson, 1999; Morris & Hough, 1987; Nelson, 2005), and critical success factors (Belassi & Tukel, 1996; Fortune & White, 2006; Pinto & Slevin, 1987; Shenhar et al., 2002) while others have focused on the business value contribution of these investments (Fitzgerald, 1998; Kaplan & Norton, 1992; Kumar, 2003). Similarly, in an analysis of the project literature over a forty-year period, Jugdev & Muller (2005) showed the evolution of our understanding within the framework of the project and product life cycles in the discourse of performance evaluation perspectives. The four (4) evolving research themes were categorized into project implementation and handover, critical success factor (CSF) lists, CSF frameworks and strategic project management paradigms (Jugdev & Muller, 2005).

It is noted that the studies on project evaluation techniques or methods have been a mix of conceptual and empirical studies, with the latter primarily involving only validation or observation of the proposed solutions. This provides an opportunity for our current examination.

2.1 Description of Alternative Project Evaluation Methods

2.1.1 The Project Performance Scorecard

The Project Performance Scorecard (PPS) proposed that several considerations are essential to help project practitioners to enhance their analysis of the contributions of IS project activities. It is argued that to effectively analyse performance in these dynamic projects, a multi-dimensional perspective is necessary to help provide a more robust view. This includes considerations of the project events throughout the life of the project, the project management processes and the impact of the project’s product. Six (6) interconnected dimensions are therefore introduced which can be used as a platform to evaluate the performance or success of the project in different organisational settings: stakeholders, project process, benefit, quality, learning and innovation, and use (Barclay, 2008), figure 1. The PPS framework relied on theoretical contributions of IS success (DeLone & McLean, 2003), Balanced Scorecard (Kaplan & Norton, 1992) and quality paradigms to help communicate its purpose.

*Stakeholders* include those that are involved in the program or have a vested interest in its outcome (PMI 2004; 2006). While there may be stakeholders with negative and positive agenda, the method concentrates on the positive view. Understanding and accounting for what is important to the stakeholders establish the framework for better management of expectations during the project because of earlier involvement and consensus of what is important to them.

*Project process* considers the processes of the project from conception to handover to the client. It incorporates the view of the classical paradigm and looks closely at the project processes to gather
insights into areas such as the financial gains, the efficient use of time, management of uncertainty and resource allocation.

*Benefit* focuses on the gains and business value that are attributable to the project including the rationale for the project selection and considerations of the project business case are embedded within this dimension. It is important to note that while organizations are primarily concerned with the bottom-line; there are other important considerations that affect performance (Kaplan & Norton, 1996).

![Figure 1: The Project Performance Scorecard](image)

*Quality* is concerned with meeting or exceeding the stated objectives. Within the IS context, considerations of reliability, usability, efficiency, maintainability, portability and functionality of the product are essential in determining the value of the product and the project process.

*Learning & Innovation* focuses on the knowledge capabilities that can be garnered from the project, including gains, advantages, and value creating capabilities that may have been arisen as a result of involvement in the project.

*Use* considers how the results of the process are being used. For example, providing a software application aimed at improving internal efficiency is only part of the contribution, assessing how the application is being used and whether it achieves its stated objectives are necessary components in determining the contribution of a project.

2.1.2 The Project Objectives Measurement Model

The Project Objectives Measurement Model (POMM) involves the elicitation and development of objectives and measures that reflect the strategic and tactical vision of the project from the perspectives of its multiple stakeholders (Barclay & Osei-Bryson, 2008). It is distinct from other methods including the PPS as it focuses on the structured development of project criteria that are representative of the project stakeholders and does not rely on particular measurement dimensions.
Three key questions are reflected throughout the framework: do the project measures reflect the fundamental objectives identified? Do the project objectives reflect the project contexts? Does the evaluation process reflect the realities of the project? To achieve these goals, the POMM is supported by two decision techniques: the Value Focused Thinking (VFT) (Keeney, 1992), and the Goal Question Metric (GQM)(Basili & Weiss, 1984). The VFT is used to help elicit and ground the values & objectives of the projects from the views of the different stakeholders, i.e. what is important to them from the context of the particular project. The GQM technique facilitates the identification of useful measures and aligns these to the identified objectives. Therefore, it can be seen that this method ensures continued collaboration with the project stakeholders to assure that their values and objectives are represented, tracked and evaluated in the project; useful and appropriate measures are identified; and the project can be evaluated based on the actual events that occurred. The POMM uses a series of iterative steps to achieve its mandate:

1. Identify key stakeholders of the project, taking into consideration the roles involved, the organisations or personnel that may be impacted by the project and its results
2. Elicit project values and objectives for each key stakeholder, including standard objectives relating to time, cost and scope
3. Apply VFT method to determine the fundamental (end) and means (facilitating) project objectives
4. Prioritize the fundamental objectives
5. Develop, review and refine (where necessary) the project means-end network that shows the relationships between the objectives
6. Apply GQM method to elicit project measures
7. Develop, review and refine (where necessary) the project objective-measure network that shows the relationships between the objectives and measures
8. Implement, monitor and take corrective actions throughout the project
9. Determine or assess the cumulative outcome of the project

3 EVALUATION METHODOLOGY

The study is influenced by the design science paradigm (March & Smith, 1995) and presents the evaluation phase of the design process (Hevner et al., 2004). Within this context, a preliminary analysis of three (3) project evaluation/measurement methods is performed to help guide the refinement and use of the alternative methods (POMM and PPS). Based on the principles of Hevner et al., (2004) a mixed approach of descriptive and analytical evaluation approaches is used in this instance. A construction of a detailed project scenario is used to demonstrate the utility of the artefacts while static analysis is used to study them for qualities such as complexity and usability. Thus a scenario-based survey was implemented among selected group of fifty-one (51) participants (business professionals and graduate students) to obtain their perception of the specific artefact in a controlled context. The participants were randomly separated into three (3) independent groups of seventeen (17) for each project evaluation method. The scenario-based survey consisted of two parts; (1) A hypothetical IS project scenario that describes a core application system implementation project. Ten (10) multiple choice questions were given to test the participants’ task performance based on the specific project evaluation method given. Documentation on the methods was also provided as it was assumed that they had no prior knowledge about any of the approaches. (2) A questionnaire with 16 questions adopted from the Maes & Poels’ (2006) conceptual model with likert-like scale of 1-5 (strongly disagree to strongly agree) was used to assess the perception of static qualities.
Surveys are practical research instruments, particularly when there is a need to move from observation to theory validation, and is useful in gathering data about individual preferences and expectations (Newsted et al., 1998). Thus, by studying a representative sample of the environment surveys seek to discover relationships that are common and facilitate the provision of generalisable statements about the phenomenon under study (Babbi 2004; Newsted et al., 1998). Further, Babbi (2004) suggested that evaluation research is undertaken for determining the impact of “some social intervention”. This information provided additional motivation to apply this research method and therefore it is used as the basis to evaluate the three (3) different project evaluation methods. Future studies will extend the survey and apply different evaluation approaches such as descriptive and experimental studies.

3.1 Test of Independent Samples

In experimental research it is often necessary to manipulate what happens to people so that casual inferences can be made (Field, 2005). Our research undertakes three experimental conditions using different sets of individuals to participate. There are several techniques available to achieve this objective through the comparison of the mean of independent groups. Suitable techniques include the Independent Samples T Test (t-test), the Mann-Whitney (M-W) test and ANOVA (Babbie, 2004; Field, 2005). The t-test and M-W test are used to evaluate the scenario-based conditions relating to an assessment of task performance, and to analyse the likert-like scaled data generated from the survey of the perception of static qualities respectively. It is noted that the techniques differ but are specifically chosen because of the types of data being collected. Moreover, the t-test on independent samples and the Mann-Whitney test have similarities in that both are testing the identity of two independent populations. The t-test compares the mean scores of two groups on a given variable, and is used when two experimental conditions and different participants are assigned to each condition (Field, 2005).

3.2 Empirical Observation Framework

3.2.1 Task Performance

To assess the three (3) methods with respect to task performance, a project scenario was used as the basis to compare the measurements methods against the performance of certain tasks. Hevner et al., (2004) posited that a useful approach that can help to demonstrate utility of an artefact is through scenario construction. Against this background the project scenario was developed in which the participants were asked, given a specific measurement method, to perform the particular set of tasks, i.e. answer the questions given based on the scenario and the measurement method given. The scenario detailed a typical contemporary event of an organisation experiencing the implementation of core application system. Within the scenario the stakeholders faced several competing objectives with varying perspectives on what is important to them, including conformance to tight deadline to implement the system, managing the budget, acceptance of the system by user, and use of application by its users. The ten (10) questions were aligned to the project process including identification of the relevant stakeholders, identifying and prioritising different types of objectives, determining suitable measures based on the scenario context, and determining what criteria would indicate successful completion of the stated project. The participants were explicitly asked to use only the information provided to answer the questions.

3.2.2 Perception of Static Quality

The Maes & Poels (2006) was adopted as the framework upon which to assess the static qualities of the measurement methods. They argued that there is a paucity of practical evaluation framework that
considers the quality of conceptual models from the user’s perspective. The Maes & Poels’ model (2006) relies on seminal IS Success models (DeLone & McLean, 1992, 2003; Seddon, 1997) which acknowledge quality as an antecedent to system success. Their model identified four interconnected constructs as necessary to help assess the quality of an artifact, which is adapted for our research context (Table 1). Perceived ease of use (PEOU) refers to “the degree to which a person believes that using a system would be free of effort” (Davis, 1989) or perceived as being difficult to use (Moore & Benbasat, 1991). Perceived semantic quality (PSQ) describes the correspondence between the information that users think the model contains and the information that users think the model should contain, based upon their knowledge of the problem domain (Krogstie et al., 1995). Thus, the users or participants can view the semantic quality of the model as how valid and complete it is with respect to (their perception of) the problem domain. Perceived usefulness (PU) relates to “the degree to which a person believes that using a particular system has enhanced his or her job performance” (Davis 1969). Within this context, the participant can evaluate the respective method based on completing certain task i.e. the project scenario activities. User satisfaction (US) is a subjective evaluation of the various consequences evaluated on a pleasant-unpleasant continuum (Seddon 1997). Against this background, the Maes & Poels’ model is suitable for this research as it applied in similar contexts, i.e. the evaluation the three IS-related measurement methods from the perspective of users. Additionally, the framework has been empirically validated as an end-user evaluation tool (Maes & Poels, 2006) and have been applied in other recent studies.

<table>
<thead>
<tr>
<th>PEOU1</th>
<th>It was easy for me to understand what the &lt;measurement method&gt; was trying to model.</th>
<th>PU1</th>
<th>Overall, I think the &lt;measurement method&gt; would be an improvement to a textual description of the project measurement process.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEOU2</td>
<td>Using the &lt;measurement method&gt; was often frustrating.</td>
<td>PU2</td>
<td>Overall, I found the &lt;measurement method&gt; useful for understanding the process modelled.</td>
</tr>
<tr>
<td>PEOU3</td>
<td>Overall, the &lt;measurement method&gt; was easy to use.</td>
<td>PU3</td>
<td>Overall, I think the &lt;measurement method&gt; improves my performance when understanding the process modelled.</td>
</tr>
<tr>
<td>PEOU4</td>
<td>Learning how to read the &lt;measurement method&gt; was easy.</td>
<td>PSQ1</td>
<td>The &lt;measurement method&gt; represents the Project measurement process correctly.</td>
</tr>
<tr>
<td>US1</td>
<td>The &lt;measurement method&gt; adequately met the information needs that I was asked to support.</td>
<td>PSQ2</td>
<td>The &lt;measurement method&gt; is a realistic representation of the project measurement process.</td>
</tr>
<tr>
<td>US2</td>
<td>The &lt;measurement method&gt; was not efficient in providing the information I needed.</td>
<td>PSQ3</td>
<td>The &lt;measurement method&gt; contains contradicting elements.</td>
</tr>
<tr>
<td>US3</td>
<td>The &lt;measurement method&gt; was effective in providing the information I needed.</td>
<td>PSQ4</td>
<td>All the elements in the &lt;measurement method&gt; are relevant for the representation of the project measurement process</td>
</tr>
<tr>
<td>US4</td>
<td>Overall, I am satisfied with the &lt;measurement method&gt; for providing the information I needed.</td>
<td>PSQ5</td>
<td>The &lt;measurement method&gt; gives a complete representation of the project measurement process</td>
</tr>
</tbody>
</table>

Table 1: Measurement Instrument for Assessing Perception of Static Qualities

3.3 Research Procedures

To accommodate the examination of the measurement artifacts, procedures were developed and made available to the participants. The following strategy was adopted for conducting the study:

1. Design survey measurement instrument and extract the project evaluation methods’ documentation. This included the development of a practical case scenario and questions on which the assessment and/or interaction with the measurement models are performed. The
second part of the evaluation includes the measurement instrument (Table 1) which is adapted from the Maes & Poels (2006) with minor modifications including assessment of a specific measurement method instead of conceptual model for the process, and project measurement process in place of a process.

2. Obtain independent review (pre-pilot) of the instrument and models documentation on the sufficiency of the variables and to confirm if the documentation is sufficiently informative and helpful in completing the scenario-based questionnaire. The scenario was later refined to improve its clarity and readability.

3. Identify and recruit pilot study participants.

4. Present each group with documentation on TCM, POMM and the PPS models accompanied with the research instrument and scenario. The instrument/questionnaire also includes the moderating variables of gender, experience and age along with other relevant data collection aids such as role and industry represented.

5. Record and analyze the responses of the study using the SPSS application

6. Refine questionnaire based on suggestions of the respondents. The documentations on the models were refined for clarity, particularly in the POMM where unnecessary areas removed (e.g. the conceptual diagram). The scenario questions were refined to improve its understandability.

7. Identify and recruit at minimum an additional 45 study participants and randomly divide them in three groups. Graduate students and experienced professionals were identified. The set of graduate students included those who are full-time and part-time students, with a large portion of them being currently employed.

8. Repeat steps 4-5.

4 RESULTS AND DISCUSSIONS

4.1 Task Performance

Our assessment of task performance was based on the average accuracy level of the participants when using the respective measurement method to interact with the scenario and complete the set of questions (i.e. tasks). We did pairwise comparisons of the three (3) methods using statistically difference of means tests. Table 2 provides a description of our results. These results suggest that with respect to task performance, PPS and POMM are attractive competitors to TCM as PPS outperforms TCM in a manner that is statistically significant, and POMM also outperforms TCM but with a difference that is not statistically significant. This suggests that participants using POMM or PPS were better able to accurately determine the most suitable project evaluation tasks including identifying complete list of stakeholders, applicable project objectives and measures and determination of success of the project. This finding coincides with objectives of the designing both PPS and POMM such as to provide practitioners with an improved alternative in performing project evaluating tasks and be better to analyse the performance of a project using multiple criteria.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Models</th>
<th>Mean</th>
<th>StdDev</th>
<th>Difference</th>
<th>Statistically Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>POMM vs TCM</td>
<td>POMM</td>
<td>6.7647</td>
<td>1.64048</td>
<td>POMM &gt; TCM</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td>TCM</td>
<td>6.2353</td>
<td>1.09141</td>
<td></td>
<td>(Signif. = 0.276)</td>
</tr>
<tr>
<td>PPS vs TCM</td>
<td>PPS</td>
<td>7.2353</td>
<td>1.67815</td>
<td>PPS &gt; TCM</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>TCM</td>
<td>6.2353</td>
<td>1.09141</td>
<td></td>
<td>(Signif. = 0.048)</td>
</tr>
<tr>
<td>POMM vs PPS</td>
<td>POMM</td>
<td>6.7647</td>
<td>1.64048</td>
<td>PPS &gt; POMM</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td>PPS</td>
<td>7.2353</td>
<td>1.67815</td>
<td></td>
<td>(Signif. = 0.414)</td>
</tr>
</tbody>
</table>
4.2 Perception of Static Qualities

The perception of static qualities is used to assess the quality dimensions identified (PEOU, PU, PSQ and US) based on the views of the participants as a result of their interactions with the requisite measurement methods. Also, pairwise comparisons of the assessment of the static qualities of the methods using non-parametric tests were also performed (Tables 3a-d). The results showed that there were no significant variations between the three methods. The average response per respondent primarily laid between 3 and 4 on the likert-like scale for all four dimensions being tested (figure 2). This means that that the respondents were fairly neutral and agreeable in relation to their perception of the quality dimensions for all methods. This has interesting implication for the newer methods in that the participants showed no significant negative reactions to them during their application. It is noted that while the results did not reflect a significantly strong positive review over and above the TCM, the opposite is also true. Thus, it can be argued that users may require more time to better understand the components of the methods, plus further simplification of the methods’ structure to facilitate ease of application may also be needed. Moreover, the results also imply that POMM and PPS may be adopted in practice with little difficulty while providing an alternative perspective in the project evaluation process.

![Comparison of Static Qualities](image)

Table 3a: PEOU: Results on Comparison of the 3 Models

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Models</th>
<th>Mean rank</th>
<th>Difference</th>
<th>Statistically Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>POMM vs TCM</td>
<td>POMM</td>
<td>19.18</td>
<td>POMM &gt; TCM</td>
<td>NO (Signif. = 0.321)</td>
</tr>
<tr>
<td></td>
<td>TCM</td>
<td>15.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPS vs TCM</td>
<td>PPS</td>
<td>17.12</td>
<td>PPS &lt; TCM</td>
<td>NO (Signif. = 0.821)</td>
</tr>
<tr>
<td></td>
<td>TCM</td>
<td>17.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POMM vs PPS</td>
<td>POMM</td>
<td>18.59</td>
<td>POMM &gt; PPS</td>
<td>NO (Signif. = 0.153)</td>
</tr>
<tr>
<td></td>
<td>PPS</td>
<td>16.41</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 3b: PUS: Results on Comparison of the 3 Models

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Models</th>
<th>Mean rank</th>
<th>Difference</th>
<th>Statistically Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>POMM vs TCM</td>
<td>POMM</td>
<td>18.94</td>
<td>POMM &gt; TCM</td>
<td>NO (Signif. = 0.390)</td>
</tr>
<tr>
<td></td>
<td>TCM</td>
<td>16.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPS vs TCM</td>
<td>PPS</td>
<td>16.71</td>
<td>PPS &lt; TCM</td>
<td>NO (Signif. = 0.636)</td>
</tr>
<tr>
<td></td>
<td>TCM</td>
<td>18.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POMM vs PPS</td>
<td>POMM</td>
<td>19.82</td>
<td>POMM &gt; PPS</td>
<td>NO (Signif. = 0.165)</td>
</tr>
<tr>
<td></td>
<td>PPS</td>
<td>15.18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 3c: PSQ: Results on Comparison of the 3 Models

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Models</th>
<th>Mean rank</th>
<th>Difference</th>
<th>Statistically Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>POMM vs TCM</td>
<td>POMM</td>
<td>19.35</td>
<td>POMM &gt; TCM</td>
<td>NO (Signif. = 0.266)</td>
</tr>
<tr>
<td></td>
<td>TCM</td>
<td>15.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPS vs TCM</td>
<td>PPS</td>
<td>19.56</td>
<td>PPS &gt; TCM</td>
<td>NO (Signif. = 0.219)</td>
</tr>
<tr>
<td></td>
<td>TCM</td>
<td>15.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POMM vs PPS</td>
<td>POMM</td>
<td>17.29</td>
<td>POMM &lt; PPS</td>
<td>NO (Signif. = 0.902)</td>
</tr>
<tr>
<td></td>
<td>PPS</td>
<td>17.71</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 3d: US: Results on Comparison of the 3 Models

#### 4.3 Implications for the 3 measurement methods

The results provide an interesting outlook on the methods under examination. The TCM continue to reveal fairly positive reviews by users despite the growing criticisms in research and practice. While both researchers and experienced professionals have suggested that the TCM does not sufficiently meet the analytical needs of current projects (Atkinson, 1999; Barclay 2008; Cohen & Graham, 2001; Nelson, 2005), the results show that it has been consistently positive in areas such as task performance, perceive ease of use, semantic qualities and user satisfaction. This may possibly help explain its continued use in practice (White & Fortune 2002) despite some identified weaknesses. This paradoxically relationship is worth additional investigations.

The newer methods POMM and PPS are also shown to suggest positive application in practice. For example, the PPS is shown to significantly perform better than the other methods with regards to task performance, i.e., evaluation activities for projects. This may be as a result of the additional dimensions (e.g. use or benefit) that can be used to help decision makers in their analysis of project contribution or performance. The POMM on the other hand facilitates better perception of static qualities in areas of usefulness, semantic qualities and perceived ease of use. This may be linked to the clear structured procedures and required stakeholder collaboration which are inherent in the approach.

While we can derive positive implications for this study, it is noted that a larger sample size may be required to draw conclusions with high(er) degree of confidence. We however acknowledge that the study underscores the need to have continued dialogue with practice and to search for improved decision tools for projects.

### 5 RESEARCH LIMITATIONS & FUTURE RESEARCH

The study provides a preliminary analysis of certain quality attributes of the measurement methods based on participants’ interaction with them. This is intended to help guide future investigations and refinement of the alternatives approaches, POMM and PPS, and improve the understanding TCM’s use in practice. Based on the resources available, a fair percentage of the participants were graduate students.

Proceedings ECIS 2009
students with relatively limited experience in IS projects although they otherwise had industry experience. Also, the time taken to learn about the artifacts (measurement models) was significantly less than what would be required in a real-world project environment. While the participants provided important insights into the perceptions of the methods, issues such as incomplete or imperfect understanding of the capabilities of the artifacts are taken into consideration. Thus, additional training may be required to facilitate a better understanding of the goals and capabilities of the PPS and POMM in particular.

We concur with the observations of Newsted et al (1998) that the survey approach is an effective research method because of its usefulness in gathering data about individual preferences and expectations, however, a key limitation is its provision of a just a snapshot of how individuals may be feeling at a particular time. This issue is taken into consideration as the participants’ state of mind may have impacted their responses and therefore the results. Consequently, further analysis preferably under real-world project settings are required to better gauge the attitudes towards them and the actual performance of these tools. Action research, field studies, case studies in diverse project, organisational and industry settings are some of the considerations for future studies.

6 CONCLUDING REMARKS

The research is intended to encourage debate on the application of different measurement methods in practice and the perception of its users. We observed that while this research area is evolving, the path to assuring that the developed artefacts satisfy the needs of the project environment for which they are developed is sometimes ignored. This is evident by the number of conceptual contributions (e.g. Atkinson 1999; Stewart 2008) and the paucity of analysis on the perception of users to the traditional paradigm for example. This study therefore attempts to redress this concern through the comparison of three (3) measurement methods including the traditional method TCM, and recently developed alternatives, the PPS and POMM. This was conducted within the ambit of utility, perceived semantic qualities (PEOU, PSQ, PUS, US) based on the perceptions of users in experimental conditions. The initial results provide encouraging results and support the value of different approaches to help users and decision makers in evaluating IS project performance.

References


Field, A. P. (2005). Discovering statistics using SPSS: (and sex, drugs and rock `n' roll) (2nd ed.): SAGE.


USABILITY OF IT-SYSTEMS IS MORE THAN INTERACTION QUALITY - THE NEED OF COMMUNICATION AND BUSINESS PROCESS CRITERIA

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0075.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Usability, Business value of IT, Evaluation, Heuristics</td>
</tr>
</tbody>
</table>
USABILITY OF IT-SYSTEMS IS MORE THAN INTERACTION QUALITY - THE NEED OF COMMUNICATION AND BUSINESS PROCESS CRITERIA

Stefan Cronholm, Linköping University, Dept of Management and Engineering, 581 83 Linköping, Sweden & University of Borås, Business and Informatics, 501 90 Borås, Sweden, stefan.cronholm@liu.se

Vince Bruno, RMIT University, School of Business Information Technology, GPO Box 2476V, Melbourne, Victoria, Australia, vince.bruno@rmit.edu.au

Abstract

The design and evaluation of IT-systems are usually supported by different usability criteria. Our hypothesis is that criteria are predominantly formulated as supporting interaction between a user and an IT-system. We are claiming that there is a need for criteria formulated at higher levels such as communication and business processes. One example of a criterion formulated at the interaction level is “Visibility of system status” and one example of a criterion formulated at the business process level is “Quality of work”. If criteria is formulated and used on the interaction level only, the impact on design and evaluation can only take place at this level. This choice will also mean that you are only able to speculate whether the IT-system is supporting higher levels. We are not saying that criteria belonging to the interaction level are unimportant; rather we are saying that there is a need for formulating complementing criteria that resides on the communication and business process level.

Keywords: Usability criteria, Usability heuristics, Design principles,
1 INTRODUCTION

There is a lot of support available for designing and evaluating IT-systems (Jones 1970; 1990; Preece et al. 1994; Shneiderman 1998). The type of support we are addressing in this paper is normally referred to as usability criteria, principles, golden rules, factors or heuristics. We will use the name criteria in this paper when we refer to any of the concepts above. A common definition of criterion reads: a standard on which a judgment or decision may be based (AskOxford 2008; Merriam-Webster 2007). When we use the word criterion in this paper we mean a basis for comparison, like a reference point against which other ‘things’ can be evaluated. The aim of usability criteria is to support the design or evaluation of IT-systems (see section 2). We define usability as “a set of attributes that bear on the effort needed for use, and on the individual assessment of such use, by a stated or implied set of users” (AS/NZS_4216 1994).

The problem we are addressing is that several of the existing criteria seem to support the interaction between a user and IT-System and not the business process. In order to support this hypothesis we present a few concrete examples that can be found in the “Ten heuristics” constructed by Nielsen & Molich (1990). The first heuristic reads “Visibility of the systems status”. The second heuristic reads: “User control and Freedom” and the third heuristic reads: “Consistency and Standards”. All these heuristics are important but they are not primarily referring to the business process. They are formulated to support the interaction between a user and an IT-System. The same observations can be made from studying the “Eight golden rules” constructed by Shneiderman (1998). One of the golden rules reads: “Enable frequent users to use shortcuts” and another golden rule reads: “Offer simple error handling”. Our claim is that usability criteria should also be formulated to support communication between users and the business process.

We are viewing IT-systems as media or instruments that are used in order to support communication between users and business actions that are performed in the business process (Goldkuhl 2005), (see section 3). Therefore the use of an IT-system per se is only a means to achieve a higher goal; to successfully manage a business action that produces value for a client (Porter, 1985). Another basis supporting this hypothesis is a famous quote from Norman (1998), “I don’t want to use a computer, I want to accomplish something”. The first part of the quote “I don’t want to use a computer” is related to the interaction between a user and IT-System. The second part “I want to accomplish something” is directly related to the business process. Norman (1998) doesn’t want to worry about the system status, IP addresses, search algorithms or database structures. Norman (1998) wants to perform business actions in order to achieve a goal of a task. The IT-system is only a tool that should support him to achieve the goal through the performance of tasks. Therefore, the tool should reside in the background and the task (the business process) should reside in the foreground. The same way of reasoning is valid for usability criteria. Besides supporting interaction, we claim that usability criteria should support the design of the IT-mediated communication between users and should support the business process.

We have identified a few criteria that support the communication between users in a business. One example reads: “Satisfy communication needs” (Cronholm and Goldkuhl 2002). The meaning of this criterion is that a user should be able to “say” what he/she wants through the IT-System. We find these criteria are important, since they are highlighting the human-to-human perspective and not the more limited user-to-computer perspective (see section 2 for a more exhaustive discussion). Thereby, the use of an IT-system can also be viewed as a social process consisting of technology mediated business communication.

We have also identified a few criteria supporting the business process. We find these criteria very important since they are oriented towards the actions or activities that exist in a business process. One example of a business oriented criterion is “Task sequencing” (see Participatory Heuristic Evaluation constructed by Muller et al., (1998)). “Task sequencing” is important since it governs the designers’ attention towards how the IT-system can useful in the business. Our view is that there is a need for complimentary criteria supporting interaction, communication and business processes.
It seems that there are at least three levels that criteria can reside on; the “interaction level”, the “communication level” and the “business process level” (see section 3). Furthermore, it also seems possible to create relations between the levels. The formulation of criteria at the “business process level” can be formulated as business goals that should be achieved. The formulation of criteria at the “communication level” can be formulated as means for achieving the criteria formulated at the “business process level”. Consequently, criteria formulated at the “interaction level” can be viewed as means for achieving the criteria formulated at the “communication level”. This hierarchy and a classification of criteria will be the focus of this research.

The relation between the three levels constitutes a hierarchy. The advantage and role of a multilevel abstraction hierarchy is discussed in Rasmussen et al. (1994). Rasmussen et al. (1994) compares a multilevel abstraction hierarchy with a means-end hierarchy and claim that a multilevel abstraction hierarchy is often used in practical problem solving processes. Furthermore, having access to several levels of abstraction is important for effective problem solving. That is, the hierarchy could be used for identifying more concrete interaction criteria that supports the fulfilment of more abstract business process criteria. Vice versa; the hierarchy could be used for searching more abstract business process criteria when that is needed. Shifts in the level of abstraction during problem solving have proved to be supportive and has been demonstrated by Wason & Johnson-Laird (1972). We are not saying that usability criteria formulated at the lower “interaction level” are unimportant; vice versa they are an important means for supporting the achievement of a higher level and to offer concrete criteria.

Research on the concept of interaction is reported within the area of Human-Computer Interaction (i.e. (Bratteteig and Stolterman 1997; Löwgren and Stolterman 1999; Preece et al. 1994). Research on the concepts of business processes is reported within the area of business processes and information systems (Davenport 1993; Hammer 1990). The growing number of criteria lists has intensified the need for practitioners to be able to understand that criteria are formulated on different levels. Our ambition is that the results will hopefully contribute to both novice and experienced practitioners; primarily as an imperative to look for criteria supporting the business process level. A second aim is to persuade criteria constructors to formulate criteria residing on all levels. This introductory section is followed by a definition of the concept of criterion (section 2). In section 3 the concepts of interaction, communication and business process are discussed. Section 4 describes our research method and in section 5 we present the findings. Finally, in section 6 we present the conclusions drawn.

2 INTERACTION, COMMUNICATION AND BUSINESS PROCESSES

The aim of this section is define the concepts of interaction and business process in order to further explain why there is a need of criteria being formulated on communication and the business process level. An interaction is defined as the interplay between a user and an IT-system (Ågerfalk 2003). To further clarify the concept of interaction with IT-system we have used the Elementary Interaction Model (EIAL) originally introduced in Ågerfalk (2003) and revised in Goldkuhl et al. (2004) (see figure 1). According to the EIAL an interaction consists of four phases: informing, execution, IT-system reaction and interpretation (see figure 1). The first phase, informing, means that a user interprets the action possibilities offered by the IT-systems and in order to reach a decision about what to do. The second phase, execution, describes that the user is performing the action chosen. The third phase, IT-system reaction, describes the IT-system’s response to the user action. Finally, the fourth phase, interpretation, means that the user is interpreting the result of the IT-system’s reaction.

In the middle of the interaction loop there is an interface containing documents with which a user interacts. A document can be a form or a web page. The screen document plays different roles in the phases, providing a multifunctional interface. In the informing phase the document is used when the user is reading the screen to figure out what to do. It contains information about the action possibilities and other conditions. In the next phase the screen document is used for execution. In this sense, the
document functions as an action medium. For example, the user enters some data in a field and clicks on a button on the screen in order to perform an action. The phase of the IT-system reaction should be understood as a response to the user execution. The IT-system’s reaction can result in changes of the screen document (as a feed-back to the user). In this sense, the document consists of action results and functions as a basis for interpretation. This interaction can be examined in more detail from the human perspective by describing the low-level capabilities that may impact the interaction, like sensory, cognitive and motor domains (Persad et al. 2007).

![Diagram](image)

Figure 1. The Elementary Interaction Loop (Goldkuhl et al. 2004)

As mentioned in section 1, IT-systems are viewed as media that support the business actions. In most businesses or social contexts there is a lot of communication. Employees are communicating with each other through IT-system and employees are communicating with clients through the IT-systems in order to achieve business goals (Goldkuhl 2005). One example of such communication is a confirmation of an order. This is a communication that takes place between an employee and a client and it is also a communication that is taking place between employees (the order agent is informing a store man about the confirmation of an order).

The theoretical basis for viewing IT-system as communication and business systems are social action theory (e.g. (Weber 1978) and language action theory (Goldkuhl and Lyytinen 1982; Habermas 1985; Searle 1969; Winograd and Flores 1986). One main point in Weber’s (1978) theory of social action is that communication is intentional. Using a social action perspective means that it is not acceptable to view IT-systems as a black box with some social and organizational consequences (Dietz 2001). IT-systems should therefore be perceived as systems for business action. The language action theory conceives communication as one type of action. Communication is not restricted to a mere transfer of information. To communicate is to establish interpersonal relationships between the sender and the receiver (Searle 1969). In a language action perspective, IT-systems are not considered as “containers of facts” or “instruments for information transmission” (Goldkuhl and Ågerfalk 2002). This perspective emphasizes what users do while communicating through an IT-system. IT-systems are thereby socio-technical systems for action in business and such actions are the means by which business relations are created.

The need of communication between employees and between employees and clients can be supported by IT-systems. The use of IT-systems is not an end in itself. The use of IT-systems is instrumental in relation to other aims and actions. The aim of using IT-systems in a business context (or a public context) is to support the communication that is taking place when performing business actions. The communication is important since it contributes to fulfilling business goals. The communication level can in this respect be considered as means for the level of business action. Business communication is therefore part of business processes. The aim of a business process is to produce product(s) and/or service(s) in order to satisfy the need of a business client and business goals.
The concepts of interaction, communication and business process are depicted as three sets or three levels (see figure 2). As pictured, the interaction is the interplay between a user and IT-System and the communication is something that occurs between two users; mediated by an IT-system. The arrow in the picture symbolizes the business; a business that produces something (a product or a service) for a client. The communication level is viewed as a subset of the business process level and the interaction level is viewed as a subset of the communication level. That means that if the IT-system is considered as providing good support for the business level it consequently also provides good support for the lower communication level and interaction level. On the other hand, if an IT-system is evaluated at the interaction level no predictions can be made of the higher communication level and business process level. Our claim is that there is a need to formulate a balanced set of criteria at all the three levels discussed, not only at the interaction level. We are not saying that criteria formulated at the interaction level are useless; ideally we would like to see criteria residing at all the three different levels and that they are coherent and complementary.

![Diagram of interaction, communication, and business process levels](image_url)

*Figure 2. Relations between interaction, communication and business process*

## 3 RESEARCH METHOD

Our hypothesis reads that there is a need for formulating criteria on higher levels than the interaction level. To verify our hypothesis we need to establish that existing criteria lists mainly reside on the interaction level. Looking closer into existing criteria lists it is obvious that some of them can be characterized as general while others are more context specific. In order to choose lists that represent a variation we have used the following criteria of selection: general/context-specific, variation in perspectives, thoroughness, easy accessible and familiar to practitioners.

The first selection criterion, general and context-specific, means that we have chosen criteria list that represents both general and context-specific criteria (see appendix). The need for using context specific criteria is touched upon, in a study performed by Beck et al., (2003). Results from a literature survey indicate that 44 of 114 papers in the area of mobile human-computer interaction (small-screen devices) utilized traditional usability evaluation techniques such as heuristic evaluation (Nielsen 1993; Nielsen and Molich 1990) and that only six of the papers employed new methods in realistic mobile use situations. A mobile phone or a PDA, provide a significant increase in supporting communication of a large range of business actions. Examples communication and business actions supported by small screen devices are buying and delivering of tickets for travelling or tickets for entertainment.

The second criterion, a variation in perspectives, means that different perspectives should be represented such as business, communication, technical and user. The third criterion, thoroughness, means that the criteria lists should have a broad coverage. A broad coverage is condition since the

Proceedings ECIS 2009
designer (or evaluator) is focused on finding as many ‘problems’ as possible and wants the results to be as complete as possible (Sears 1997). The lists of criteria chosen must also be accessible to us, that is, they should be well described and easy to understand. Finally, as far as possible we have chosen lists of criteria that are well known.

We have also selected a research oriented, or an “academic”, criteria list that is not well known to practitioners. The reason for choosing this list is that it is based on a communication perspective and it has good thoroughness. The idea of incorporating this list is to understand if and how criteria generated from a communication perspective will relate to the more familiar criteria lists. Our final selection consists of six different criteria lists. All criteria lists (except the “academic” list) meet the selection criteria. Two of the lists can be characterized as being of a general character; these lists are the 10 usability heuristics (Nielsen 1993) and the Eight Golden Rules (Shneiderman 1998). Two of the lists represent the context of small-screen devices; the chosen lists are design guidelines for small screen devices (Kärkkäinen and Laarni 2002) and context-aware mobile applications (Häkkilä and Mäntyjärvi 2006). The final two lists represent the communication and business context. The chosen lists are Participatory Heuristic Evaluation (Muller et al. 1998) and the “academic” criteria list Actability Principles (Cronholm and Goldkuhl 2002).

The model presented in figure 2 has served as an analysis model (see section 2). Two coders (the authors) have classified the formulation of each criterion as belonging to the interaction level, the communication level or to the business process level. Both coders have been involved with the concept of usability for the past ten-year, in a pedagogical and research perspective. Each coder was provided with the criteria lists, which included the description text published by each criteria list author. The coder performed the categorisation of each criterion in the criteria list individual (separately). The coders identified that some of the criteria are formulated in a way that they can belong to more than one level. In these cases the coders needed to make a judgement for a primary classification based on the emphasis placed on the criteria’s description. Through classifying criteria an understanding of the scope and limitations of each criteria list has been achieved. The result of the classification of the criteria is based on the interpretation of the definitions and clarifications of the criteria presented in the literature. The results of the individual coding were then compared using the Cohen Kappa coefficient to ascertain the level of agreement between coders (Cohen 1960). This measure provides an easy calculation that test whether agreement between coders exceeds the chance levels of the coder classification. The calculation of classification of the criteria shows the level of agreement between coders (see section 5). The translation of the Kappa Statistics into strength of agreement is described by Landis and Koch (1977).

Once agreement levels were ascertained the similarities and discrepancies in the interpretation between coders of the criteria have been discussed. Where a disagreement existed, the criterion description was examined. The classification was then decided on the wording of the criterion and related to the level with the strongest relation. This process is not dissimilar to the conducting of a heuristic evaluation (Nielsen and Molich 1990), where inspectors examine an interface individually on two separate occasions and then come together to discuss their findings. In this way, both coders have ultimately agreed upon the final classifications discussed in the findings.

4 FINDINGS

The findings are presented according to the different contexts of the criteria lists chosen. First, we present the findings for criteria lists categorized as being of a general character (i.e. 10 usability heuristics (Nielsen 1993) and the Eight Golden Rules (Shneiderman 1998)). Second, we will present findings concerning the small screen devices (i.e. design guidelines for small screen devices (Kärkkäinen and Laarni 2002) and context-aware mobile applications (Häkkilä and Mäntyjärvi 2006)). Third, we will present our findings concerning the communication and business oriented criteria lists (i.e. Participatory Heuristic Evaluation (Muller et al. 1998)) and Actability Principles (Cronholm and Goldkuhl 2002)). Finally we will present a summary of the classifications. For each context the
criteria classification is made according to one of the levels interactive, communicative or business process. All the criteria lists are presented in the appendix.

4.1 Classification of criteria lists of general character

The findings show that most criteria belong to the interaction level, a few to the communication level and none to the business process level (see Table 1). This means that there is a heavy weighting towards the interaction level. The level of agreement between authors for this context is a Kappa value, $\kappa = -0.1$, that shows a poor agreement level. This disagreement was mainly due to the different interpretations of the wording of the description for criteria. The proportion of classifications made where authors agreed was $P_o = 0.6$. The main contention from the classification was with the interaction and communicative level.

Examples of criteria proposed by Nielsen (1993) that are classified as belonging to the interaction level are “Visibility of system status”, “Match between system and the real world”, “Consistency and standards” and “Aesthetic and minimalist design”. Another example proposed by Shneiderman (1998) reads: “Offer informative feedback”. The reason for classifying them at the interaction level is that they are all primarily referring to the interplay between a user and an IT-system. Examples of criteria belonging to the communication level are “Flexibility and efficiency of use” (Nielsen 1993) and “Design dialog to yield closure” (Shneiderman 1998). The reason for classifying these at the communication level is that they support the communicative intent of an action.

<table>
<thead>
<tr>
<th>Name of criteria list</th>
<th>Number of Criteria</th>
<th>Interaction level</th>
<th>Communication level</th>
<th>Business Process level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heuristics</td>
<td>10</td>
<td>7</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Eight Golden Rules</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1: Classification of criteria lists of general character

4.2 Classification of criteria lists of small screen devices

The findings concerning the criteria lists of small screen devices show that most criteria belong to the interaction level, a considerable number to the communication level and one to the business process level (Table 2). It is clear the concentration of criteria is predominantly found in the interaction level. It is interesting that the classification of the criteria list “context aware mobile applications” resulted in five of the criteria being classified as belonging to the interaction level and four of the criteria as belonging to the communication level. The level of author agreement for this context was a Kappa value, $\kappa = 0.4$, that is considered a fair level of agreement between the authors. The proportion of classifications made where authors agreed was $P_o = 0.7$. The disagreements in classifications were spread across all levels.

Examples of criteria proposed by Kärkkäinen and Laarni (2002) that are classified as belonging to the interaction level are “Present the most important information first at the top of the hierarchy”, “Re-think the navigational aids”, “Indicate the links clearly” and “Optimize the reading process”. Another example proposed by Häkkilä and Mäntyjärvi (2006) is “Avoid information overflow”. Clearly, the primarily aim of these criteria is to support interaction within the context of small screen devices.

Examples of criteria proposed by Häkkilä and Mäntyjärvi (2006) that are classified as belonging to the communication level are “Consider the uncertainty in decision-making situations” and “Prevention from interruption”. These criteria support the communicative intent of the performance of a task. These criteria are especially important for small screen devices because the mobility of the device can find the user in various situations that can have an impact on the communicative intent of IT-System (mobile device). Examples of criteria belonging to the business process level are “Determine the
The purpose of the site / service” (Kärkkäinen and Laarni 2002) and “Usefulness” (Häkkilä and Mäntyjärvi 2006). The argument for classifying these criteria at the business process level is that they are primarily referring to purpose and usefulness to the business.

<table>
<thead>
<tr>
<th>Name of criteria list</th>
<th>Number of Criteria</th>
<th>Interaction level</th>
<th>Communication level</th>
<th>Business Process level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Screen Devices</td>
<td>8</td>
<td>6</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Context aware mobile apps</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2: Classification of criteria lists of small screen devices

### 4.3 Classification of criteria lists of communication and business process

The findings concerning the criteria lists of communication and business processes show that most criteria belong to the interaction level, a fair number to the communication level and a fair number to the business level (Table 3). Some interesting differences are that the actability principles are predominant at the interaction level and the communication level while the participatory heuristics evaluation criteria are predominant at the business process level (the +5 in the column “number of criteria” refers to the extension of the 10 heuristics (Nielsen 1993) and number between brackets refers to the 10 heuristics). The level of agreement by the authors for this context was a Kappa value, $\kappa = 0.9$, that is considered an almost perfect level of agreement. The proportion of classifications made where authors agreed was $P_o = 0.9$. The one difference was the criteria “Understand used concepts (familiar and understandable vocabulary)” (Cronholm and Goldkuhl 2002), which was classified within the interaction level by one author and communication level by the other author. This difference was discussed placed primarily in the interaction level based on the emphasis placed in the wording of the criteria’s description.

<table>
<thead>
<tr>
<th>Name of criteria list</th>
<th>Number of Criteria</th>
<th>Interaction level</th>
<th>Communication level</th>
<th>Business Process level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actability Principles</td>
<td>10</td>
<td>6</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Participatory Heuristic Evaluation</td>
<td>+5 (10)</td>
<td>1 (7)</td>
<td>0 (3)</td>
<td>4 (0)</td>
</tr>
</tbody>
</table>

Table 3: Classification of criteria lists of communication and business process

Examples of criteria proposed by Cronholm & Goldkuhl (2002, 2005) that are classified as belonging to the interaction level are “Clear action repertoire”, “easy to navigate and “Familiar and understandable vocabulary”. Another example is “Skills” (Muller et al., 1998). The aim of these criteria is to support interaction. Examples of criteria proposed by Cronholm & Goldkuhl (2002, 2005) that are classified as belonging to the communication level are “Satisfy communication needs, “Understand the communicative intent of different messages” and “Offer good support for business actions”. All these three examples are primarily aimed at supporting the communication between users. Examples of criteria proposed by Muller et al. (1998) that are that are classified as belonging to the business process level are “Task Sequencing” and “Quality of work”. The reason for viewing these criteria as belonging to the business process level is that they refer to the business tasks or work. They are aiming at providing value for a client.

### 4.4 Summary of Classifications

The findings concerning all the criteria lists show that a huge majority of the criteria is classified as belonging to the interaction level, some to the communication level and a few to the business level
Five of the six analyzed criteria lists consist of a majority of interactive criteria. It is not surprising that the business process orientated contexts embrace more criteria that are characterized as related to the communicative or business process dimension. What is notable is that both the contexts of general and small screen devices have a strong focus on interaction. The analysis shows that the communicative and business process levels have been neglected in several of criteria lists. An interesting observation is that even criteria lists characterized as being communication and business oriented, embrace criteria that mostly belong to the interaction level.

In summary, for all the criteria lists analysed, the proportion of classification made where authors agreed, $P_o = 0.78$, and the proportion of categorisations made where agreement is expected by chance, $P_c = 0.51$. This describes a very high level of agreement on classifications between authors, based on $P_o$. The coefficient of agreement, Cohen’s (1960) Kappa, for all of the criteria lists classified, $\kappa = 0.56$, is a moderate level of agreement between authors, as describe by Landis and Koch (1977), see figure 3. This provides an excellent starting point for the two authors to come to a consensus in the classification of criteria. The heavy weighting towards the interaction level has also been identified by Sears (1997). Sears (1997) discusses this gap from an evaluation technique perspective and proposes a combination of a heuristic evaluation (Nielsen 1993) with a cognitive walkthrough (Nielsen and Mack 1994) in order to achieve a better usability outcome. Our interpretation of the research by Sears is that he attempts to compensate for the gap between the interaction level and the business level. All the lists contain gaps in one or two of the proposed levels. It seems that the inventors of the lists have applied, consciously or not, a specific perspective while constructing the lists. The best coverage of all the levels can be found in Participatory Heuristic Evaluation (Muller et al. 1998).

<table>
<thead>
<tr>
<th>Context</th>
<th>Set of criteria</th>
<th>No of criteria</th>
<th>IL</th>
<th>CL</th>
<th>BPL</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>Nielsen’s Heuristics</td>
<td>10</td>
<td>7</td>
<td>3</td>
<td>0</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td>Eight Golden Rules</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Small Screen Device</td>
<td>Small Screen Device</td>
<td>8</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>Context Aware Mobile Devices</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Business Orientated</td>
<td>Actability Principles</td>
<td>10</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>Participatory Heuristic Evaluation</td>
<td>+5 (10)</td>
<td>1 (7)</td>
<td>0 (3)</td>
<td>4 (0)</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Table 4: Summary of all criteria classification

5 CONCLUSIONS

Our hypothesis reads that there is a need for formulating criteria at higher levels than the interaction level. The analysis has shown that criteria: mainly support the interaction between a user and an IT-System, are not formulated to support a human-to-human communication (mediated by an IT-System) and are not formulated to support the higher level of business processes. The definition of usability (see section 1) puts the onus on the practitioner to define the “set of attributes” that defines the usability goals for a project. The aim of the model presented in section 2 is to support the practitioner to explicitly define and select criteria at different levels.

Before drawing conclusions we would like to say that criteria are not something that should be used in a mechanical way. You should always be open to other problems and strengths that are not covered by the criteria used. But, using criteria means to intentional focusing on certain aspects that correspond to the criteria selected. That’s why a selection is important. If criteria formulated on the interaction level only are used, design and evaluation can only take place at this level. This choice will also mean that you are only able to speculate whether the IT-system is supporting the higher levels, i.e. the communication level and the business process level. If a designer or evaluator is using criteria belonging only to the interaction level, there can be a risk of unconsciously overlooking other
important aspects of the IT-systems. We are not saying that criteria belonging to the interaction level are unimportant, rather we are saying that there is a need for formulating complementing criteria residing on the communication and business process level.

In order to get a broad coverage of different criteria lists, we selected three different contexts. The importance of identifying different contexts is acknowledged by Henninger et al. (1995). They claim that “If the potential of usability guidelines as an interface design technique is to be fully realized, they need to be augmented with context-specific guidelines and examples that synthesize isolated guidelines into domain-specific solutions to design problems”. One conclusion is that the contexts do not differ that much when looking at each list’s relative distribution of interactive, communicative and business process oriented criteria. Most of the criteria, considering all the contexts, are formulated as interaction criteria. Otherwise, there are no major differences between the lists, except that the lists “Context Aware Mobile Devices” (Häkkilä and Mäntyjärvi 2006) and “Actability principles” (Cronholm and Goldkuhl 2002) that tend to lean more towards the communicative level.

Our conclusion that most criteria are oriented towards interaction is supported by Hartson et al. (2001) who are claiming “Usability is seated in the interaction design”. Another observation that supports our conclusion is the criteria list constructed by Muller et al. (1998). We interpret the criteria added by Muller et al. (1998) to Nielsen’s 10 heuristics (1993) as a reaction towards Nielsen’s formulations. The five added criteria by Muller et al. (1998) are mostly formulated at the business process level. Table 4 clearly shows that the work done by Muller et al. (1998), has identified the imbalance in the original criteria, proposed by Nielsen (1993). Hornbæk & Erik Frøkjær (2008) support the importance of using criteria residing on higher levels. They claim that problems discovered will have higher business relevance.

The result is based on the analysis of six usability criteria lists and the results have to some extent been compared to existing theory. Our opinion is that the direction of the results works as a good basis for further elaborations. There is a need of further elaborations considering other criteria lists and other types of IT-systems. As future research we also propose a deepened elaboration on the relations between the three levels described.

References


AS/NZS_4216 "Information technology—Software product evaluation—Quality characteristics and guidelines for their use," 0 7262 9071 8, Australian/New Zealand Standard, Homebush NSW 2140 Australia, Wellington 6001 New Zealand, p. 20.


Appendix

**Nielsen’s (1993) Heuristics**
- Visibility of system status
- Match between system and the real world
- User control and freedom
- Consistency and standards
- Error prevention
- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design
- Help users recognize, diagnose, and recover from errors
- Help and documentation

**Shneiderman’s (1998) Eight Golden Rules**
- Strive for consistency.
- Enable frequent users to use shortcuts.
- Offer informative feedback.
- Design dialog to yield closure.
- Offer simple error handling.
- Permit easy reversal of actions.
- Support internal locus of control.
- Reduce short-term memory load.

**Häkkilä & Mäntyjärvi’s (2006) Context Aware Mobile Application Guidelines**
- Consider the uncertainty in decision-making situations.
- Prevention from interruptions.
- Personalization.
- Avoid information overflow.
- Secure the user’s privacy.
- Remember mobility.
- Secure the user control.
- Access to context.
- Visibility of system status.
- Usefulness.

**Kärkkäinen & Laarni’s (2002) Small Screen Device Guidelines**
- Determine the purpose of the site / service

Re-evaluate the interface metaphors
Present the most important information first at the top of the hierarchy
Re-think the navigation aids
Indicate the links clearly
Optimize the reading process
Use markers while scrolling or paging text
Use pictures with caution

**Cronholm & Goldkuhl’s (2002; 2005) Actability Principles**
- Easy to understands what can be done with the system (clear action repertoire)
- Able to “say” what he/she wants through the system (satisfy communication needs)
- Can easily move to another document (easy to navigate)
- Understand consequences of proposed and performed actions (action transparency)
- Can immediately see if the intended action is executed (clear feedback)
- Can easily access information of what has been done previously (easy access to action memory)
- Know who has said what (personalized information)
- Understand used concepts (familiar and understandable vocabulary)
- Understand the communicative intention of different messages
- Offer a good support for business actions

**Muller et. al.’s (1998) Participatory Heuristic Evaluation (the added five criteria to Nielsen’s 10 heuristics)**
- Task Sequencing
- Skills
- Pleasurable and Respectful Interaction with the User.
- Quality Work
- Privacy

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0146.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Grounded Theory, Business Process Management, Research methodology, Theory Building</td>
</tr>
</tbody>
</table>
USING GROUNDED THEORY FOR STUDYING BUSINESS PROCESS MANAGEMENT PHENOMENA

Seidel, Stefan, University of Liechtenstein, Fuerst-Franz-Josef-Strasse 21, 9490 Vaduz, Principality of Liechtenstein, stefan.seidel@hochschule.li
Recker, Jan, Queensland University of Technology, Brisbane QLD 4000, Australia, j.recker@qut.edu.au

Abstract

Grounded theory has emerged as a promising research approach for generating theory in emerging and relevant domains of IS research. In this paper we advance the use of grounded theory in one important and under-researched IS domain, business process management, which is characterized by an abundance of conceptual work with very little thorough theory development. We provide an illustrative case study that shows how grounded theory allowed us to develop theories about creative organizational processes, and the influence of creativity requirements on the management of such processes. Based on the experiences gathered, we then suggest a set of key issues that highlight why and when grounded theory can be particularly suitable for the study of phenomena associated with the business process management domain. We also provide examples for other inquiries that could benefit from the use of grounded theory.

Keywords: Grounded Theory Method, Business Process Management, Research Methods, Qualitative Research
1 INTRODUCTION

Development of theory is key to Information Systems research and to any scientific endeavour pertaining to the development of a deep and meaningful understanding of certain real-world phenomena that we observe in day-to-day life. Theory development can take many forms, from analysing available literature and extending existing theoretical models, to using common sense or experience (Eisenhardt, 1989). Over the last decades, other forms of theory development have also emerged, using case study research (Eisenhardt, 1989) or the grounded theory method (GTM) (Glaser & Strauss, 1967; Strauss & Corbin, 1998).

These types of qualitative research have gained increased popularity in IS research (e.g., Urquhart, 2001; Fernández et al., 2002), suggesting that grounded theory is becoming a well-accepted research method in IS research. The grounded theory approach essentially attempts to explore for, and develop, generalised formulations about the general features of a phenomenon while simultaneously grounding the account in empirical observations or data (Martin & Turner, 1986). One of the key advantages – and challenges – of the grounded theory approach is that it is applicable to research domains that are characterized by their emergence and lack of substantive theory.

One such domain is the area of business process management (BPM), a recently emerging key area of IS research that is characterized through high relevance to practitioners (Gartner Group, 2009) but also the absence of rigorous or mature theoretical foundations (Rosemann et al., 2006). It is not the case that no empirical work or theory development has been carried out. Known studies report on the use of Delphi studies (Indulska et al., 2009), case studies (Bandara et al., 2005), focus groups (Radulescu et al., 2006), experiments (Mendling et al., 2009) or surveys (Recker, 2008) in their endeavours of exploring the BPM domain for generalizable insights and testable propositions. Continuing alongside the work of our colleagues, in this paper, we advance the Grounded Theory Method as an alternative research method for studying phenomena typically associated with business process management.

Over the last years, we have conducted a grounded theory study in order to investigate organizational creative processes (e.g., Seidel et al., 2008; Seidel, 2009). In this paper, we use examples from this study to illustrate the benefits and challenges in applying GTM to business process management research. Accordingly, our paper seeks to contribute to the IS body of knowledge by discussing the use of GTM in studying BPM phenomena. In doing so, we present a set of issues that are grounded in findings we have made conducting a BPM study where GTM was used in order to analyse data and generate theory.

We proceed as follows. The next section provides a brief overview of the use of GTM in IS research as well as a brief introduction into the area of BPM including a review of related BPM research. We then advance GTM as an appropriate research method for BPM. This is followed in Section 3 by the description of a GTM study on organizational creative processes. Then, we discuss in Section 4 issues that emerge when studying BPM phenomena through the lens of GTM. The paper concludes with a discussion of limitations and contributions and provides an outlook to future research endeavours.

2 BACKGROUND

2.1 Grounded theory

Over recent years, the grounded theory method has increased as a popular research approach employed in a number of IS studies (e.g., Orlowski, 1993; Urquhart, 1998; Trauth, 2000; Fernández, 2003). Furthermore, several papers discuss issues around the usage of GTM in order to study IS phenomena (Urquhart, 2001; Fernández et al., 2002; Urquhart & Fernández, 2006). Usually, these
studies focus on some kind of IT-related processes. Urquhart (1998), for example, investigated processes of early requirements gathering, while Orlikowski (1993) studied the adaptation and use of CASE tools over time. The nature of the phenomena that were studied highlight that GTM can be used in order to gain contextualized, rich description of certain phenomena.

Grounded theory, in general, is seen as a method for building theory, not verifying it (Strauss & Corbin, 1998). The approach is based on two main principles. First, the process of theory-building is highly iterative, during which theory and data are constantly compared. This process can be referred to as comparative analysis. Second, GTM builds upon theoretical sampling as a process of data collection and analysis that is driven by concepts that emerge from the study and appear to be of relevance to the nascent theory (Strauss & Corbin, 1998).

Two main streams of grounded theory can be distinguished. Grounded theory was first introduced as a general qualitative research method by Glaser & Strauss (1967). Strauss (1987) and later Strauss & Corbin (1998) then revised this general understanding and provided detailed procedures and techniques for the analysis of data. In IS research, both types of approaches have been used in studies, those that are rather “Glasarian” (e.g., Fernández, 2003) and those that are rather “Straussian” (e.g., Urquhart, 1998). Urquhart (2001) notes that whether the more rigid approach suggested by Strauss (1987) works may depend on the nature of the phenomenon studied.

In the study conducted we applied the paradigm suggested by Strauss & Corbin (1998) – with some adoptions, which will not be discussed in detail in the interest of brevity – as their approach promised to lead to meaningful results with regard to the particular BPM phenomenon we were investigating. The Straussian approach is conducive to studying processes because “by answering the questions of who, when, where, why, how, and with what consequences, analysts are able to relate structure with process” (Strauss & Corbin, 1998 p 127). Straussian GTM studies typically follow a basic three stage procedure:

- **Open Coding:** In general, open coding is a process of identifying a set of themes or categories that appear to be relevant in order to describe and explain the phenomenon under investigation. Such category is a conceptual element of a theory – an abstract representation of something identified through the data as being significant. Moreover, the researcher identifies properties and dimensions of categories.

- **Axial Coding:** In axial coding, relationships between categories are identified. To do so, Strauss & Corbin (1998) distinguish different types of sub-categories, which they label conditions, actions/interactions, and consequences. By grouping categories, the researcher starts to link structure with process. This involves taking the categories that emerged during open coding and reassembling them with propositions about their relationships. These emerging propositions form a theoretical framework.

- **Selective Coding:** In selective coding the core category, the central phenomenon of the research, is identified and other concepts are related to the core category. Thereby, the nascent theory is integrated by linking the different categories to the central research phenomenon.

In summary, grounded theory is thought to ensure that nascent theory is informed by the data; it is not preconceived or forced upon the data but rather emerges from it. However, it has been argued in the literature that grounded theory studies do not completely omit any existent theory. Urquhart and Fernandez (2006), for instance, call this misconception the myth of the “researcher as a blank slate.” Strauss and Corbin (1998) suggest that researchers may consult literature or experience in order to uncover examples of similar phenomena. When doing so, researchers should be self-reflective, so as to be aware of biases that may result from their own background knowledge. In the study we discuss in this paper, this suggestion encouraged us to constantly ‘step back from the data’, and to question ourselves whether concepts and relationships were actually grounded in the data, or were imposed by us through preconceived knowledge.
2.2 Research on BPM phenomena

Business process management (BPM) is a holistic management practice that (re-) emerged during the 1990s based on work on business process re-engineering (Hammer, 1990), total quality management (Powell, 1995) and process innovation (Davenport, 1993). BPM denotes a holistic, complex practice to organizational management that builds upon the notion of a business process, a series of tasks or activities that need to be carried out in order to collectively realise an organisational objective or policy goal, and a set of conditions that determines the order of the tasks (Hammer, 1990). These processes span an organization from end-to-end and include the involvement of organizational as well as IT resources, policies, regulations, contextual environments, risk factors and other phenomena. In its attempt to manage these processes effectively and efficiently, BPM addresses various aspects of strategic alignment, governance, methods, IT, people and culture (Rosemann et al., 2006) – indicating not only the complexity of the problem but also the complexity of the management approach.

The complexity and richness of business process management makes it certainly attractive as a research domain. And indeed, recent years have seen an increasing number of scholarly works examining various, discrete phenomena and artifacts pertaining to BPM. Much work, for instance, has been carried out in areas of process modeling artifacts (Rosemann et al., 2008), the development of process-aware information systems (van der Aalst et al., 2007), or approaches towards process optimization (Reijers & Mansar, 2005). This work has in common that it is mostly conceptual or analytical in nature, and considers discretely (i.e., in isolation) a particular BPM artifact, stripping away the contextual environment in which such artifacts are being put to use in practice.

This situation indicates a lack of substantive or holistic theory on the management of processes, the lifecycle of related phenomena (such as adoption or use of tools or methods), suitable governance mechanisms, decision making strategies and other related aspects of interest (de Bruin, 2007). While the lack of foundational theory denotes a challenge to academia and industry, it also denotes a window of opportunity for IS scholars to contribute to theory development in this emerging and relevant area of IS research. And indeed, some exploratory work has been carried out. For example, on the basis of case studies, Bandara et al. (2005) developed a theoretical model of the critical success factors of process modeling. Rosemann et al. (2006) use a multiple case study approach together with a Delphi study to develop a theory of BPM progression in organizations; and Recker (2008) used semi-structured interviews and the survey approach to building and testing a theoretical model of process modeling standard usage.

These studies are stimulating examples for theory development work, using ‘typical’ research methods prevalent in Information Systems. Continuing this work, we now advance the grounded theory method as an alternative, complementary or additional research method that we deem highly applicable to the BPM research domain.

2.3 On the Applicability of GTM to the Challenge of Studying BPM Phenomena

Business process management is concerned with the management of complex social, organizational processes that involve various stakeholders as well as other organizational or IT-based resources, policies or artifacts (Rosemann et al., 2006). We argue that the BPM domain, due to the complex and contextualized nature of its associated phenomena, is a primary application area of the GTM method. More specifically, we advance that GTM as an inductive, data-driven method can provide new insights into BPM phenomena.

This argument rests on three key observations. First, business processes are highly dependent on, and influenced by, the contextual environment in which they are embedded (Rosemann et al., 2008). Grounded theory aims at generating substantive theory and places special emphasis on the premise that contextual complexities and particularities need to be incorporated into an understanding of a particular phenomenon (Orlikowski, 1993). Thus, GTM as an inductive approach to generating theory...
can help researchers to gain an in-depth understanding of business processes and themes related to these.

Second, BPM research focuses on practical issues and, thus, emphasizes practical relevance. Similarly, GTM studies are suggested to generate theories that are not only rigorously developed but also provide useful insights for practitioners. Glaser (1978) argued that GTM has the potential to generate theory of high practical relevance. His arguments include the capacity of grounded theories to support practitioners in providing additional aspects that were not empirically known to them as well as increasing the experts’ ability to deal with new situations (Fernández et al., 2002).

Third, process management philosophies have at their core the notion of a business process as a collection of activities that takes one or more kinds of input and creates an output that is of value to the customer (Hammer, 1990). Such processes are dynamic in nature and range from highly structured, to highly unstructured, even artistic processes (Hall & Johnson, 2009). Grounded theory is seen as being particularly fitting to study such types of processes because it facilitates “the generation of theories of process, sequence, and change pertaining to organizations, positions, and social interaction”(Glaser & Strauss, 1967 p 114). Orlikowski (1993) frames GTM as a research approach that specifically includes elements of process and change, thereby implying a close connection to the process phenomenon that builds the core of BPM.

In the following section we provide further details about the applicability of GTM to the study of BPM phenomena, by describing our experiences gathered in a GTM study of organizational creative processes. We complement our experiences with the consideration of extant literature on the use of GTM in IS research.

3 A STUDY OF ORGANIZATIONAL CREATIVE PROCESSES

3.1 Study objectives

The research aimed at studying business processes that are characterized by creativity. Creativity is a process, carried out in complex social settings, that leads to creative products that are both novel and purposeful (Woodman et al., 1993). While a number of theories exist that describe the occurrence of creativity, creative behavior and creative products (e.g., Rubenson & Runco, 1992), they are only of limited value when considering a process perspective. This is because, typically, business processes that contain creativity are constituted by both transactional, well-structured parts as well as creative parts where the first and the latter are closely interrelated. Thus, organizations are required to gain a holistic understanding of such end-to-end business processes; they need to understand what factors shape the process and what process-related strategies and IT systems can be applied in order to deal with particular characteristics, such as particular risks being caused by high demands for flexibility and high uncertainty with regard to the creative output. Understanding the phenomenon of creativity at a process level is imperative to the development and adaptation of information systems artifacts, as well as of organizational change within creative organizations, so as to ultimately enhance organizational effectiveness without compromising creativity. Besides this high practical relevance, such a study also exhibits an academic challenge, as it can contribute meaningfully to the IS body of knowledge by providing substantive theory that can then be used as a starting point to conduct research in other substantive areas in order to proceed to more general theory in the field of BPM. Consequently, the themes that were identified within the contextualization stage of this research were business processes that contain creativity, and the question of how these processes can be managed. This could be expressed in terms of a quite general research problem. Correspondingly,

**Research problem:** How does creativity influence business processes and business process management?
Note that this rather broad research problem was further developed into a set of more specific research questions which evolved during the course of this study. Specific research questions aimed at the identification of contextual factors, strategies and information technology that organizations apply in order to manage their processes, as well as intended and unintended consequences. In the following we demonstrate how the different stages of the approach (as described in Section 2.3) were utilized in order to study a particular BPM phenomenon. We also provide reflections hinting at different issues that came up throughout the study.

### 3.2 Data Collection

Data was collected from three Australian organizations from the field of film and visual effects production. Processes in this industry can be described as highly dependent on creativity, interdependent, intensively involving the client, complex, but also repetitive (Seidel, 2009). The products, such as visual effects sequences, are highly creative as they are characterized by both novelty and purposefulness.

Glaser and Strauss (1967) encourage researchers to include multiple data collection techniques, to allow the researcher to consider multiple viewpoints from which an emerging concept can be analyzed, substantiated, and its properties be developed. Our data collection strategy involved semi-structured interviews, process modeling and document analysis. More than 30 individuals were interviewed over two years. Process models provided the researchers with an advanced understanding of the nature of the processes being studied. This knowledge was then used in order to conduct further semi-structured interviews and theoretical sampling. For instance, the process models allowed the researcher to ask more focused questions with regard to IT systems, involved actors, or required resources. The analysis of documents, such as press releases, helped to further contextualize this research. Due to space limitations, in the present paper we limit our illustrations to data that was collected in semi-structured interviews.

### 3.3 Open Coding

Open coding started with a process referred to as microanalysis (Strauss & Corbin, 1998). That is, a line-by-line analysis of semi-structured interviews to identify initial codes was carried out. This enabled comparisons within the same group as well as comparisons with different groups (i.e. the case organizations). A close examination of data to identify both differences and similarities allows for fine discrimination and differentiation among categories” (Strauss & Corbin, 1998 p 102). Table 1 provides some exemplary open codes.

<table>
<thead>
<tr>
<th>Quote</th>
<th>Open Codes / Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>I think, managing expectations [managing expectations]. And the role that the producer has in terms of delivering to a client [delivering to a client], because that would often be rare that along the line the client is directly involved. So, it's not that it's rare, but [...] the producer is much more involved with the project every single day and they are reporting obviously outside to a client [reporting to a client]. And when the client comes in [client touch points] that's the moments when you are getting direct feedback from the clients [getting feedback from clients]. But, the role the producer has in managing those expectations [managing expectations] I think is really important.</td>
<td>Reporting to a client</td>
</tr>
<tr>
<td></td>
<td>Getting feedback from clients</td>
</tr>
<tr>
<td></td>
<td>Managing expectations</td>
</tr>
<tr>
<td></td>
<td>Delivering to a client</td>
</tr>
<tr>
<td></td>
<td>Client touch points</td>
</tr>
</tbody>
</table>
By grouping concepts into categories, the number of units the researcher works with can be reduced. In our case, Table 2 provides an overview of how concepts were integrated into higher level categories. For example, various concepts could be identified as supporting the strategy of communication with the client, whereas other could be identified as supporting the category of managing creativity internally. Similar, different potential roles of persons in creative organizations could be grouped under the category of the artist, whereas other roles were grouped under the category of the creative supervisor. For example, leads, visual effects supervisors, producers, and team leaders were grouped under the category of the creative supervisor as these concepts share various characteristics that are of relevance in order to explain how processes in creative environments are managed.

Table 2: Developing categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managing the scope of creativity</td>
<td>Avoid endless exploring, channeling creativity, encouraging creativity,</td>
</tr>
<tr>
<td></td>
<td>creative buy-in, giving latitude, granting bending rights and responsibilities, providing stimuli to creative people,</td>
</tr>
<tr>
<td></td>
<td>putting people under time pressure, restricting creativity</td>
</tr>
<tr>
<td>Task allocation and team building</td>
<td>Challenge creativity, finding a balance between technical and creative,</td>
</tr>
<tr>
<td></td>
<td>identification of complementary skill sets, task allocation, trial and</td>
</tr>
<tr>
<td></td>
<td>error, getting the right team</td>
</tr>
<tr>
<td>Artist</td>
<td>Visual effects artist, editor, sound editor, compositor</td>
</tr>
<tr>
<td>Creative supervisor</td>
<td>Lead, visual effects supervisor, producer, team leader</td>
</tr>
</tbody>
</table>

Having formed categories, these could be further developed by identifying their properties and dimensions. As properties can vary along dimensional ranges, they give concepts variation. For example, artists are characterized by a certain attitude towards IT. This attitude varies on a continuous range from IT-averse to IT-affine.

Arguably, even though GTM concerns the emergence of theory from data, such research also starts with some (inevitable) preconceptions. Yet, through the open coding exercise, a number of apparent concepts emerged during the research that the researchers had not thought about beforehand. Moreover, open coding enabled the development of categories with great explanatory power that was expressed by properties and dimensions that enabled to cover a broad range of various situations. Thus, open coding provided rich insight into the processes that were studied. For instance, it became apparent that managing processes in creative environments requires the managers to gain an in-depth understanding of a complex interplay of various factors in order to decide what strategies and tactics can be used and combined in a particular situation.

3.4 Axial Coding

The central purpose of axial coding is to relate structure with process. In this study it was sought to understand the structure of processes in creative environments, that is, involved stakeholders, resources, etc., as well as how organizations effectively deal with such processes. In order to identify relationships, Strauss & Corbin (1998) propose a paradigm that differentiates categories into conditions (contextual, causal, and intervening), phenomena, actions/interactions, and consequences. Eventually, this process of relating categories results in the formation of conjectures or hypotheses. To that end, categories are related at a dimensional level.

In the present study, for example, requirements specifications to the creative product was identified as a contextual factor that impacts the context of organizational creative processes. The data further suggested that requirements specifications may vary from low to very detailed. Thus, requirements
specifications can be qualified dimensionally and then be related to the category of the creativity-intensive process. This insight, in turn, leads to the following preliminary proposition:

**Detailed requirements specifications to the creative product (dimensionally qualified factor) lead to low variance in outcome (dimensionally qualified consequence) of the process (phenomenon).**

This way, the conditional structure was identified and thus structure was linked with process. The paradigm was useful to studying the phenomenon under observation. The distinction between contextual factors, strategies, and consequences enabled the researchers to generate an understanding of the studied processes, which very much represents common perspectives on business processes, as, for instance, evidenced in the wide range of different required capability areas for progression BPM in organizations (de Bruin, 2007).

### 3.5 Selective Coding

Axial coding was followed by selective coding, where the core category representing the central phenomenon of the study was identified and other (sub-) categories were related to it in order to proceed to an integrated theoretical scheme. The **creativity-intensive process** representing organizational creative processes was identified as being the core category of this research. When the act of selective coding was started, we could see that there was need to further group concepts in order to densify categories. For example, in open coding and axial coding two main groups of strategies could be identified: those referring to internally managing creative processes, and those referring to externally managed creative processes. Table 3 shows how the categories of communication with the client and managing creativity internally were grouped under a higher-level category strategies in managing creativity-intensive processes. The category of strategies in managing creativity-intensive processes by itself became a sub-category of creativity-intensive processes, explaining how creative organizations deal with the phenomenon of creativity within business processes.

**Table 3: Relating categories to sub-categories**

<table>
<thead>
<tr>
<th>Category</th>
<th>Sub-category</th>
<th>Specific strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategies in managing</td>
<td>Communication with the client</td>
<td>Creative brief, matching requirements with capabilities, providing stimuli, showing</td>
</tr>
<tr>
<td>creativity-intensive processes</td>
<td></td>
<td>references, ongoing communication, showing work in progress, approval and review</td>
</tr>
<tr>
<td></td>
<td>Managing creativity internally</td>
<td>Task allocation and team building, resource allocation, managing the scope of creativity,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>internal reviews, internal task breakdown</td>
</tr>
</tbody>
</table>

All contextual factors, strategies, and consequences were mapped to the creativity-intensive process. As business processes provide a cross-functional perspective onto an organization, they are of such complex nature that they can be viewed as a concept with the capacity of integrating both structural and dynamic aspects of the emergent theory. Moreover, selective coding led to a substantial reduction to few categories with greater explanatory power. Thus, a parsimonious theoretical frame was established that consisted of only a limited number of concepts grounded in a multitude of incidents.

### 4 ON THE USE OF GTM FOR STUDYING BPM PHENOMENA

Orlikowski (1993) justified the usage of grounded theory through its focus on contextual and processual elements, and its focus on the action of key players associated with organizational change. Obviously, these aspects are of relevance also to BPM: the researcher studies processes as well as the action of key players (usually referred to as actors) who participate in organizational processes or see to the governance or management of these.
As the above described example shows, understanding BPM issues not only requires understanding what tasks are carried out in what order and what resources are required and what is the process output (i.e., a mere ‘process’ focus). It also requires the researcher to understand a variety of contextual factors that impact the business process and also strategies and tactics that individuals and organizations apply in order to handle these complex phenomena. Moreover, one must understand both intended and unintended consequences. Such strategies and tactics, for example, often involve the application of IT systems (van der Aalst et al., 2007), involve the assessment and mitigation of organizational risks (Lambert et al., 2006) or other contextual elements (Rosemann et al., 2008).

In the BPM study conducted, the grounded theory study allowed us to uncover that organizational processes in film and visual effects production organizations are characterized by a complex interplay of artists, clients, creative supervisors, and organizational resources. Creativity in such business processes causes high unpredictability with regard to process structure, required resources and process outcome. Therefore, the processes are characterized by the iterations between creating a mutual understanding of the requirements of the creative product, generating the creative product, and reviewing it. In order to manage this process, creative supervisors act as process intermediaries and apply various strategies in both internally managing the process and communicating with clients while they pursue both creative and operational process performance.

The experiences in our study strongly suggest that GTM as an inductive, theory-building research method is appropriate to provide insights into BPM-related phenomena. We found that particularly the GTM focus on involving a variety of data sources as a valuable approach to research. The triangulation can help researchers to consider various viewpoints onto a particular phenomenon (e.g., a process), and, in turn, develop rich and comprehensive insights. Yet, with all benefits considered, we also encountered challenges and issues in our study. Table 4 provides an overview of such issues that advocate the use of GTM in order to study BPM related phenomena. The different issues are illustrated by examples from the above study.

Table 4: Using GTM for studying BPM issues

<table>
<thead>
<tr>
<th>#</th>
<th>Issue</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Practical Relevance</td>
<td>GTM aims at producing substantive theory of high practical relevance.</td>
<td>The substantive theory that was developed is applicable to the particular context. It provides practitioners with profound knowledge about factors that need to be considered when managing organizational creative processes. Moreover, the theory explains what strategies and tactics can be used and combined in order to handle specific situations.</td>
</tr>
<tr>
<td>2</td>
<td>Starting point for developing formal theory</td>
<td>Substantive theory can be the starting point to proceed to more general theory.</td>
<td>In our study we first generated a theory on business processes applicable to the substantive area of the creative industries. By engaging with existent theory we proceeded to a more general theory. In later work such theory may then be formalized and/or tested.</td>
</tr>
<tr>
<td>3</td>
<td>Accounting for contextual factors</td>
<td>GTM produces theory that accounts for both intervening and causal conditions and, thus, recognizes the context of business processes and business process management issues.</td>
<td>Particular attention was paid to the various factors that impact business processes in creative environments. It turned out that processes highly depend on both internal and external factors that must be considered when managing processes.</td>
</tr>
<tr>
<td>4</td>
<td>Inductive nature</td>
<td>The inductive, emergent nature of GTM enables researchers to discover new concepts and relationships that have not been considered in previous studies.</td>
<td>The GTM approach led to the development of complex categories with great explanatory power. For each category a number of properties was identified that enabled to explain a variety of situations. Thus, the method enabled not only to identify ‘obvious’ concept (e.g., tasks and roles) but also other concepts (e.g.,</td>
</tr>
</tbody>
</table>
personal traits and abilities of actors) that need to be considered when managing processes.

Various techniques of data collection were used, including semi-structured interviews and the usage of process modeling techniques and related process documentations. This triangulation across methods provided rich insights incorporating different viewpoints.

It turned out that the method could capture the dynamic nature of organizational processes. Different stages of processes could be identified and various strategies and tactics could be linked to these stages.

The detailed procedures of GTM helped the researchers to rigorously develop theory without being lost in vast amounts of data.

5 Using multiple methods of data collection

Grounded theory enables the researcher to include a variety of data sources; for example, process models, process descriptions or voiced process experiences, that can often be found in organizations.

6 Studying processes

GTM is particularly fitting in order to study process-related issues.

7 Guidelines

GTM provided guidelines how to actually conduct a study

Over and above the issues and examples described in Table 4, GTM also comes with the label of being highly relevant to practice. Glaser (1978) argues that grounded theorists can contribute a great deal by providing the ‘man in the know’ with substantive theory, and expresses a number of arguments suggesting the high practical relevance of GTM studies (see Table 5, leftmost column). Arguably, the BPM domain is a field characterized by high relevance to practice (Gartner Group, 2009). In this domain, the relevant ‘man in the know’ may be a BPM professional, process owner, or manager responsible for a particular process. We argue that GTM-generated theory does in fact provide these stakeholders with information and insights relevant to their task settings. Accordingly, in Table 5 we provide exemplary evidence for the arguments provided by Glaser (1978) to substantiate of relevance.

Table 5: Practical relevance of GTM studies of BPM phenomena

<table>
<thead>
<tr>
<th>Arguments for practical relevance, adopted from (Glaser, 1978)</th>
<th>Practice of BPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practitioners get the ability of identifying additional consequences, conditions, and strategies beyond those that were empirically known to them</td>
<td>Particularly those BPM practitioners that are relatively new to a field will profit from theory that explains essential conditions, strategies and consequences. Thus, grounded theories can be used to inform practitioners about both intended and unintended consequences and, for example, may facilitate risk management.</td>
</tr>
<tr>
<td>Practitioners are supported in expanding the description and meaning of incidents by placing them into the greater context. Increasing the expert’s capacity to know by introducing a limited number of concepts that can be applied in various situations.</td>
<td>The BPM practitioner is able to abstract from a specific situation (such as a specific process failure), reflect on it, and develop new solutions to the problem situation at hand.</td>
</tr>
<tr>
<td>By expanding the practitioners’ theoretical knowledge, they are enabled to expand their capacity to deal with new, more complex situations, as their knowledge is not limited to particular incidents anymore. The theory helps experts to emancipate from restrictions of their specific expertise as theory enables them to become more open to change as they begin to see that certain incidents are mere patterns of a process.</td>
<td>Grounded theorists try to generate categories with great explanatory power. Such powerful concepts can be easily remembered by BPM practitioners and applied in various contexts.</td>
</tr>
<tr>
<td></td>
<td>BPM practitioners can assume responsibility for more complex processes. By applying theoretical knowledge they are no more limited to only knowing certain instances of particular processes.</td>
</tr>
<tr>
<td></td>
<td>Business processes in many industries are highly dynamic and require BPM practitioners to constantly adapt to new challenges. Thus, emancipation can be seen as an important enabler for process change.</td>
</tr>
</tbody>
</table>
Practitioners see their knowledge in a theoretical view and can capitalize on it through sharpened judgment and the improved ability to identify new strategies etc.

BPM practitioners become sensitized and are empowered to identify new strategies and practices in order to successfully manage processes.

5 CONCLUSION

In this paper we advanced the use of grounded theory as a research method appropriate to the study of phenomena typically associated with business process management. Based on the experiences gathered during a study of organizational creative processes, we outlined arguments for the use of GTM in general, and the ways in which it can be used for theory development in the context of Business Process Management in particular.

With our paper we seek to contribute to the ongoing discussion on the application of qualitative research methods to the IS discipline. We have argued that GTM is particularly fitting to study BPM phenomena. This is mainly reasoned by the nature of such phenomena and related studies: (1) BPM phenomena are often bound to particular contexts, (2) BPM research is thought to provide practical insights, and (3) business processes are characterized by interactions between various actors and IT artefacts to, collectively and collaboratively, carry out organizational processes.

We hope that our example stimulates fellow colleagues to also employ GTM in the area of BPM research. We believe that such studies can develop rich and comprehensive understandings into complex and dynamic phenomena such as business processes, and the management of these. We particularly would like to encourage fellow scholars to use GTM in multi-method approaches, for example, to inductively develop theories and hypotheses and, then, to test the propositions using typical deductive theory-testing methods such as experiments or surveys. Recently, we have seen some studies employing some form of multi-method research (Recker, 2008).

We believe that GTM, and more generally speaking theory development – could be highly beneficial to a range of BPM-related phenomena, such as the factors critical to success – or failure – of process optimization projects, the phenomenon of organizational adoption of BPM standards, the key contextual factors facilitating process change, or the embodiment of a process-aware culture in organizations, to name just a few.

References


CONSTRUCTING COMPARABLE BUSINESS PROCESS MODELS
WITH DOMAIN SPECIFIC LANGUAGES – AN EMPIRICAL
EVALUATION

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0367.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Conceptual modelling, Domain ontology, Business Process Management, Empirical study</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
CONSTRUCTING COMPARABLE BUSINESS PROCESS MODELS WITH DOMAIN SPECIFIC LANGUAGES – AN EMPIRICAL EVALUATION

Becker, Jörg, University of Münster, Leonardo Campus 3, 4149 Münster, Germany, joerg.becker@ercis.uni-muenster.de

Breuker, Dominic, University of Münster, Leonardo Campus 3, 4149 Münster, Germany, dominic.breuker@ercis.uni-muenster.de

Pfeiffer, Daniel, University of Münster, Leonardo Campus 3, 4149 Münster, Germany, daniel.pfeiffer@ercis.uni-muenster.de

Räckers, Michael, University of Münster, Leonardo Campus 3, 4149 Münster, Germany, michael.raeckers@ercis.uni-muenster.de

Abstract

The objective of this paper is to evaluate the semantic building block-based approach as a means for improving comparability in business process modelling. It is described whether and why the semantic building block-based approach reduces the variations in comparison to traditional modelling approaches. Our argumentation is grounded on the assumption that business process modelling projects in large organisations have to be conducted in a distributed manner. However, the goal of these projects is to integrate single models into a consistent process landscape. This allows the organisation to mine the processes for potential improvements. A lack of comparability could deteriorate the quality of the process landscape and the analysis performed on its basis. In a laboratory experiment the variations of distributed process modelling in the traditional and the building block-based approach have been compared. Results indicate that the semantic building block-based approach leads to considerably fewer variations between business process models and, thus, improves the comparability of them.

Keywords: Model Comparison, Domain Specific Languages, Business Process Modelling, PICTURE.
1 INTRODUCTION

Comparability is an important quality criterion for business process diagrams (BPD). In large business process modelling projects, as they are often executed in huge companies, the documentation of relevant domain knowledge has to be accomplished in a distributed manner. This is even more important if the modelling project crosses organizational boundaries. The challenge of such projects is not to represent the domain knowledge in single process models that are later on analyzed separately, but rather to combine all the models to a process landscape providing an overview over the companies business. The involvement of multiple actors in the modelling process causes the resulting models to differ in terms of vocabulary, level of detail and level of abstraction (Hadar & Soffer 2006). Creating an integrated description based on such models is thus hampered. Hence, an overall view on the organization is blurred by deviating ways of describing it. The possibilities to set up a holistic business process management based on the complete set of process models is therefore limited (Becker & Algermissen & Falk & Pfeiffer & Fuchs 2006).

The problem of variations between business process models can be addressed in two different ways. On the one hand, variations can be seen as given, which makes it necessary to resolve conflicts. This can either be done in an automated fashion (van Dongen & Dijkman & Mendling 2008), which, however, delivers only approximate results, or manually (Pfeiffer & Gehlert 2005), which involves significant efforts. The second way of dealing with variations is to avoid them during the construction of a BPD. This can be accomplished by constraining the choices a modeller can make when he creates a BPD. Domain specific modelling languages lower the ambiguities during the construction of conceptual models (Pfeiffer 2007). Thus, we claim that the application of a domain specific modelling language reduces the variations between BPDs compared to a traditional process modelling language. The aim of this paper is to deliver empirical evidence to support this assumption.

The paper proceeds as follows: In the next section the different variations within and between business process diagrams are described that can emerge in distributed modeling projects. In the following section of this paper the fundamental characteristics of the semantic building block-based approach are discussed. The specific structure of this approach is confronted with the properties of traditional modeling languages. Subsequently, the semantic building block-based approach is evaluated in a laboratory experiment. The distributed modeling conflicts that can emerge in the building block-based approach are compared to the traditional approach. The paper closes with a short summary of the main results and an outlook to future research.

2 NOTIONS OF VARIATION IN BUSINESS PROCESS MODELING

A real world phenomenon can be represented through BPDs in many different ways. BPDs are constructed by using two different languages. The first one is the modelling language. Its meaning can be formally specified, which makes this part of a process diagram unambiguous. The other component of a BPD consists of a domain language. It is used to make understandable statements about real world phenomena. In order to create a BPD, both languages must be applied together to create (human) readable process descriptions. Domain languages are owned by a linguistic community that decides on the meaning of its statements by shared conventions, which have been established implicitly by using the language. Because of the ambiguity of such natural languages it is possible to express the same meaning by different combinations of constructs and domain statements.

Variations in BPDs arise from both, differing perceptions of reality and from the process of explicating this perception. A variation is a semantic or syntactic deviation between different BPDs which refer to the same or a similar real world phenomenon. They can be due to two different reasons (Soffer & Hadar 2007).

- Variations due to varying mental representations: The mental representations of two model creators are most likely not exactly the same. This means the model creators perceive or structure
real world phenomena differently. Likewise, they can, consciously or unconsciously, consider deviating aspects of the phenomenon as relevant. This can lead to BPDs at diverse levels of abstraction. Likewise, in these models the sequence of activities can vary or the model elements can be annotated with a different number of details.

- **Variations due to the explication:** Even when the model creators share “the same” mental representation variations can arise. These variations result from a different explication of the mental representations. Domain and modeling languages offer certain degrees of freedom to express a given fact. Model creators can utilize this freedom in diverse ways. For example, different domain statements can be chosen to express a specific aspect of the mental representation. Similarly, a model creator may have the choice between multiple constructs to describe a given fact. Thus even with an equivalent mental representation, different BPDs with corresponding conflicts can emerge.

Deviations between models have been investigated empirically especially in the context of structural models. UML Class Diagrams have been analyzed in multiple modelling experiments (Hadar & Soffer 2006, Lange & Chaudron 2006, Soffer & Hadar 2007). Other empirical studies have focused mainly on the advantages of specific constructs in comparison to alternative forms of representation, such as entity types and attributes (Shanks & Nuredini & Tobin & Moody & Weber 2003), properties of relations (Burton-Jones & Mesu 2002, Burton-Jones & Weber 1999), optional properties (Bodart & Patel & Sim & Weber 2001), or whole-part relations (Shanks & Tansley & Nuredini & Tobin & Weber 2002). There are only a very few empirical studies that refer to variations in process diagrams. Mendling et al. (2006), for example, have analyzed the SAP Reference Model to identify errors and inconsistencies. Gruhn and Laue (2007) have investigated the role of OR-connectors in Event-driven Process Chains (EPC). Beneath these empirical studies, conflicts between models have theoretically been discussed in the database schema matching and integration literature (Kashyap & Sheth 1996, Parent & Spaccapietra 1998), in publications about metamodeling (e.g., Rosemann & zur Mühlen 1998), and ontology engineering (Davis & Green & Milton & Rosemann 2003). In this paper we draw upon Pfeiffer (Pfeiffer 2008) who has derived a comprehensive theoretical analysis of the variations in the context of business process modelling. The different variations are described in Table 1.

<table>
<thead>
<tr>
<th>Variation name</th>
<th>Variation description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type variation</td>
<td>Two model elements have the same meaning but a different construct (type) assigned.</td>
</tr>
<tr>
<td>Synonym variation</td>
<td>Two model elements have the same meaning but different labels.</td>
</tr>
<tr>
<td>Homonym variation</td>
<td>Two model elements have the same label but a different meaning.</td>
</tr>
<tr>
<td>Abstraction variation</td>
<td>Model elements in two different models have a deviating level of abstraction.</td>
</tr>
<tr>
<td>Control flow variation</td>
<td>The number of outgoing or incoming control flows of two corresponding model elements differs.</td>
</tr>
<tr>
<td>Annotation variation</td>
<td>A model element in the first model is annotated with a different number of model elements than a model element with a similar meaning in the second model.</td>
</tr>
<tr>
<td>Order variation</td>
<td>The order of the two model elements is permuted between two BPDs.</td>
</tr>
<tr>
<td>Separation variation</td>
<td>There is a model element that has no corresponding model element in the second model with the same, a more general, or a more specific meaning.</td>
</tr>
</tbody>
</table>

*Table 1: Description of the variations between business process diagrams*

In the following we will focus on type, synonym, abstraction, order and separation variations as part of the analysis. Homonym, control flow, and annotation variations are not taken into account. The reasons for this limitation are discussed in Section 4.
3 TRADITIONAL AND SEMANTIC BUILDING BLOCK-BASED PROCESS MODELLING

The application of traditional business process modelling languages leads to business process diagrams that are hard to compare. Every model created with a traditional language can include many of the variations described in the previous section of this paper. For instance, an EPC basically consists of events and functions, whose semantics are essentially defined by the domain statement the modeller assigns to it (Kellger & Nüttgens & Scheer 1992). Only by applying various rules and modelling conventions, comparability between the BPDs can be achieved (Schütte & Rotthowe 1998). The creation as well as the implementation of such regulations within a specific modelling project involves significant efforts.

By using a business process modelling language which belongs to the semantic building block-based approach, the comparability of the resulting business process diagrams can be significantly improved. These semantic building block-based languages (SBBL) achieve this advantage by avoiding the conflicts that occur when traditional modelling languages are used (Becker & Pfeiffer & Räckers 2007d, Pfeiffer 2007). The semantic building block-based approach guides the modeller through the modelling process and restricts him in his decisions. By decreasing the choices a model creator can make during the model construction, the comparability of the BPDs can be increased (Pfeiffer 2008).

Analysing BPDs created with a SBBL on a semantic level can be more cost-effective compared to the analysis of BPDs created with traditional languages. While the effort of creating a SBBL is relatively high, various analysis methods can be predefined and executed automatically on models created with it (Pfeiffer 2008), which may compensate the creation effort. Using traditional languages to create BPDs requires business process experts to perform every analysis manually or semi-automatically. Such an approach does not scale.

The main modelling construct of the language class SBBL is the so-called process building block (PBB). PBBs limit the degree of freedom within the process of model creation. Unlike traditional business process modelling languages the SBBLs employ PBB as their most important modelling constructs. Every PBB represents one or more reoccurring activities from a particular domain (Baacke & Rohner & Winter 2007, Becker et al. 2007d, Lang & Glunde & Bodendorf 1997). The difference between a PBB and a modelling construct from a traditional language is that the PBB already incorporates a domain statement. Modellers do not create and assign a domain statement to a construct, they can only choose from a given set of PBBs and, thereby, from a given pool of statements. Thus, the PBB are semantically specified and have a defined level of abstraction (Rupprecht & Funffinger & Knublauch & Rose 2000). If additional information is needed, the PBB can be further described by a predefined set of attributes.

Concerning their semantics, the PBB are unambiguously and mutually exclusively defined. To specify the constructs of a SBBL, a domain ontology is used. Every PBB stands for a set of elements taken from this ontology. Hence, the meaning of a PBB is explicitly defined. With the aid of the ontology, it is possible to ensure that no element of a SBBL contains semantics already covered by another element of this language. Given a real world phenomenon, there exists only a single possibility to represent it in a SBBL-based language. In ideal, every construct would be derived from the domain ontology, but from a practical perspective it is often necessary to include at least some constructs from other languages. For instance, this could be a construct to split up and join the control flow. In Figure 1, for example, the ontology element ‘encash/receive a payment’ has been incorporated into a SBBL as a PBB. Also the corresponding attributes of the PBB are taken from the domain ontology. This encompasses not only the attributes themselves, but also their possible values. In the given example, the attribute ‘Information System’ has only three allowed values: ‘Open Office’, ‘MS Office’ and ‘MS Money’. The available labels for the PBB, which specify the domain task more detailed, are defined in the same manner. For the PBB ‘encash/receive a payment’, the labels
’encash/receive a cash payment’, ‘encash/receive a credit card payment’, and ‘encash/receive a money transfer’ are allowed.

Figure 1: A Process Building Block and a Section from a Domain Ontology

Languages from the class SBBL either avoid or at least decrease the previously described variations between BPDs. By using the semantic building block-based approach, some types of variations between models can be fully eliminated. As the PBB do not offer multiple ways to express a specific fact, variations due to explication cannot occur. However, variations due to varying mental representations are still possible. But, because of the ontology incorporated into the language that is guiding the modeller through the modelling process, their frequency can be significantly reduced. In the following the impact of the language class SBBL is discussed with regard to the five variation types considered:

- **Synonym variations**: Because of the fact that the constructs of languages from the class SBBL are derived from an ontology, they offer a controlled vocabulary to the modeler. Synonyms can be detected in the ontology, which makes it possible to eliminate them in advance of the model creation. Hence, as long as the modeler can only choose from the given vocabulary of a SBBL, no synonym variations can occur.

- **Type variations**: During the language construction, it is ensured that no semantically overlapping modeling constructs are included in the SBBL. If every PBB and every attribute of the language is semantically disjoint, it can be proven that no type variation can occur (Pfeiffer 2007). For every observable real world phenomenon only one single construct exists which is able to represent it within the language. Therefore, every modeler who wants to describe the phenomenon is forced to use same construct.

- **Abstraction variations**: The type in combination with the label defines the semantics of a PBB. Because every PBB is semantically disjoint from the others, every modeler has to choose the same PBB to express a specific matter. Thus, the number of possible choices for the selection of domain statements and, thereby, also the number of abstraction variations is reduced. To completely avoid
them, a specific level of the ontology has to be defined from which all the domain statements of a model have to originate.

- **Separation variations:** This type of variation cannot be entirely removed from models created with the language class SBBL. Nevertheless, it can be at least reduced because during model construction the modeler is guided by the ontology-based PBBs he can choose from. With the meaning of the PBBs in mind, he focuses on the semantics covered by them. Therefore, the models better fit to each other concerning the semantics they express.

- **Order variations:** Just like the separation variations, this type of variation cannot be completely avoided. In traditional modeling languages, it is hardly feasible to make any statements about the correct order of specific elements on the basis of their type. In contrast to that, the semantic building block-based approach allows to define heuristic order rules based upon the predefined semantics of the PBBs. For example, it is reasonable that the activity ‘approve’ always follows the activity ‘perform a formal verification’.

The creation of languages from the class SBBL can only be accomplished successfully with a specific domain in mind. In order to be able to express every real world phenomenon by using a modelling language of this type, it is necessary to restrict the application to a specific domain. Otherwise, no appropriate ontology can be created due to the complexity of the real world. Hence, languages from the class SBBL are domain specific languages. A well documented example for such a language is the PICTURE-language, which is specifically designed for public administrations (Becker & Algermissen & Falk 2007a, Becker & Algermissen & Pfeiffer & Räckers 2007b, Becker et al. 2007d, Falk 2007). It consists of 24 PBB and over 50 attributes. The PBBs in PICTURE can only be connected in a sequential form. For an in-depth description of the language, we refer to (Becker et al. 2007a).

### 4 EVALUATION OF THE SEMANTIC BUILDING BLOCK-BASED APPROACH

#### 4.1 Setting of laboratory experiment

The hypothesis to evaluate is that modelling with a semantic building block-based language results in a smaller number of variation compared to traditional modelling languages. In order to do this, an empirical evaluation was conducted. EPC was chosen as an example of a traditional modelling language, PICTURE as an example for a domain specific one.

Within a laboratory experiment, twelve graduate students from the University of Muenster were asked to create an EPC and a PICTURE model independently from each other based on a given case description. This case description was used to examine the variability between BPDs in both languages. This experimental setup simulates the process of distributed modelling and facilitates the validity of the analysis for two reasons. Firstly, all participants are modelling the same situation, which eliminates the case description as a source of variability. Secondly, every participant creates both an EPC and a PICTURE model. Thus, all variations resulting from a different understanding of the case description or from deviating opinions about the adequate degree of detail or abstraction influence the modelling process of both languages in the same way. The remaining variations can be fully explained by the process of explicating the mental representations of a participant in the form of a process diagram. Because the order in which the models had to be created was not specified, the participants were able to correct mistakes they made during the creation of their first model if they notice a misunderstanding during the creation of their second one. Thus, in case the understanding of the case description changed over time due to a learning process, the validity of the results is not threatened.

As PICTURE is a domain specific language designed to be used in public administrations, the case description was taken from this domain. It is about handling an application for a resident parking permit. First, the application has to be checked according to various criteria. Depending on the results, it is either accepted or rejected. In the latter case a rejection letter is created and send to the applicant. If the application is accepted, the application data must be entered into a software system, the parking
permit has to be created, an annual fee must be encashed and the parking permit has to be delivered to
the applicant. Various methods of payment as well as different ways to deliver the parking permit are
possible. The analysis has been carried out in two steps:

- **Automated analysis:** In the first step, both EPC and PICTURE models were tested for similarity
  with an automated comparison algorithm (van Dongen et al. 2008). This algorithm has been
designed to quantify the similarity of the process flow as well as to detect and resolve problems
resulting from the ambiguities of natural languages. The applicability of the algorithm has been
demonstrated empirically by using the SAP Reference Model.

- **Manual analysis:** The second step was, in contrast to the first one, conducted manually to
  reconfirm the results from the automated comparison. In order to do this, the authors analyzed the
  BPDs from both groups to find and quantify variations from the types described above. If a high
degree of similarity between the two models is found in the automatic analysis a small number of
variations can be expected in the manual analysis. The automated analysis of the models only
provides a percentage value of similarity. As the analysis is conducted manually in the second step,
the nature of the variations can be explored in more detail.

### 4.2 Characteristics of the Automated Analysis

The comparison algorithm which has been used to determine the degree of similarity between the
BPDs can be used for both PICTURE and EPC models in the same way. This is ensured by the fact
that the models themselves are not used for the similarity calculation. Instead, the result is computed
by using what is called a *causal footprint*. This is a construct that can be derived from the BPD. It is a
directed graph whose vertices represent the various activities in the process. Vertices are connected by
arcs whenever the corresponding activities of the vertices are always performed either before or after
one another. In the first case, the arc is called a look-back-link, in the second case it is a look-ahead-
link (van Dongen & Mendling & van der Aalst 2006). If, for example, there is an arc connecting the
vertices A and B, this means that, depending on the type of the arc, activity A is either always
performed before activity B or after it. In order to finally execute the comparison, the causal footprints
of the models must be transformed into vectors. Their similarity is then determined by the deviation of
their directions. For more details concerning the transformation, we refer to (van Dongen et al. 2008).

The comparison algorithm is able to identify ambiguities of natural languages within the labels of the
model elements. To calculate the similarity of BPDs, common elements must be identified. Therefore,
equivalent vertices need to be identified in order to compare two footprints. Natural languages allow
expressing the same real-world concepts in different ways. This hampers the automatic identification
of similar or equivalent activities. In order to deal with this problem, the comparison algorithm uses
the lexical database WordNet, which allows to detect synonyms (Miller 1995). With the aid of this
information, the semantic similarity of activities can be computed. Comparing the similarity score of
an activity and of all elements connected to it, it is possible to map equivalent activities of different
process diagrams (van Dongen et al. 2008).

The comparison algorithm determines the similarity of process diagrams regarding their content and
their respective process flow. The causal footprint consists of both the vertices representing activities
themselves and look-ahead as well as look-back-links, which stand for the procedural relations of the
activities. Therefore, the comparison does not only consider the similarity regarding the content, but
also takes the process flow into account.
4.3 Results of the Automated Analysis

Figure 2: Average similarity degrees for PICTURE and EPC models

12 BPDs from each group were compared pair-wise with each other. This resulted in a total of 66 comparisons for each group. Within the group of the EPC models, an average similarity of 0.54% has been measured. The maximum similarity was 4.02%, the minimum was 0%. This means that the comparison algorithm perceived the BPDs as being totally different. In contrast, the PICTURE models achieved an average similarity of 43.75%. Some comparisons resulted in a value of 100%, which means that the models were identical. Other PICTURE models scored lower values as well. The minimum value was 13.99%.

Detailed results are described in Figure 2. In this diagram the average similarities of the individual BPDs compared to all other models are depicted. Figure 2-I presents the similarity values for the PICTURE and the EPC group on a single scale, Figure 2-II uses separate scales instead.

4.4 Characteristics of the Manual Analysis

Detailed statements about the nature and the degree of variability between BPDs can only be given manually. A framework, which classifies possible variations between process diagrams into different categories, was introduced in Section 2. To identify these variations in process diagrams, a semantic analysis of BPDs is necessary. Thus, a specific meaning needs to be assigned to every model element according to the modeller’s intention. By this means, an ontology which describes the whole semantic of the case description has been developed. Thereafter, it was possible to assign statements of this ontology to every model element. The intended meaning had to be carefully explored by the authors. With the resulting assignments, the basis for the identification of variations was established.

When variations are identified they need to be counted in compliance with strict rules to assure a reasonable quantification of the variability. With the previously given definitions, variations can easily be identified. But the definition alone was not sufficient to generate a meaningful result. A set of rules for quantifying the identified variations had to be developed. They allowed for a consistent and uniform measurement. For example, rules were designed to prevent counting some variations multiple times. Different types of variations were not weighted, because there was no information about the extent to which an individual type of variation influences the comparability of BPDs. To measure synonym variations, all words in the EPC models were stemmed and frequently occurring words like “a”, “an” or “and” were not considered in the comparison. Taking variations resulting from such differences into account would distort the results, because transforming strings into the described form can be easily accomplished automatically (Porter 1980).

With the given experimental setup, a reasonable measurement of homonym, control flow, and annotation variations was not possible. All models were created on the basis of the same case
description. This makes the measurement of homonym variations difficult, because they occur when different concepts are expressed by the same terms. This usually happens in complex systems of different BPDs, however, not within a single case. Annotation conflicts were not measureable because no attributes were used within the EPC and only a fixed set of attributes within the PICTURE models. The PICTURE as well as the EPC language has strict rules concerning the incoming and outgoing control flows. In fact, only the AND, OR, and XOR operators from the EPC language allow for deviating numbers of control flows. Hence, no control flow variations were detectable during the analysis.

4.5 Results of the manual analysis

Within the variation analysis an average of 31.93 variations between EPC models were identified. An average of 12.59 of these variations were synonym variations, 5.95 were abstraction variations, 10.70 were separation variations, 2.15 were type variations, and 0.53 were order variations. The group of the PICTURE models scored an average value of 4.59 variations. It consists of 0.63 synonym variations, 0.83 abstraction variations, 1.77 separation variations, and 1.32 type variations. Order variations were not found between PICTURE models. A comparison of the combined results can be found in Figure 3.

![Figure 3: Numbers of variations for PICTURE and EPC according to the different variation types](image)

4.6 Discussion of the results

The results of the automated similarity calculation are confirmed and further detailed by the manual analysis. While the automatic analysis can hardly find any commonalities between EPC models, it provides very good results for PICTURE models. In compliance with these results, the manual analysis shows a significantly higher number of variations of any kind for EPC models compared to PICTURE models. These results support assumption that the automated analysis is correct and further specify the results by categorizing the variations.

The semantics of BPDs that contain natural language elements cannot be captured automatically. The use of ontology-based labels for the PBBs in PICTURE actually results in a massive reduction of synonym variations compared to EPC. Although the algorithm used is build to detect synonyms, the low similarity degrees for EPC models imply that it fails to do so in most of the cases. The avoidance of many synonym variations by PICTURE in parallel with the high similarity degrees indicates that synonym variations cannot be resolved automatically.
The degree of detail and abstraction are fixed when using a SBBL-based modelling language. The limitation of the number of choices a modeller can make within the modelling project when he is using a SBBL in fact increased the comparability of the created models. A significant decrease of abstraction and separation variations in the manual analysis supports this conclusion.

It remains to be demonstrated that the expressiveness of a SBBL is sufficient. The increased comparability of models created with a SBBL leads to a decreasing expressiveness because of the predefined semantics of the PBBs. It is possible that the modeller is limited in his decisions that he is not able to represent all relevant real world facts by using the PBBs. Hence, the creation of a SBBL is very time consuming and error prone. This analysis only shows that the language class SBBL produces models with a higher degree of comparability, but it does not take the expressiveness of the models into account. Although there is a study focusing on this issue by evaluating PICTURE against workflow patterns (Becker & Algermissen & Pfeiffer & Räckers 2007c) and despite of the fact that PICTURE has been successfully used to create a total of more than 1,000 process diagrams in 12 german administrations until now (Bergener & Pfeiffer & Räckers 2009), a comprehensive empirical study is up to future research.

5 CONCLUSION

In the beginning of this paper we argued that BPD created in a distributed process modelling project exhibit a significant degree of variability which is a major obstacle to the creation of a sound overview over the organizations process landscape. In order to overcome this problem in BPD, we proposed to use languages from the SBBL class. Such languages reduce the variations between BPD by limiting the freedom of choice during model creation. To support our argumentation, an empirical evaluation of the PICTURE language, a well known example for a language coming from the class SBBL, was conducted. In a laboratory experiment that simulated a distributed modelling project the potential advantages of the language class SBBL have been analyzed. Both an automated and a manual approach were chosen to compare the performance of the two languages EPC and PICTURE. The results of the analysis demonstrate that the type of the language has a strong influence on the number of variations in the resulting BPDs. PICTURE considerably decreased the number of variations and, thereby, improved the quality of the corresponding BPDs.

However, the number of variations is only one component of the evaluation of the semantic building block-based approach. Furthermore, it is necessary to assess the efficiency and the effectiveness of the resulting languages. Efficiency means that a SBBL-based modelling language is able to acquire a specified number of processes at minimal cost. The creation of a domain ontology, whose existence is a necessary precondition for a language from the SBBL class, requires a considerable amount of effort, which could outweigh quality improvements. Effectiveness requires that a language of the class SBBL is expressive enough to describe the relevant phenomena of the domain at hand. In other words, effectiveness makes sure that the modelling language can indeed be successfully applied in a given domain. An empirical analysis of these two aspects is open to further research.

ACKNOWLEDGEMENTS

The work published in this paper is partly funded by the European Commission through the STREP PICTURE. It does not represent the view of European Commission or the PICTURE consortium and the authors are solely responsible for the paper's content.
References


# KNOWLEDGE CLUSTERS: DEALING WITH A MULTILEVEL PHENOMENON

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0169.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Case Study, Knowledge integration, IT Project Management, Multilevel research</td>
</tr>
</tbody>
</table>
KNOWLEDGE CLUSTERS: DEALING WITH A MULTILEVEL PHENOMENON

Bhandar, Mamata, U21 Global, 5 Shenton Way, #01-01, UIC Building, Singapore 068808.
mbhandar@u21global.edu.sg

Pan, S-L, National University of Singapore, 3 Science Drive 2, Singapore 117543.
pansl@comp.nus.edu.sg

Tan, Bernard, National University of Singapore, 3 Science Drive 2, Singapore 117543.
btan@comp.nus.edu.sg

Abstract

The central idea of this paper is to comprehend knowledge integration in inter-organizational IS projects through a knowledge cluster view of inter-organizational IS projects. The proliferation of such projects combined with the complexity of managing them, motivated this study. The study is based on a single case study which involves four organizations collaborating on a project. Findings suggest that knowledge integration in inter-organizational projects occurs through knowledge clusters and involves the interaction of complementary specialized knowledge bases within a structure and influenced by the clusters’ perceptions towards the project. Implications of the findings and future research opportunities are discussed.

Keywords: Knowledge cluster, knowledge integration, multi-level analysis, inter-organizational projects

1 INTRODUCTION

Inter-organizational arrangements for mutual benefits in the form of partnerships, alliances and joint ventures are very common. Such arrangements take place for several reasons including resource efficiency, resource acquisition, and skill enhancement (Simatupang et al 2002). Of the different types of inter-organizational arrangements, this study focuses on inter-organizational IS projects, an organizational form in which multiple organizations collaborate on an IS project for a specific period of time and that involve a client-vendor relationship. The simplest of such projects may involve one client and one vendor, but the more complex may have multiple clients or/and multiple vendors. With companies increasingly outsourcing all or some of their IS activities (Lacity and Hirschheim 1993), including IS development, it is very common indeed to see IS projects that involve multiple organizations.

Development of Information systems is a complex, intensive, and dynamic activity that requires close cooperation and coordination among diverse stakeholders who house the knowledge required for the execution of the project. In inter-organizational IS projects, knowledge required for the project is situated in the diverse stakeholders that belong to different functional groups (departments, users, management, IT etc) and are part of different organizations. Typically for a project some representatives from each of these departments/functional groups/stakeholders, referred to as knowledge clusters in this paper, are involved in the project. Specifically, knowledge clusters in this study refers to a group that possesses certain specialized knowledge essential for the project (e.g. users, IT department representatives) and that is governed by similar rules and boundaries (e.g.
departments) and hence the members share similar perceptions towards the project. This definition of knowledge clusters is similar to that adopted by Yayavaram and Ahuja (2008), although they have not explicitly defined the term and they base their definition on technologies. Other studies that have defined the term are in the context of industrial clusters, as cluster of firms, a more macro level view than our consideration. Knowledge from multiple knowledge clusters in each collaborating organization can be highly differentiated and specialized and therefore has to be integrated for the project.

Knowledge integration, in this context is conceptualized as the process through which disparate, specialized knowledge located in multiple knowledge clusters across organizations is combined, applied and assimilated. For instance, in an IS project, the users from the client organization communicate system requirements to the vendor’s IT consultants who use their software expertise and knowledge from the users to build the system. Users then assimilate the system by making necessary changes to their work practices (Faraj and Sproull 2000; Huang et al 2001). Knowledge integration is essential since if knowledge from a particular cluster is missing or is not integrated, for example, lack of user participation and hence unclear requirements, the project outcome may suffer. The several constraints like time, budget, specifications require knowledge integration in an inter-organizational IS project to be effective, i.e. knowledge for the project has to integrated within the stated constraints.

Managing the knowledge integration process between the organizations in a project is a crucial task (Walz et al 1993) and is a challenge not only because knowledge is often dispersed, differentiated and embedded (e.g. Tsoukas 1996) in various knowledge clusters but also because the clusters have their own agendas within organizations that are intrinsically different, that may possess diverse competencies (Pisano 1994) and conflicting interests. The fact that many of the systems are developed under extreme time constraints and often with the help of external consultants exacerbate the challenges associated with combining diverse forms of expertise on particular projects (Levina 2005). Client-vendor relationships in outsourced IS development projects have always been presumed to be adversarial (e.g. Chowdhry and Sabherwal 2003) and the associated factors, along with the difficulties in obtaining quick feedback, meeting frequently, and building interpersonal relationships, make their management an arduous task.

How then is knowledge across multiple knowledge clusters from multiple diverse organizations integrated for effective knowledge integration in inter-organizational IS projects? To address this question and increase our understanding of this complex phenomenon in a challenging environment an empirical study of an inter-organizational IS project is conducted. Related literature is reviewed and knowledge cluster concept explained in the following section.

2 LITERATURE REVIEW

2.1 Knowledge and Knowledge clusters

This study bases itself on the view that knowledge exists in the individual and the collective (Nonaka 1994) and both include the tacit and explicit dimensions of knowledge as suggested by the early works of Polanyi (1966). Individuals in their clusters are assumed to be the prime vehicle of knowledge delivery to the project irrespective of its nature and location. The notion of collective knowledge is important because Collective knowledge can also refer to knowledge situated in various departments, groups or teams. The representatives of certain departments, groups, teams are referred to as knowledge clusters in this study and they house the different types and forms of specialized knowledge (e.g. process knowledge, requirements) required for the project. The concept of knowledge cluster is represented in figure 2.1.
Each knowledge cluster groups representatives from a certain knowledge area that will be involved in the project. This group can include some representatives from one knowledge base, more than one knowledge base or the entire knowledge base. For instance, knowledge cluster can be a few members from a certain department (e.g. users), representatives from two or more departments (e.g. multiple user departments) or the entire department (e.g. when the user department is small and everybody’s knowledge is required for the project). These three possibilities are shown in figure 1 above.

Yayavaram and Ahuja(2008) used the term knowledge cluster more in the context of their definition of a knowledge structure with respect to a firm’s technical knowledge. According to them a knowledge cluster can be a group of people, but that need not always be the case and sometimes it can be a group of related technologies. Although the usage of the term is similar the distinction in our adoption of the concept is mainly due to the context. Knowledge clusters in this paper are a group of people, encompassing individuals’ knowledge, the collective knowledge and embedded routines and practices thus including related technical knowledge as well. In identifying clusters for the current study, groups that were to contribute the same type of functional knowledge were considered; i.e most of the identified clusters belonged to different functional areas such as users, IT, management, or different vendor organizations with their specialized knowledge. The specialized knowledge of the clusters also provided a sense of the cluster’s role in the project, attitude, perception hence knowledge integration behaviour of the cluster towards the project.

2.2 Knowledge Integration and Inter-organizational IS Projects

Scholars have proposed different definitions of knowledge integration (table 2). For this study the knowledge integration as proposed by Bhandar et al (2007) is adopted since it incorporates a process view and is easily applicable to an inter-organizational context. Based on their view, knowledge integration for this study is viewed as the process through which relevant knowledge from different clusters is combined, applied and assimilated for the goal of the project. This view implies that knowledge integration is achieved through several activities starting from project negotiations to the post-implementation stages and is also influenced by certain antecedent conditions like reason for initiating the project. It also suggests that knowledge integration process requires the knowledge clusters to be involved in the process to contribute, apply and assimilate knowledge. Okhyusen and Eisenhardt (2002) add that knowledge integration is not about assembling discrete pieces of knowledge, but depends on how members know and integrate their individually held knowledge and so is influenced by the surrounding environment.
The process of knowledge integration in inter-organizational IS projects as described involves ongoing interaction between the clusters to contribute knowledge and share common experiences to redefine shared beliefs through social interaction (Huang et al. 2001) thus entailing that the clusters have a motivation to be part of the project. For instance, top management may need the system but the user department may not. This can affect the users’ assimilation of the system hence their participation in the knowledge integration process. Okhyusen and Eisenhardt (2002) highlight the importance of structures to improve interactions among group members for knowledge integration and Van Den Bosch et al. (1999) and Grant (1996) suggested the importance of the ability of the firm for knowledge integration. Clusters need to possess the ability to comprehend the ideas and perspectives being exchanged so as to contribute knowledge to the project and reinforce the common knowledge base, i.e. the overlap of knowledge that exists between the network members (Demsetz, 1991; Grant, 1996). Knowledge integration in inter-organizational IS projects is not easy given it involves the integration of knowledge spanning cross functional capabilities (Carlile and Rebentisch 2003) which is more complicated compared to integrating one kind of knowledge across individuals or groups (Grant 1996), notwithstanding the inherent characteristics of knowledge that can make its integration difficult. The common knowledge that exists in an inter-organizational set-up is modest and the fact that knowledge in each organization exists in multiple entities further adds to the challenge. Additionally, these projects involve managing knowledge clusters from different organizations. The clusters are affected by the inter-organizational dynamics such as distinct competencies (Pisano 1994) and conflicting interests that have to be overcome before their knowledge can be integrated (Pan et al. 2001). Conflicts inherent in a client-vendor relationship (Lacity and Hirschheim 1993) although healthy from the perspective of the organization can affect the behavior of knowledge clusters towards knowledge integration. So how then is knowledge integrated in inter-organizational IS projects? To address this question and understand this complex phenomenon in a challenging set-up, an empirical study of a four organization project is undertaken. The methodology followed for the study is discussed next.

### 3 METHODOLOGY

Qualitative research method was adopted for this study since it allows an emphasis on processes and meanings (Denzin and Lincoln 1994) essential for this study investigating a knowledge integration process. The case study method was deemed appropriate for data collection since the phenomenon of knowledge integration is closely intertwined with the context of the inter-organizational project (Yin 2003). The study also required informants to reveal sensitive data that required comprehension of the context (e.g. to interpret the quotes in light of their relationships) which was possible through long and

<table>
<thead>
<tr>
<th>Definition</th>
<th>Author/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application of knowledge</td>
<td>Grant 1996</td>
</tr>
<tr>
<td>Synthesis of knowledge bases</td>
<td>Alavi and Tiwana 2002</td>
</tr>
<tr>
<td>Process view: an ongoing collective processes of constructing, articulating, and redefining shared beliefs through social interaction of organizational members</td>
<td>Huang et al 2001</td>
</tr>
<tr>
<td>Distinguish between knowledge integration process and knowledge integration per se. The process involves the actions of group members by which they share their individual knowledge within the group and combine it to create new knowledge. By contrast, knowledge integration is the outcome of this process, consisting of both the shared knowledge of individuals and the combined knowledge that emerges from their interactions</td>
<td>Okhyusen and Eisenhardt (2002)</td>
</tr>
<tr>
<td>Defines knowledge integration specifically in the context of IS projects as the process of embodying business application domain knowledge with technical knowledge in the design of the software.</td>
<td>Tiwana (2004)</td>
</tr>
<tr>
<td>Process view- process through which relevant knowledge is combined, applied and assimilated</td>
<td>Bhandar et al (2007)</td>
</tr>
</tbody>
</table>

Table 1: Knowledge integration views and definitions
informal interviews. This project was chosen for the study based on three criteria: the project was recently completed to ensure that participants could recall events, permission to study the project was granted by the top management so as to allow access to rich data, and it provided a right context for the study, an inter-organizational project with four organizations and client-vendor relationship.

The main source of data was face-to-face interviews conducted with representatives of each organization involved in the project at different hierarchies (top management, middle management, team members and users). Questions were asked to understand the motivations/expectations/views of each organization/cluster for the project, their account of how the project progressed, the conflicts and resolution of conflicts. Secondary data was collected from organizational websites (e.g. organizational background), articles, and third parties (employees of the companies not involved in the project). The multiple sources provided for triangulation (Stake 1994) of evidence, ensured that facts stated by one cluster could be verified by the other and also provided multiple perspectives on issues.

Data analysis was done in iteration with data collection (Myers 1997; Walsham 2006). Initially analysis was conducted at an organizational level, through which the behaviour of some members could be explained. The researcher then adopted a cluster level analysis, since they saw behaviour of members could be better explained from the cluster they area associated with rather than the organization. Data collected was transcribed in consideration with recording media for qualitative studies (Walsham 1995, 2006). Themes were identified using open-coding (Strauss and Corbin 1990) that influenced the organizations knowledge integration behavior throughout the project. For instance, ‘prior experience’ and ‘lack of motivation’ were identified as themes that influenced ‘requirement gathering’ since they affected time taken/outcome for that activity. Interesting comments, surprising revelations, special notes/observations made during the site visits or interviews were also considered. For example, highly sceptical tone, formal atmosphere, and interviewees being very guarded in disclosing facts were all noted. The case data collection details are presented in table 2.

<table>
<thead>
<tr>
<th>Organization</th>
<th>Interviewees</th>
<th>Interview details</th>
<th>Interview background</th>
</tr>
</thead>
<tbody>
<tr>
<td>Client</td>
<td>Project manager</td>
<td>2 Both interviews lasted for over an hour and were face-to-face</td>
<td>The client project manager gave detailed accounts of the project in two interviews at different stages of the project. Others’ substantiated on work practices and issues in the project. The external project manager and projects department provided background information on client IT project management.</td>
</tr>
<tr>
<td></td>
<td>Programmers</td>
<td>2 Questionnaires</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Project manager not involved in this project</td>
<td>1 Face-to-face interview lasting over 45 minutes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Projects department</td>
<td>1 Face-to-face over lunch meeting lasting 45 minutes</td>
<td></td>
</tr>
<tr>
<td>Vendor 1 –Banking solutions firm</td>
<td>Technical lead consultan</td>
<td>2 Both interviews lasted over and hour</td>
<td>Lead representatives gave detailed accounts of the project and their perspectives into the issues. Common issues arose and allowed for triangulation of evidence for key findings. Each gave access to their own team members who filled out questionnaires. Team members commented on the nature of relationships in the team, the camaraderie and the influence of social activities.</td>
</tr>
<tr>
<td></td>
<td>Consultants</td>
<td>5 Questionnaires</td>
<td></td>
</tr>
<tr>
<td>Vendor 2-Consulting firm</td>
<td>Lead representative consultants</td>
<td>1 Interview lasted over and hour</td>
<td></td>
</tr>
<tr>
<td>Vendor 3- Systems Integrator firm</td>
<td>Lead representative consultants</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

*Table 2: Case Data details*
4 CASE DESCRIPTION AND ANALYSIS

This project was the development of a data warehouse and it involved four organizations. The client was a leading bank in Asia and the three vendors; a banking solutions firm, systems integrator and business process consulting firm that were all engaged on a time and material contract\(^1\). The four organizations were multinational firms with distinct competencies and were represented by an expert each that were all data warehouse experts with 20 plus years of experience. A project team with representatives from all four organizations was set-up and collocated in the client’s premises. The project was on for two years at the time of data collection.

4.1 The knowledge clusters and their roles

Based on the case data and the definition of knowledge clusters the various knowledge clusters from each organization, their roles and representation to the project team are summarized in table 3. From the client there were four clusters; Client_Management, Client_Users, Client_IM and Client_IT. Client_Users comprised of business users and data analysts and this cluster had to communicate process flows to be incorporated in the data warehouse model and had to use the data model as it was being developed to assess suitability and usability. Client_IT was involved in building the data warehouse along with the vendor organizations. Client_IM was the Information management department which initiated and owned this data warehouse. It coordinated the project and the head of Client_IM was the project manager. Client_Management was involved initially to grant funding for the project after which it followed up with Client_IM after every quarter on the project progress. From the three vendor organizations there was one expert each that was the main coordinator and each brought in a small team for the area of expertise they were hired for. Each vendor organization is therefore considered as a single cluster. The three vendor clusters were; Vendor_Bankingsolutions, Vendor_Systemsintegrator and Vendor_Consulting.

<table>
<thead>
<tr>
<th>Org</th>
<th>Clusters</th>
<th>Team</th>
<th>Role</th>
<th>Knowledge Integration behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Client</td>
<td>Client_Users</td>
<td>∗</td>
<td>Data analysts &amp; business users. Contribute requirements &amp; use system</td>
<td>Did not see need for the system and was against it</td>
</tr>
<tr>
<td></td>
<td>Client_IT</td>
<td>∗</td>
<td>IT dept. Involved in system development</td>
<td>No prior experience in implementing data warehouse and resisted system development method</td>
</tr>
<tr>
<td></td>
<td>Client_IM</td>
<td>∗</td>
<td>Information management dept. Managed project and system development</td>
<td>Did not receive much cooperation from other clusters</td>
</tr>
<tr>
<td></td>
<td>Client_Management</td>
<td>∗</td>
<td>Acceded to the project and monitored project progress</td>
<td>Minimal involvement.</td>
</tr>
<tr>
<td>Banking solutions</td>
<td>Vendor_Bankingsolutions</td>
<td>*</td>
<td>Involved in system development technical leader</td>
<td>Passionate about project but no access to Client_Users knowledge. Had crashes with other experts</td>
</tr>
<tr>
<td>vendor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systems integrator</td>
<td>Vendor_Systemsintegrator</td>
<td>*</td>
<td>Information quality experts</td>
<td>Passionate about project. Did not get support from users and had clashes with other experts</td>
</tr>
<tr>
<td>vendor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business process</td>
<td>Vendor_Consulting</td>
<td>*</td>
<td>Business process consulting</td>
<td>Passionate about project. Did not get support from users had clashes with other experts</td>
</tr>
<tr>
<td>consulting vendor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Knowledge Clusters in the Case

\(^1\) Contract where vendors are paid for whatever time and material they put into the project.
4.2 Interaction of the knowledge clusters

The knowledge clusters’ interaction in this project is shown in figure 2. Although there was representation of every cluster to the project team, some of the links were not effective. The dotted lines indicate ineffective links/impeded knowledge flow from those clusters to the project team.

The client had several hierarchies and knowledge flow between the hierarchies was ineffective. Client_IM was the project manager for the project and this cluster had little clout on the other clusters in terms of getting their buy-in for the project and enthusing their involvement and participation. Client_User and Client_IT did not perceive a need for the system which meant lack of their involvement and hence knowledge contribution for knowledge integration. Client_IM said: “Client_Management agreed to the project, but the user departments were only passively involved in the project. Client_Management also never followed up with them.” Due to the lack of perception of strategic benefit, Client_Users did not use the system when it was ready thus affecting system assimilation for knowledge integration as well. Client_Management acceded to the project but was seldom involved in the project activities had very little communication and interaction with the other clusters so its commitment was not visible to the other clusters. Restructuring in the client further disrupted knowledge from some clusters since key people left. Communication with the vendor clusters was simpler since each had a lead representative who was the single point of contact to that cluster. The project team was collocated in the client’s office which did help in technical team communication between all those who were actively involved.

Although representatives were assigned from Client_User to streamline communication between the vendors and Client_Users, they seldom participated in the project activities and hence their knowledge contribution to the project was affected. Effectively it was only Client_IM that served as a link from the client to the project team. Vendor_Consulting said: “Although each of the vendors was assigned a user representative, they never participated in the project and were hardly present for meetings and so it was very difficult to get the knowledge needed from users.” Client_IT was also in the project team to assist in building the data warehouse but during system development it had issues in accepting the development methodology proposed by the vendors despite having a lot of experience with IT projects. The vendor’s attributed this to Client_IT’s lack of experience in building data warehouse which required a methodology different from standard systems development life cycle followed for normal IT applications. This friction between impeded knowledge integration since more time was spent in convincing each other. One of the vendors said: “Client_IT’s lack of experience in data warehouse meant lot of time convincing them on the development methodology and necessary process changes.”

Figure 2: Interaction of the knowledge clusters
If knowledge integration was affected due to the lack of common knowledge in terms of development methodology between Client_IT and the vendors, knowledge integration was also affected due to too much common knowledge. The lead representatives of all the three vendor clusters had over 20 years of data warehouse experience and were all equally passionate about data and this project. Though hired for expertise in a particular area the overlap in their knowledge bases made technical knowledge communication between them easier but also led to frequent clashes on the best way of doing things. Client_IM had to frequently resolve such clashes and admitted: “We needed all of their expertise, but unfortunately they came from different organizations and there was too much overlap of knowledge. There should be just enough common knowledge to have a shared understanding of things, but not overlapping expertise.” Knowledge integration was affected since more time was spent resolving issues, there was lack of knowledge dependency between the clusters, lack of appreciation for specialized knowledge base of other clusters and hence lack of harmony in the process.

Client_management intended to have the data warehouse within a certain time and budget but the vendor clusters were hired on a time and material contract that provided little motivation for them to speed-up the project when delays were imminent. Vendor_bankingsolution said: “Our contract does not state that we have to complete the project within the time frame.” Even when there were process changes in the client organization which meant delays Vendor_Systemsintegrator was unfazed: “It is alright for me if they have changes, the customer is the king, We are hired on time and material basis, so any delays does not affect us since we get paid for every man-hour we put in.” Knowledge integration suffered since the time and material contract on which the vendor clusters were engaged did not ensure a concerted effort on part of all clusters to work towards achieving effective knowledge integration in terms of finishing the project on time. This is in addition to ineffective knowledge integration due the lack of knowledge from some clusters reaching the project team.

5 DISCUSSION AND FINDINGS

In recent years organizational researchers have tried to examine multilevel theoretical perspectives for concepts such as creativity, learning and system usage (e.g. Burton-Jones and Gallivan 2007). Many knowledge management scholars (e.g. Merali 2001) also suggest the importance of considering knowledge management approaches across different organizational levels and also acknowledge (Nonaka and Takeuchi 1995) the difficulties in doing so. This is something evidenced in this study as well and as mentioned in the methodology section, the use and change of theory and analysis was subject to the data and preliminary analysis. Initial analysis for this study was conducted at the organizational level but many issues could not be explained very well. For instance, although a typical client vendor relation, where in one client hired three vendors for a project and motivation for all parties should be obvious, this was not witnessed and it could not be explained at the organizational level. The problem was the various clusters within the client organization had different perceptions towards the project. A cluster level analysis was then used and that could explain the behavior of the different groups; Client_User and Client_IT, toward knowledge integration. Based on this understanding of knowledge cluster and the case, inter-organizational knowledge integration is explained as a process that involves complementary specialized knowledge clusters interacting within certain structures and influenced by the clusters’ perception towards the project.

For the project, complementary specialized knowledge bases are identified from within the organization (e.g. users, IT departments) and outside the organization (the vendors). This is an important process considering organizations often have problems identifying the content, location, and use of the knowledge for software engineering (Inkpen and Tsang 2005). This is also consistent with the importance accorded to partner selection in alliances (Dodgson 1992) considering the length of time needed to build effective communication paths between organizations and the nature of knowledge to be shared. The concern is alike in long term collaborative IS projects. The case shows that in identifying the knowledge clusters, organizations base their selection in trying to find complementary knowledge bases that are needed for the project and make some effort to facilitate
exchange between the clusters by building a common knowledge base through social activities and team collocation, for instance. Common knowledge is the common understanding of a subject area (Demsetz 1991) that enables easier transfer, sharing or integration of knowledge (Alavi and Tiwana 2002). Although the three vendors were selected for their specialized knowledge in different domains, the representatives were all data warehouse experts and the common knowledge they shared did make knowledge exchange easier but also led to conflicts. Too much common knowledge is shown to be impeding for knowledge integration as much time was spent in resolving conflicts. The specialized knowledge bases of the vendors were not sufficiently complementary and hence dependency between the clusters was lacking. The lack of common knowledge between the vendor clusters and Client_IT in terms of system development methodology and between the vendor clusters and Client_Users in terms of understanding of data warehouse benefits also hampered knowledge integration. The case therefore suggests that when specialized knowledge of the clusters is sufficiently complementary there is a dependency between the clusters and more harmonious interaction for knowledge integration while at the same time common knowledge between the clusters enables a shared understanding of issues for knowledge integration but too much redundancy can also be ineffective. This emphasis on a balance between specialized and common knowledge is the contribution of this finding to existing literature.

The knowledge clusters interact within their own organizational structures (clusters from within an organization) as well as within the project structure (project team) and that serves as means through which their knowledge reaches the project. Gulati (1995) affirms there are usually no preexisting reporting relationships in an alliance or systems that serve as natural conduits for information and therefore structures have to be created for that purpose. This is typically the case in inter-organizational IS projects where the project team consists of members from different clusters across organizations and a reporting structure can be cross organizational and is created for the duration of the project. In the case, clusters within the client were all represented to the project team but the representatives were not actively involved in project activities (refer figure 2) and hence access to their knowledge was inefficient. Infrequent involvement of the Client_Management and Client_Users, lack of ties between them and Client_IM and restructuring in the client hampered their knowledge flow to the project. The structure was present but was ineffective in providing access to certain clusters’ knowledge thus blocking required knowledge to the project. The case shows that knowledge flows, enabled by the structures, should exist within the various knowledge clusters involved in the project and within the core project team. The main idea of the structures should be to provide a channel for knowledge from all the involved clusters to reach the project. Prior literature has stressed on communication channels for knowledge flows (e.g. Ravasi and Verona 2001). Grant (1996) argues that structures that provide organizations with a comparative advantage in managing the various knowledge processes are a critical strategic variable while Huang and Newell (2003) suggest that efficient collaboration in a team requires the explicit definition of the communication channels desired within the team. The distinction on the findings in this paper is the observation that when the structure allows effective communication between all the knowledge clusters and not just between organizations and simultaneously between the clusters and the core project team knowledge integration better facilitated.

The involvement of the knowledge clusters in the knowledge integration process is contingent on their motivation for the project determined by factors such as: their need for the system, effort and cost incurred, value for their business and contractual terms. Lack of motivation of the knowledge clusters affects their knowledge integration behavior in terms of lack of effort, involvement and knowledge contribution to the project. In this case, Client_Management intended to have the system within a certain time and budget but the vendors were hired on a time and material contract that provided them little incentive to hasten the project. Their goals were not synchronous and that led to ineffective knowledge integration due to lack of concerted effort on the part of all clusters to expedite the project as Client_IM confirmed: “The distinct KPIs of all four organizations affected the project.” Further, the lack of perception of benefits form the data warehouse affected Client_Users effort in contributing and assimilating knowledge for the project. Motivation in different forms has been identified as an influencing factor for several knowledge processes (e.g. Kankanhalli et al 2005) but the focus was on...
individuals or organizations. The contributions of this finding is the emphasis of motivation at the level of knowledge clusters as against at the individual/organizational level. Project benefits are usually mutual in a client-vendor relationship (client needs system; vendor needs business) but concerted effort towards knowledge integration (Huang and Newell 2003) is not always present as shown in the case because of differing perceptions among the knowledge clusters. The study also finds that contractual terms used for legal protection and as means to control behavior of partners (e.g. Choudhury and Sabherwal 2003) can help synchronize goals in the project towards effective knowledge integration.

The above discussion and understanding of inter-organizational knowledge integration are summed up in table 3 below as three findings.

<table>
<thead>
<tr>
<th>Findings</th>
<th>Contributions</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sufficiently complementary knowledge bases of the clusters ensure dependency and some common knowledge between them facilitates knowledge exchange for effective knowledge integration.</td>
<td>-Emphasizing cluster level complementarities rather than organizational level -Emphasis on balance between complementary and common knowledge between clusters</td>
<td>-Identify clusters such that there is complementarity/dependency between them -Measures to build/develop common knowledge</td>
</tr>
<tr>
<td>Structures instituted provide a platform for knowledge clusters to interact and a channel for their knowledge to reach the project.</td>
<td>-Knowledge based objective for structures -Cluster representation vs organizational representation</td>
<td>-Design structures to represent all clusters and provide channels for knowledge flows. -Keep structures stable</td>
</tr>
<tr>
<td>Knowledge integration is effective when there is concerted effort on part of the clusters toward the project.</td>
<td>-Knowledge cluster motivation -Importance of concerted effort</td>
<td>-Measures/incentives such as contract need to be instituted/phrased accordingly</td>
</tr>
</tbody>
</table>

Table 3: Findings for effective knowledge integration

6 IMPLICATIONS AND CONCLUSION

One of the key contributions of this study is the concept of knowledge cluster (figure 1) that affords conceptualization of the complex phenomenon of inter-organizational knowledge integration by abstracting the multiple levels (individual, group, organizational and inter-organizational) and their interactions in an inter-organizational context. The concept of knowledge clusters allows escalating issues at each level to a knowledge level, thus making it easier to comprehend the issues from a unified perspective. Knowledge clusters stem from and exist within the organizational structure thus encapsulating the inter-organizational dynamics, organizational and inter-departmental dynamics, as well as the dynamics of the individuals within the cluster. The concept is also a contribution to multi-level research and it addresses a call for such studies by Burton-Jones and Gallivan(2007). In a setting such as collaborative project involving knowledge integration it is essential to understand the correct behavior of organizations within the collaborative set-up. As the case shows it is at times hard to understand the behavior of the organization as such since the multiple clusters within behave differently so when grouped as clusters this behavior and hence possible solutions are clearer. This concept is of significance to knowledge integration literature in various inter-organizational contexts like joint ventures and alliances as well as to any knowledge intensive contexts like R&D and new product development. The understanding of inter-organizational knowledge integration through processes involving the knowledge clusters contribute to the underlying inter-organizational knowledge integration theory and go beyond those identified by earlier studies and restricted to intra-organizational contexts. Although the processes identified by Huang et al(2001) can be also be applicable to this context, they do not incorporate multiple dimensions critical to the
context (structural, motivational and cognitive) and aspects critical to the inter-organizational context (synchronizing goals of all clusters across the organizations).

The challenge in managing IS/IT projects, especially those involving multiple organizations, is well known and acknowledged. This paper provides a knowledge integration perspective to address the knowledge related challenges that traditional project management strategies did not address. This is essential since the goal of the project is to integrate the various knowledge bases. Based on this approach project management strategies can be developed to manage the knowledge clusters within the project environment to influence their knowledge integration behavior for the goal of the project. Specifically (column 3 in table 3), this study suggests that for effective knowledge integration, knowledge clusters for the project should be identified keeping in mind the complementary nature of their knowledge bases and providing mechanisms for fostering common knowledge between them.

Secondly, all clusters buy-in for the project should be obtained so as to ensure their concerted effort towards the project. Although this finding is not new, the implication is on focusing on the knowledge clusters rather than departments/organizations/individuals since it is the knowledge of the clusters that has been identified as essential for the project. Lastly, structures instituted should enable interaction between the clusters and effective knowledge flows.

This study was conducted in an Asian country and there was an element of conservatism shown by the interviewees in revealing data that may have affected some of the insights. To overcome this limitation, multiple people were interviewed on the same subject. The organizations also hesitated to share too many project related documents and to make up for this multiple people were interviewed to get oral confirmation of the data. It must also be noted that this study has the inherent limitations of a case study in terms of it being very context specific. The goal of this paper is therefore not to make generalizations applicable to all settings, but to be able to contribute to the underlying knowledge integration theory. The concept of knowledge clusters introduced in this study is novel and should be examined more rigorously especially their evolution and behaviour over the life cycle of the project and suitable structures for their interaction in different client-vendor settings. Another interesting aspect would be to explore their dynamics in specific inter-organizational arrangements (one client vs. one vendor, multi client vs. one vendor). The observations in this study are based on a single case and need to be further researched through questionnaires or more detailed case studies.

References


Gulati ,R.,(1995), Social Structure and Alliance Formation Patterns: A Longitudinal Analysis, Administrative Science Quarterly, 40, 619-652


CONCEPTUAL CONVERGENCES: POSITIONING INFORMATION SYSTEMS AMONG THE BUSINESS DISCIPLINES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0217.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Knowledge coordination, IS Philosphy, IS Journals, Convergence</td>
</tr>
</tbody>
</table>
CONCEPTUAL CONVERGENCES:  
POSITIONING INFORMATION SYSTEMS AMONG THE  
BUSINESS DISCIPLINES

Hovorka, Dirk S.; Faculty of Business, Technology, and Sustainable Development, Bond University, Gold Coast, AU, dhovorka@bond.edu.au

Larsen, Kai; Leeds School of Business, University of Colorado, Boulder, CO, USA, kai.larsen@colorado.edu

Monarchi, David; Professor – Information Systems (retired) University of Colorado, Boulder, CO, USA, david.e.monarchi@colorado.edu

Abstract

The structure and interrelationships of formal knowledge produced in the scientific disciplines have long been of interest to academics. One increasingly important domain of Information Systems (IS) research is the study of the creation and evolution of knowledge. Recent discourse about the intellectual structure of IS has revealed Latent Semantic Analysis (LSA) as an analytic technique that reduces problems associated with human categorization and citation analysis. This exploratory research positions IS within a ‘science of business’ using LSA to analyze semantic relationships in 24,841 abstracts from core business journals and begins to chart research concepts around which IS and other business disciplines converge. Results indicate that IS research has greater conceptual similarity to the disciplines of Management, Operations, Strategy, and Marketing than previously thought, and that these disciplines are converging on semantically similar research topics and concepts. This organizational-behavior-technical domain is distinct from an economics focused domain dominated by Finance and Accounting. The convergence suggests that IS is contributing to research and knowledge creation useful to other business disciplines and that strategic and functional dependence among the disciplines is increasing. This potentially leads to more integrated systems-oriented knowledge and greater practical relevance for both IS and other business disciplines.

Keywords: Business science, Latent Semantic Analysis, reputational networks, knowledge production, functional dependence, strategic dependence.

* corresponding author
1 INTRODUCTION

The discipline of Information Systems (IS) has a long history of research into the design and management of socio-technical systems for the capture, storage, dissemination, retrieval, and management of information and knowledge. One increasingly important domain of IS that has received less attention is the study of the creation and evolution of knowledge. Recent discourse regarding the intellectual structure and core concepts of IS (Larsen et al. 2008; Sidorova et al. 2008) has illuminated Latent Semantic Analysis (LSA) (Deerwester et al. 1990) as an analytic technique that reduces some of the problems associated with human categorization and citation analysis that has been observed in prior studies. Furthermore, studies of the status of IS as a reference discipline (Grover et al. 2006, Wade et al. 2006) examine the positioning of IS among related business and computer disciplines. But as Straub (2006 p. 242) notes: “scholarly endeavor should not draw artificial boundaries within the domain of information systems, and without any persuasive logic or reasoning, cut itself off from topics that are clearly intimately related to the practice of the IS profession.” In response, this paper uses LSA to analyze semantic relationships among 24,810 abstracts from a selection of core business journals, thereby broadening our perspective to encompass the context of the IS discipline among the business disciplines. We are not seeking to identify IS’s own structure or asking whether IS has become a reference for other disciplines, but instead we are beginning to chart the creation and evolution of concepts around which IS and other business disciplines orbit.

The structure and interrelationships of formal knowledge produced in the scientific disciplines have long been of interest to academicians. Formal knowledge derived from business school research has been seen increasingly as an economic resource (Machlup 1962; Whitley 2000). Questions regarding the basis of knowledge production, academic rewards, rigor, and relevance of research results have been raised by the business community, government, and business schools themselves. Business research is rarely considered holistically due to institutional structures that preclude integration (Campbell 1969; Whitley 2000). The production of knowledge has been dominated by natural categories “in terms of discipline (business economics, industrial sociology, occupational and organizational psychology) or functional orientation (operations, marketing, finance, HR)” (Tranfield 2002 p. 409). External professional associations and job markets are organized by discipline, and internally, most business schools are organized by discipline – with each discipline its own cost center. Often, each business discipline is presented pedagogically as ‘the most important’ to students.

An alternative view holds that because industry requires functional areas to be highly applied, functional, and integrated, business school research may emulate this model (Starkey and Madan 2001). Businesses are striving to create integrated value (arguably the basis for much of the reengineering revolution of the 1990s). To some extent this is the model followed by most business schools when designing undergraduate and MBA educations. If business school research is less fundamental in nature and more applied, then there should be a parallel between academic research and business in practice. However, in practice, whereas businesses serve customers who are not concerned about the academic reputation of a finance or marketing department, business school departments develop independent reputations and consider other departments in the same discipline as the primary consumers of their published research.

This exploratory research addresses the position of IS within a ‘science of business’ which is seeking to created integrated knowledge about business problems, practices, and opportunities. LSA was used to analyze the semantic content of 24,8141abstracts from top journals in the core areas of business as identified by Trieschmann et al. (2000), including Accounting, Finance, Information Systems, Management, Marketing, Operations, Insurance/Risk and Real Estate. Following Whitley’s (2000) framework of science as “the product of the social transformation of intellectually constructed objects” in a reputational system (p. 35), the functional and strategic dependence of disciplines was used as lens to examine the conceptual convergence of business disciplines over a thirty year timespan. An increasing degree of consensus on core themes and topics in the academic community would indicate the emergence of a paradigmatic ‘business science’ that would position the discipline of IS as moving toward integration. Divergence would indicate that academic business research is best represented as moving towards an adhocracy of stove-piped silos of knowledge, in which IS as a discipline is marginalized.
2 FROM CITATIONS TO CONCEPTS

Banville and Landrey (1989) suggest that the “preoccupation about the actual state and future evolution of MIS as a scientific field” (p 48) is a warranted and essential epistemological endeavor. Numerous studies that examine the structure, core issues, reference disciplines, and position of IS as a legitimate, mature discipline have been published (see Larsen et al. 2008 for a review). These studies have been widely influential in the discipline and most rely on citation analysis or human derived content analysis in attempts to capture the conceptual relationships in a large set of wide-ranging literature.

There have also been numerous bibliographic studies in the business disciplines that examine the emergence, growth and spread of ideas, which also trace changes in the intellectual structure of disciplines. Disciplines analyzed include Strategic Management (Ramos-Rodríguez and Ruíz-Navarro 2004), Operations (Pilkington and Liston-Heyes 1999), Management (Podsakoff et al. 2008), and Information Systems (Cheon et al. 1991; Culnan 1986), but generally these empirical studies do not involve a comparison among different disciplines. However, recent citation studies have examined IS as a reference discipline to other business disciplines (Grover et al. 2006; Wade et al. 2006) and determined interdisciplinary citation patterns (Biehl et al. 2005). Although citation studies can show patterns of knowledge movement, they generally do not address the specific content of the knowledge itself – the semantic relationships of research terms, concepts, and problems addressed by the disciplines in question.

How intellectual disciplines are organized, how they become mature, how they are controlled, and how they coordinate and orient research have been the focus of research on knowledge production. Scientific knowledge is increasingly seen as the product of social transformation of intellectually constructed objects, and scientific change is increasingly viewed as the outcome of social processes of negotiation, conflict, and competition (Whitley 2000) or “as a complex knowledge market, the constitution of which is, and has been, subject to shaping by institutional forces” (Ramiller et al. 2008, p. 5). These views are in contrast to Kuhn’s (1962) internalist view of mature sciences as being characterized by "uniformity and inevitability in knowledge development" (Whitley 2000 p. 3) rather than subject to external economic and political controls. In this latter perspective of formal knowledge production, disciplines are organized and controlled in different ways to produce different knowledge.

One basic type of knowledge production system is produced by political and organizational structures that compartmentalize university departments. Well-bounded and distinct departments support local reputational networks, determine employment criteria, and institutionalize training programs, which together lead to fragmented and internally-referential silos of knowledge (Shove 2000). A contrasting type of knowledge production is based in the social unit of the intellectual discipline in which moderately-bounded and distinct organizations "control and direct the conduct of research on particular topics...through the ability of their leaders to allocate rewards according to the merits of the intellectual contribution" (Whitley 2000 p. 7). Disciplines gain reputation through the production of novel research regarding concepts which have utility in other disciplines. The social systems (e.g. conferences, working groups, journals) provide coordination of task outcomes through access to rewards. Historically, the intellectual disciplines “have been very dominant in the organization of the science system, in the reward system, and in the career system...” (van den Besselaar and Heimeriks 2001, p. 1). Formal public communication and distribution channels (e.g. journals, conferences) provide the arena for conflicts over reputations as well as interpretations of the relative importance of research concepts and ideas. Selection of representative sample sets becomes a critical factor in determining the structure and relationships among and within disciplines. The different views of the IS discipline presented in Larsen et al. (2008) and Sidorova et al. (2008) show this quite clearly, with the former relying on a expansive set of journals (65 total) to examine the IS discipline, and the latter making the assumption that three high profile journals are representative of the core concepts in the IS discipline. The reputation of the high profile journals will influence researchers to pursue questions that align with the concepts that appear in those journals.

Therefore an important characteristic of reputation in intellectual disciplines is the degree of control the researcher has over work processes and research goals. Where research concepts, terms, and priorities are similar to commonsense ones, or are borrowed from other disciplines, it is more difficult to maintain unified control of research than in disciplines where vocabularies and work methods are distinct and arcane. In addition, highly selective distribution channels, result in greater dependence of researchers on
the gate-keepers of such channels, and thereby an increased disciplinary control of reputation and research direction. Finally, with increased diversity of legitimate audiences come increasingly differentiated research goals, leading to limitations on coordination and integration of intellectual priorities.

Rather than examining which disciplines cite other disciplines as reference sources, we examine the conceptual linkages among the disciplines to determine the position of IS in relation to other business disciplines. At one extreme, business science is comprised of isolated disciplinary silos, each with its own distribution channels, in which researchers do not utilize concepts from other disciplines and distribution channels limit the range of research topics, concepts, and problems they will accept. In this conception, reputations are forged within disciplines, and researchers depend less on other disciplines for coordinating integrated knowledge and building valid knowledge claims. At the other extreme, there is a “paradigm,” loosely defined as an understanding of shared topics, concepts, and models for the science of business that is recognized and accepted by its sub-disciplines. Researchers who often focus in one area are also observed in working in other areas as attempts are made to coordinate and expand research to increase the value of intellectual contributions.

Whitley's (2000) framework for reputation proposes that the major objective of scientific endeavor is couched within a reputational system. Progress in reputation is determined by the degree to which the contribution offers novelty, and the extent to which it is useful to others in the discipline (Figure 1). Conflicts between the two demands create particular tensions between scientists, and variations in their mutual balance affect the organization of knowledge. Structural characteristics of a discipline can influence the ways in which these objectives are sought and knowledge is produced. These foundation characteristics are: Strategic Dependence, Functional Dependence, and Strategic Task Uncertainty.

Our research focuses on two of these factors:

1. Strategic Dependence can be defined as the "extent to which researchers have to persuade colleagues of the significance and importance of their problem and approach in order to attain high reputation from them" (Whitley 2000 p. 88). Although this is a complex construct, for this study it was operationalized by determining the association of distribution channels. The correlations among the representative journals of each business discipline was based on the shared topics that were published (e.g. the amount of concept sharing between academic disciplines).

2. Functional Dependence can be defined as the "extent to which researchers have to use the specific results, ideas, and procedures of fellow researchers in order to construct knowledge claims that are regarded as competent and useful contributions" (Whitley 2000 p. 88). In this study, functional dependence was measured by the convergence/divergence of different business disciplines around specific research topics and ideas.

The science of business can be represented as a continuum between an integrated bureaucracy at one end, and a fragmented adhocracy at the other. An integrated bureaucracy would have high strategic dependence and high functional dependence between sub-disciplines. Positive indicators would include consensus around important concepts and topics for business research, and the valuation of opinions of researchers in other disciplines. If the science of business is a fragmented adhocracy, there would be low strategic dependence and low functional dependence among sub-disciplines. All sub-disciplines work in their own domain using different methods, focusing on different topics, and seeking internal reputations.

If the science of business is becoming a conceptually integrated bureaucracy, it is creating knowledge that is contextualized and shared among the disciplines. This would be recognized in the convergence of
disciplines around fundamental topics (functional dependence) and strong associations among the
distribution channels (strategic dependence). If, on the other hand, the science of business is a fragmented
adhocracy, then it is creating important and perhaps more fundamental knowledge within sub-disciplines
but is losing out on the contextualization and integration of knowledge between disciplines that is
necessary to create value. Prior citation analyses of business sub-disciplines (Cheon et al. 1991; Grover et
al. 2006; Wade et al. 2006) have shown a limited boundary-crossing of intellectual research topics among
sub-disciplines. The degree of fragmentation/integration also stimulates questions of research relevance.
The practice of business benefits from knowledge of integrated strategies and business functions that are
not separated by the artificial stovepipes that characterize academe (Campbell 1969). In a strongly applied
science such as business, research relevance can be increased by production of integrated knowledge that
views business activities from a systems perspective (Alter 2004).

3 METHOD

In contrast to prior scientometric studies based upon citation analysis, this study analyzed the co-
ocurrence and proximity of words as a measure of semantic similarity or meaning. We argue that this is a
semantic analysis, as it determines the similarity in meaning of “a word, phrase, sentence or text.” In
contrast to citation patterns as indicators of diffusion of ideas among disciplines, semantic analysis
emphasizes identifying the similarities and differences among disciplines regarding topics and concepts
presented in the research and envisioning a semantic network that clarifies the conceptual focus of research
in different business disciplines. Abstracts are a distilled and valid representation of the conceptual
meaning of research papers, and have been used in analysis of intellectual communities in the IS discipline
(Larsen et al. 2008). LSA is a computational technique that decomposes the semantic content of a textual
corpus (abstracts, in this study) into a numerical representation of the ‘meaning’ of each text unit. The
underlying theory of LSA is that the aggregate of all the word contexts in which a specific word does (and
does not) appear is a measure of the similarity of the meaning of words or of texts, to other words or texts
(Landauer 2007). This technique assumes that written/verbal meaning comes from the relationships that
are represented and activated by collections of words. Furthermore, it is these abstract relationships that
"make thinking, reasoning, and interpersonal communication possible" (Landauer 2007 p. 8) and the
relationships among the words in an abstract are precisely what the human, and the LSA technique, uses to
determine meaning. LSA is a method, although new to Information Systems, that is part of a long history
of computational research on semantic relationships, theory of meaning, and cognition (Jurafsky and
Martin 2000; Manning and Hinrich 1999) which assumes that human interpretation is not the only path for
representing human knowledge. LSA allows for rapid analysis of very large text-based data sets, while
minimizing many of the problems associated with human categorization. LSA offers: 1) the means to
handle very large data sets (24,000+ abstracts in this study) and 2) a consistent and replicable technique
that allows for comparison between researchers, data sets, and time horizons. This consistency provides the
ability to state that the differences (or similarities) in "meaning" we ascribe to the data are not due to
methodological or interpretative differences. If you reanalyze data with the same parameters, you will get
the same result. The converse is also true: if you use LSA with the same parameters on a different data set,
then any differences in meaning arise from the data, and not from the method or human interpretation.

3.1 DATA COLLECTION

As an exploration and proof of concept, this study used Trieschmann et al.’s (2000) determination of a
warranted set of A-journals from each business discipline, based on the proportion of academicians in each
discipline. They concluded that eight disciplines and 20 journals best represented the business disciplines:

Our goals were to develop a research "bread basket" of business school journals that was
representative of research in the 13 business school disciplines defined by AACS (1998). Within
each discipline, there are many "good" journals. For example, Glick, et al. (1997) and Johnson
& Podsakoff (1994) identify more than 30 "good" journals in the management group, which
accounts for about one fifth of business school faculty (3,457 out of 15,474). While this number

* http://wordnetweb.princeton.edu/perl/webwn?s=semantics
of journals is useful for evaluating one discipline, it becomes challenging to find and analyze the 150 or so "good" journals this number would imply for the business school as a whole... Instead, we decided to focus on a smaller set of the "best" journals in each discipline. Thus our measure of business school research productivity is deliberately biased to only one form of business research: publications in leading research journals. We exclude other forms of research such as articles in "good" but lower ranked journals, conference papers, books, book chapters, and articles in practitioner journals (Dennis 2000).

Our research utilized 24,841 abstracts from Trieschmann et al.'s (2000) list of 20 journals over a period of 30 years. Twelve of the journals provided full datasets beginning in 1973, and eight journals initiated publication between 1974 and 1990. Only two disciplines (Management Information Systems (1977), Production/Operations Management (1980)) began publication after 1973, but were included beginning with the first year of their publication. Abstracts were obtained from ProQuest's database of journal abstracts. Exhibit A lists the business disciplines and journals published by Trieschmann et al. (2000).

3.2 LATENT SEMANTIC ANALYSIS

This research represents the first application of Latent Semantic Analysis (LSA) to a longitudinal inter-field dataset, and the first empirical analysis of research topics within all the major business disciplines. LSA can be used to determine the conceptual similarity among text units (artifacts), is capable of handling large datasets, and aims to:

- develop a reproducible representation of artifacts (e.g., documents, interview data, survey data, etc.) and an approach to labeling that representation in a way that would (a) reduce problems of human interpretation of the data; and (b) allow the application of quantitative techniques based on cardinal, rather than ordinal or nominal, data. Such an advance would offer an alternative as well as a complement to some existing methods for categorizing and labeling qualitative data (Larsen and Monarchi 2004, p. 351).

LSA begins by treating each abstract as a set of words without structure. “Stop words,” or words that have little or no meaning when taken out of context, but that provide structure to the sentence, are removed. Stop words typically include articles, prepositions, pronouns, and conjunctions, as well as common adjectives and adverbs.

The remaining words are stemmed to avoid having multiple forms of a word (its morphological variants) represented in the analysis. Stemming converts a word to a related form, i.e. it “conflates” the word. For example, removing an “s” or “es” will convert some plurals to singulars, and stemming the words “walks,” “walking,” and “walked” reduces all three to “walk.” Stemming reduces processing time as well as increases identification of similar words.

After stemming, words reflecting research methods or techniques, such as “regression,” “correlation,” “ethnographic,” etc. are removed to allow the focus to be completely on the topics of interest rather than the specifics of research approaches. Although the meaning of objects is constituted by the epistemological perspective, we assumed that the contextual similarity of concepts would still be present. A sparse matrix of unique stems as rows and abstracts as columns with the number of occurrences of a specific stem in a specific abstract as the cell value is created next. This matrix is submitted to a singular value decomposition (SVD), which creates a high-dimensional space in which each abstract and each stem in those abstracts occupies a specific location. Furthermore, by aggregating the SVDs of the abstracts, a centroid for those abstracts may be located in the high-dimensional space representing that collection of abstracts (for example, all abstracts published in a specific journal over 30 years). This representation allows measurement of distance (or cosine or angle) between individual abstracts in all disciplines to the centroid representing a specific discipline.

3.3 DATA ANALYSIS

Data analysis consisted of two steps. First, the 30 years of abstracts in Trieschmann et al.’s list of journals was subjected to LSA analysis. Once each abstract was represented as a vector, the midpoint for all abstracts in each journal was calculated, and the "distance" from each of the 24,841 abstracts to every one
of the 20 journal midpoints was measured. The results were compiled in a table of 20 columns and ~24K rows which was used in a factor analysis using principal component analysis with varimax rotation. As a test of the association of distribution channels (strategic dependence), a correlation table between the semantic content of all journals was calculated and the resultant variance used to map a semantic network.

Second, functional dependence was examined by calculating the relative distance in semantic space of research topics which characterize the journals from each discipline over time. A sliding window of a 3 year period, centered on the target year, was used to select the abstracts for each discipline within that window. The centroid of each discipline’s abstracts for a given period was plotted against the centroid for each of the other disciplines for the same time period. This pairwise comparison reveals relative convergence/divergence of concepts in the journals from each discipline over time.

4 FINDINGS

The analyses revealed two surprising results: 1) The distribution channels for the business disciplines are semantically more closely related than prior studies indicate, and 2) within the business disciplines, there exists a significant degree of conceptual convergence on two distinct domains of topics and concepts.

4.1 ASSOCIATION OF DISTRIBUTION CHANNELS

As a test of the association of distribution channels (journals) in this sample, semantic analysis of the total aggregate set of abstracts for each journal was performed to find the centroid for each journal. These results were then submitted to two subsequent analyses. First, multi-dimensional scaling (MDS) of the distances was calculated to determine the relative proximity of the journals centroids (the S-stress score was small, indicating that the resulting two-dimensional map is quite accurate). The relative positions of the journal centroids were construed to represent the relationships among the disciplines they represent (Figure 2). This mapping suggests the formation of two distinct domains of business interest based on conceptual csimilarity: a organizational-behavioral-technical domain composed of Management, Strategy, IS, Marketing and Operations, and an economics domain composed of Finance, Accounting, Real Estate and Risk & Insurance. This result is in remarkable agreement with the structural equivalence map based on citation analysis of the Financial Times basket of business journals (Biehl, 2001).

Second, using the same dataset as for the factor analysis, correlation coefficients between journals were calculated. Using the significant correlations (after Bonferroni correction) between journals in the different disciplines, a diagram of the network of interdisciplinary relationships between journals, and again by extension between disciplines, was developed and overlaid onto the MDS perceptual map (Figure 2). In this semantic network the network nodes are the centroids of each journal and the weight of the network links is determined by the correlations between journal abstracts. As the journals themselves are representative of the disciplines, the resultant relationships articulate "the groups…who share common understanding, those who have idiosyncratic meanings…and those who serve as liaisons and boundary spanners (Monge and Contractor 2003, p. 187).
The semantic network suggests that over the 30-year span of journals included in this study, the business disciplines as represented by these journals are more tightly linked regarding the research topics and concepts they publish than previously thought. The network map shows two distinct domains that exhibit a significant degree of shared interest and consensus about important research concepts, topics, and problems as evidenced by the proximity of research articles in semantic space. By showing the semantic associations between journals, we can visualize the strategic dependence between the disciplines the journals represent. The line weighting of the ties indicates the percent variance for which each tie accounts (from 10-50%). The line weighting is therefore an indicator of the relative strength of the semantic association between journals. Overall, this network map indicates that there is a significant degree of strategic dependence within these two discrete business domains. The strongest associations are between Strategy and Management, IS and Management, and between Finance and Accounting. At lower degrees of association there are significant semantic relationships between Management and Information Systems, IS and Operations, and Marketing and Operations. Interestingly, Real Estate and Risk & Insurance are relative isolates with weaker associations with Finance.

4.2 POSITIONING IS AMONG BUSINESS DISCIPLINES

The final analysis reveals the changes in semantic content of each business discipline over time in comparison to each other discipline (Figure 4 A-F). An increase in the semantic distance (the angle) between the target discipline (listed first) and any other specific discipline indicates that the topics, and concepts in the journals are diverging. A decrease in distance indicates greater convergence on semantically related concepts. As words pertaining to research methods, epistemologies, and ontologies were removed from the analysis, any convergence/divergence of the centroids representing the discipline reflects changes in the meanings of the abstracts.

Given the limitations of representing discipline centroids in a two-dimensional semantic space, for the purpose of displaying the data one discipline must be used as a baseline to be compared to each other discipline. In Figure 3 A-F, the baseline discipline is noted in the figure heading and does not appear in the figure. We found that Real Estate and Risk/Insurance were very insular and did not converge or diverge with any other discipline. Therefore, for clarity, they were not included in these figures. Lastly, only significant relationships have been displayed in each figure.

Figure 3A shows that the semantic content contained in IS abstracts is converging with the semantic content in the abstracts of the other six disciplines, with the exception of Accounting, which is diverging in
semantic content (i.e., the "distance" in semantic space is becoming greater). An analogous pattern is observed when Accounting is held as the baseline (Figure 3B). In this figure, we observe that all business disciplines are diverging in semantic content from Accounting with the exception of Finance, which is converging.

Surprisingly, this pattern holds when each of the other business disciplines is held as the baseline - as a group, Operations, IS, Management, Strategy, and Marketing are converging with each other. Accounting and Finance are also converging with each other but are simultaneously becoming more distant from the organizational-behavioral-technical domain. Although there are annual oscillatory variations, in each case the trend lines are converging at significant values.

<table>
<thead>
<tr>
<th>A: Information Systems vs. all other disciplines</th>
<th>B. Accounting vs. all other disciplines</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Graph A" /></td>
<td><img src="image" alt="Graph B" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C. Operations vs. all other disciplines</th>
<th>D. Strategy vs. all other disciplines</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Graph C" /></td>
<td><img src="image" alt="Graph D" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E. Management vs. all other disciplines</th>
<th>F: Marketing vs. all other disciplines</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Graph E" /></td>
<td><img src="image" alt="Graph F" /></td>
</tr>
</tbody>
</table>

**Figure 3 A-F Comparison of Conceptual Trends in Business Journals**

Examination of these figures show that in the formative years of the disciplines much of the research published in these journals was semantically less similar regarding the research topics, concepts and problems than they were in recent publications. Although there is no baseline distance in semantic space
that serves as a demarcation between fragmentation and integration, these patterns show that the business disciplines were conceptually more fragmented 30 years ago than now and that in general, there is increasing agreement on conceptually related concerns and problems.

It is clear that the business disciplines are increasingly focusing on semantically similar research topics, concepts, and problems over time. There are two discrete but associated "domains" within business science: a convergence on financial concepts (Finance and Accounting, with weaker associations with Real Estate and Risk & Insurance); and a convergence around strategic, managerial, and technical concepts. These large-scale domains indicate a growing degree of functional dependence among sets of the previous more fragmented business disciplines. In addition, if research topics are held in common more frequently, there must be a greater association among distribution channels over time. This is a logical argument, rather than one with direct empirical support. But it suggests a greater degree of strategic dependence among the business disciplines within the separate domains.

Prior citation studies of business disciplines have shown that “the management field is becoming more integrated and interdisciplinary” (Biehl et al, 2006, p. 363) while other disciplines are becoming more insular. Our evidence suggests a broader pattern of conceptual integration and a long-term trend. It is important to note that semantic similarity is a continuum, not a point of demarcation. Our study indicates a greater degree of conceptual similarity among IS and other business disciplines than has been observed in prior citation studies. It should be noted that citation analysis may not always indicate consensus on the concepts under inquiry, but rather may reflect issues of method, history, or areas of distinction. Our semantic analysis of concepts shows that the research concepts published in top IS journals are similar to research in highly ranked journals in Management and Operations and to a lesser degree, Marketing and Strategy.

4.3 LIMITATIONS

This is an exploratory study and the set of business journals included do not necessarily capture the full extent of research topics and concepts in the business disciplines. For example, as shown in Larsen et al. (2008), the conceptual breadth of IS studies is not fully represented by MISQ and ISR. As noted in Straub (2006) the choice of journals will have a strong effect on any study of disciplinary structure and relationships. Additionally, the “basket of words” approach used in LSA flattens the semantic content of abstracts to a single point and may uniquely categorize studies that address multiple research terms. But these findings show long-lived trends of conceptual change in the research topics and concepts published in these highly ranked journals. Future research should include a broader data set which includes European and Asia-Pacific journals to determine whether these changes hold across the disciplines globally. In addition, the concepts or themes around which the journals publications are converging should be identified.

5 DISCUSSION

This study analyzed the semantic content of abstracts in a warranted set of representative journals of the eight business disciplines identified by Trieschmann et al. (2000). By determining the proximity of the aggregated semantic content that contributes most to the position of a journal in semantic space, it is possible to show conceptual relationships among journals and the disciplines that they represent. By examining the actual terms and concepts underlying research, this research shows that two domains of business disciplines are more related, both conceptually and by distribution linkages, than previously thought. In one domain, Management holds a central position between IS and Strategy, and IS has centrality between Operations and Management. In the other large-scale business domain, Finance and Accounting research concepts are tightly intertwined with a lesser shared interest in some concepts within Real Estate and Risk & Insurance.

This study also highlights the importance of longitudinal analysis in scientometric research. A sliding window analysis clearly demonstrates conceptual convergence semantically related topics among journals in the organizational-behavioral-technical with a simultaneous divergence from an accounting/finance domain. The use of LSA was critical, in that it allows consistent semantic analysis of large data sets and enables comparison across time periods.
The focus on concepts and research problems shared with other business disciplines potentially leads to more integrated knowledge and greater practical relevance for IS. It is important to recognize the importance of this finding without overstating it. This does not indicate the direction or timing of boundary spanning knowledge, and no claims are made about IS as a reference discipline. But at the same time, it does suggest an increasing centrality of IS concepts as a contribution to knowledge regarding a broad set of problems that other business disciplines are concerned with.

As noted by Biehl et al. (2001), “PhD students are trained to focus on a single discipline...tenure is easier to obtain in a fairly focused research stream...and publishing in a variety of cross disciplinary journals is often seen in a negative light” (p. 369). But despite these institutional structures and the contrasting result from interdisciplinary citation studies (Biehl et al. 2005; Grover et al. 2006; Wade et al. 2006) these data suggest that IS and business disciplines are trending towards greater conceptual integration. Research, particularly in IS, is not as much transdisciplinary as neo-disciplinary, and is at risk of being marginalized by traditionally bounded business disciplines concerned with the same concepts. Although there is evidence that “business academics tend to publish in discrete and mostly non-overlapping disciplines” (Biehl et al. 2005, p 368), the disciplines themselves are converging on semantically, and arguably conceptually, similar research ideas, concepts, and problems.

Although the sources and direction of knowledge movement among disciplines is not apparent, this research suggests that IS is more conceptually related to other business disciplines in an organizational-behavioral-technical domain than prior studies would indicate. This implies that IS is becoming a central part of a maturing ‘science of business’ that has a network of functional and strategic dependencies that may lead to interdisciplinary perspectives on common research problems. These reputational networks can influence the institutional direction of research funding and programs. The allocation of subject domains to the sciences is not due to clear demarcations of content. Instead academic disciplines themselves exist due to a wide variety of internal, external, and historical forces, and reinforce ethnocentrism due to “the tribalism or nationalism or in group partisanship in the internal and external relations of university departments, national scientific organizations and academic disciplines” (Campbell 1969 p. 328).

Recognition of the role played by IS in producing integrated knowledge has implications for collective coordination of task outcomes and for academic reward structures. Supporting, encouraging, and rewarding the trend toward interdisciplinary studies may have the effect of aligning research with the institutional and economic forces that drive research. This may increase both integrated systems-based knowledge and the relevance of academic knowledge to practice.

References


**Exhibit A**

<table>
<thead>
<tr>
<th>Table 1 Trieschmann et al (2000) Academic Business School Fields</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fields and Journals</strong></td>
</tr>
<tr>
<td>Field 1: Accounting</td>
</tr>
<tr>
<td>Journal of Accounting &amp; Economics [JAE]</td>
</tr>
<tr>
<td>Field 2: Finance</td>
</tr>
<tr>
<td>Field 3: Insurance, International Business &amp; Real Estate</td>
</tr>
<tr>
<td>Field</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
AN EXPLORATION OF THE REAL OR IMAGINED CONSEQUENCES OF INFORMATION SYSTEMS RESEARCH FOR PRACTICE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0463.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>knowledge transfer, Bibliometrics, Research methodology, Practice</td>
</tr>
</tbody>
</table>
AN EXPLORATION OF THE REAL OR IMAGINED CONSEQUENCES OF INFORMATION SYSTEMS RESEARCH FOR PRACTICE

Avison, David, ESSEC, Cergy-Pontoise, France, avison@essec.fr
Gregor, Shirley, Australian National University, Canberra ACT 0200, Australia, Shirley.Gregor@anu.edu.au

Abstract

In this paper we attempt to assess the impact of IS research on practice. It is important to show a positive impact, if we can, so that academia, prospective students and the outside world more generally sees that the IS discipline has made a positive and major impact on organizations, people and society. Our research fills a gap as there has been little or no work to date that systematically assesses the effective relevance of IS research to practice. We find that the use of bibliometric indicators to trace the flow of ideas from research published in our top journals to practitioner journals does not detect any strong link between IS research and practice. We suggest several alternative means of assessing effective, rather than imagined, relevance and also consider how more effective diffusion of our IS research might be encouraged.

Keywords: Information systems, impact, research, practice, bibliometric indicators
1 INTRODUCTION

The positive impact of using information systems (IS) in practice is well demonstrated, with studies of the business value of information technology (IT) showing contributions to competitive advantage and productivity (see for example, Dedrick et al., 2003). However, the role that IS academic research plays in the adoption and effective use of IT is not clear. We would like to think that IS research has relevance and contributes to practice so as to be of benefit to individuals, organizations and society. Moreover, government research funding bodies are increasingly asking that external impact be implicated in research proposals (for example, European Union, 2008; National Science Foundation, 2008).

The impact of IS research on other research disciplines is evidenced by Baskerville and Myers (2002), and is extremely important, but this does not evidence impact on practice. Many research papers make claims that the work described is ‘relevant to practice’: for example, many of those describing action research, case studies and organization surveys. However, making this statement is not enough. In a discipline like IS, as in many applied disciplines, it is vital that we demonstrate that our research is actually being used by practitioners and so increase our claims to legitimacy.

In this paper, therefore, we address the question: Does IS academic research have demonstrable consequences for practice concerned with the use and implementation of IT? In answering this question, we found we also had to address a secondary question: How can we show that ideas and knowledge flow between IS researchers and professional practice?

The motivation for addressing these questions is strong. It is not enough to espouse relevance: it must be demonstrated. If we show that IS research has an impact on practice, then we achieve several useful ends. We can, for example:

1. Use the evidence as a weapon to show academia and the outside world that the IS discipline (not just technology) has made a positive and major impact on organizations, people and society.
2. Provide our students with a context which demonstrates the impact of IS research on organizations, people and society so that they appreciate its present and its potential impacts and thus see the study of IS as potentially worthwhile.
3. Demonstrate that we have means for assessing the flow of ideas from IS research to practice that can be used for benchmarking and for making arguments to funding bodies.

In contrast, if it appears that little of our research ideas flow through to practice, then we believe we are being alerted to a situation which our community should make considerable effort to address. If we cannot provide convincing evidence of how ideas flow between academia and practice, because there are no obvious means to do so, then we need to devise suitable indicators and measures.

We begin from an underlying position that IS academic research should have practical relevance, as IS is an applied discipline; a view recognized by our journals when they ask that articles indicate potential practical significance. Our work is significant because we could find no prior work that attempted to answer our questions for IS academic research, research that is published in our leading journals.

When we set out to address our primary question, we did not know the answer! Was our research being used by practitioners? However, even more interesting perhaps, we did not know how we could provide the answer. How could we demonstrate that our research was having such an impact (or not)? This paper discusses our exploration as we began to address these questions.
2 BEGINNING OUR EXPLORATION: SEARCHING FOR EFFECTIVE RELEVANCE AND KNOWLEDGE TRANSLATION

Benbasat and Zmud (1999, p 13) argued that ‘the IS field does not possess the evidence with which to illustrate the impact of its research …. This is an important question that academics should investigate’. Almost ten years later, we also could not find research that addressed this question. Rosemann and Vessey (2008) make suggestions on how the relevance of IS research can be improved, specifically through using what they call applicability checks when research projects originate. These authors argue that there is “limited demand on the part of practice for the outcomes of IS academic research” (p. 2). However, the only evidence they give for this view is that the numbers of practitioners who subscribe to IS journals or who attend academic conferences such as the International Conference in Information Systems (ICIS) are rather low.

To address our primary question, we turned to a familiar debate in IS to continue our exploration, that of rigor versus relevance (for example, see Agarwal & Lucas 2005, Benbasat & Zmud, 1999). We did not do so to join in this debate ourselves, but to see if there was evidence of what we call effective relevance as embodying some demonstrable evidence of practitioner take-up as against the mere suggestion that the research ‘should be’ or ‘might be’ relevant (as often claimed in the conclusion sections of papers). Despite the lengthy debate of rigor versus relevance there has been little demonstration of effective relevance in this body of literature nor has there been debate about how it can be demonstrated.

The research impact of IS papers has been assessed through bibliometric methods both with regards to the impact of IS researchers within the discipline and also to researchers in other disciplines (Baskerville & Myers, 2002, Lowry et al. 2007). But we are addressing a different question when we ask about the impact on practice, that is, the flow and reception of ideas or knowledge between IS academic and practitioner communities. This issue is one of knowledge translation or transfer which “assumes that practical knowledge (knowledge of how to do things) in many professional domains derives at least in part from scientific knowledge” (Van de Ven 2007, p. 3).

In some related areas where a more distinct artifact is produced, measures of impact such as patentable ideas or start-up companies can be used. In IS, however, the products of research are likely to be less tangible outputs such as methods for IS development or the evaluation of applications, the translation of which from research to practice are more difficult to track.

The problem of the gap between practitioner and academic communities in terms of knowledge creation and transfer was addressed in a special issue of the Academy of Management Review (2001, 14, 2). In introducing this issue, Rynes et al. (2001, p. 340) note that “a substantial body of evidence suggests that executives typically do not turn to academics or academic research findings in developing management strategies and practice”. Some observers have suggested that academics and practitioners inhabit basically different communities, with different frames of reference and different ways of sense-making (Shrivasta & Mitroff, 1984). Further, Rynes et al. (2001) note the paucity of empirical work on the topic of knowledge transfer between academics and practitioners, particularly in the organizational sciences.

There are many ways in which knowledge transfer from research to practice can occur, through textbooks, practitioner-oriented books and publications, teaching, consulting, executive education, and personal links. In this exploratory study, however, we have chosen to study knowledge transfer by investigating how academic research results flow to trade journals and practitioner magazines. We recognize that this is only one of the approaches that could be employed. In defence of our choice, ideas in practitioner journals have been shown to be an important means for the diffusion of knowledge (Rogers, 1995, Van Steijn & Rip 1988) and there is a substantial literature that points to the flow of ideas and knowledge through communication channels as crucial for innovations and innovative behavior (see Rogers, 1995). To give one example, Nederhof and Meijer (1995) argue that
Trade journals are the most single important source of knowledge for farmers and horticulturists. Further, Spencer (2001), in looking at the extent to which corporate and academic researchers in the flat panel display industry exchanged knowledge, noted the importance of publishing in practitioner journals if knowledge transfer is to occur.

To sum up, in our study we chose to explore whether the impact of IS research on practice could be demonstrated by looking at the use of IS academic research in outlets with a professional readership. We explain our approach to achieve this goal in the next section.

3 EXTENDING OUR EXPLORATION: SEARCHING FOR IMPACT OF IS ACADEMIC RESEARCH IN PRACTITIONER JOURNALS

We chose to explore the impact of IS research in practitioner journals by extending traditional bibliometric analyses, as these methods can be applied in a systematic fashion. The method that we adopted was to examine the extent to which a selection of high-profile ‘practitioner’ journals referenced IS research in 2006. This method involved four inter-related steps, none of which were carried out without difficulty, and all these required some degree of iterative exploration and judgment. The four steps were as follows:

1. Choice of the sources of IS research that were to be included.
2. Choice of the high-profile practitioner journals that were to be examined for evidence of references to IS research; that is, referencing the sources selected in (1).
3. Choice of a database that could provide evidence of the linkages between the IS research in (1) to the practitioner journals in (2).
4. Analysis of the linkages between IS research and the practitioner journals in 2006, using the database from (3).

Each of these steps is described in more detail below:

3.1 Choice of sources representing high-quality IS research

We decided to use the ‘basket’ of journals selected by the Association for Information Systems (AIS) as being representative of high-quality research devoted specifically to the information systems field. This basket comprises (in alphabetical order) European Journal of Information Systems (EJIS), Information Systems Journal (ISJ), Information Systems Research (ISR), Journal of the Association for Information Systems (JAIS), Journal of Management Information Systems (JMIS) and Management Information Systems Quarterly (MISQ). Unfortunately, JAIS had to be omitted because it is a relatively recent addition to the Web of Science/ISI database and citation data for it could not be obtained. Although data was difficult to obtain, our understanding is that only MISQ (with around 3000 subscribers) and ISR (with around 2000 subscribers) have substantial subscriptions for their paper editions (of the basket of six, only JAIS is electronic only). Table 1 shows details of the five journals included in the analysis.

<table>
<thead>
<tr>
<th>Journal Title</th>
<th>Abbreviated Title</th>
<th>ISSN</th>
<th>Impact Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>European Journal of Information Systems</td>
<td>EJIS</td>
<td>0960-085X</td>
<td>0.862</td>
</tr>
<tr>
<td>Information Systems Journal</td>
<td>ISJ</td>
<td>1350-1917</td>
<td>1.543</td>
</tr>
<tr>
<td>Information Systems Research</td>
<td>ISR</td>
<td>1047-7047</td>
<td>2.537</td>
</tr>
<tr>
<td>Journal of Management Information Systems</td>
<td>JMIS</td>
<td>0742-1222</td>
<td>1.818</td>
</tr>
<tr>
<td>MIS Quarterly</td>
<td>MISQ</td>
<td>0276-7783</td>
<td>4.731</td>
</tr>
</tbody>
</table>

Table 1: Selected High Quality IS Research Journals
3.2 Choice of high-profile practitioner journals and magazines

Table 2 shows the practitioner journals and magazines selected for inclusion in our analysis. These are hereafter referred to as practitioner journals for the sake of convenience. The sample is a judgment sample from journals of interest to IS that met the following criteria:

- They appear in the chosen database (the Web of Science/ISI);
- They are mainly aimed at practitioners;
- They have relatively high circulation counts, and thus a potentially large readership (figures taken from Ulrich’s Periodicals Directory 2008).

Some judgment was required to identify journals that are aimed at least in part at practitioners. Ulrich’s Periodicals Directory makes a distinction between ‘scholarly/academic’ journals and trade journals, but it was found not to be useful for our purposes, as its categorization appeared to be problematic. For example, *Communications of the ACM (CACM)* is shown as an ‘academic/scholarly’ journal in Ulrich’s Periodicals Directory, but the Association for Computing Machinery (ACM) itself refers to this publication as a ‘magazine’. Rather than using Ulrich’s, therefore, we examined the home pages of journals that could be expected to reference IS research. We made a decision as to whether a journal could be classed as having a practitioner audience by considering its stated aims on its website. The circulation count was also taken into account, as this is evidence of reaching an audience beyond academia.

Only journals with a circulation count of 5000 or above were included. The result of this process was the journals shown in Table 2.

<table>
<thead>
<tr>
<th>Journal Title</th>
<th>Abbreviated Title</th>
<th>ISSN</th>
<th>Impact Factor</th>
<th>Circulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communications of the ACM</td>
<td>CACM</td>
<td>0001-0782</td>
<td>1.509</td>
<td>85,000</td>
</tr>
<tr>
<td>California Management Review</td>
<td>CMR</td>
<td>0008-1256</td>
<td>1.429</td>
<td>6,500</td>
</tr>
<tr>
<td>Harvard Business Review</td>
<td>HBR</td>
<td>0017-8012</td>
<td>1.505</td>
<td>250,000</td>
</tr>
<tr>
<td>IBM Systems Journal</td>
<td>IBM</td>
<td>0018-8670</td>
<td>0.747</td>
<td>30,000</td>
</tr>
<tr>
<td>Journal of the American Medical Informatics Association</td>
<td>JAMIA</td>
<td>1067-5027</td>
<td>3.979</td>
<td>5,000</td>
</tr>
<tr>
<td>MIT Sloan Management Review</td>
<td>Sloan</td>
<td>1532-9194</td>
<td>0.888</td>
<td>25,000</td>
</tr>
</tbody>
</table>

Table 2: Selected High-Profile Practitioner Journals/Magazines

3.3 Choice of Citation Database

The two primary databases utilized in our analyses were the *Web of Science* (2008) and *Scopus* (2008). The Web of Science refers to services offered by Thomson Scientific and was previously known under the label of the Institute for Scientific Information (ISI). This database was used as the primary source as it is well-known and used frequently for bibliometric and citation analyses. It was supplemented by data from the Scopus database for some analysis where Scopus provided the data in a more convenient or complete form.

The Thomson/ISI database has some drawbacks. Clarke (2008) has pointed out some deficiencies with respect to IS, including the omission of about 60% of the core body of IS publications, patchiness of data and some serious anomalies. The database also requires some expertise to use and has some peculiarities. Specific journals can be hard to find, even using the international standard serial number (ISSN), which was found to be the most reliable search criterion. Some journals are only in the Social Science Edition (e.g. *ISJ* and *ISR*), some only in the Science Edition (e.g. *EJIS*) but some are in both (e.g. *MISQ*).
In a final step we used Google Scholar (2008) for some further tentative analysis. Google Scholar is an experimental service, still in ‘beta’ release in 2008, but has a more extensive collection and the powerful Google search engine (Clarke, 2008). Unless otherwise noted, however, the data used is from the Web of Science.

3.4 Analysis

Each of the five leading IS journals were examined in turn and the number of citations by the selected practitioner journals in 2006 were noted, using the Journal Citations Report facility in the Web of Science database. Note that all ‘articles’ shown in the database were included, even though these will include some editorials, reviews and similar material. Table 3 shows the results of this analysis.

It can be seen that there are relatively few citations of IS research in the practitioner journals, both in absolute terms and as a percentage of the total number of citations of the IS research articles. As an example, for MISQ, the table shows that there were a total of 49 citations of MISQ articles in the selected practitioner journals in 2006. However, in 2006 the total number of times that MISQ articles were cited by all journals in the Web of Science database was 3186. The last line in Table 3 shows the percentage of citations for the journal that were in the selected practitioner journals compared with the total citations for that journal. So, for MISQ, just 0.02% of its citations were in the selected practitioner outlets. In total, there were only 80 citations in the practitioner journals in 2006 that formally referenced any work in the leading IS research journals. These 80 citations represented, on average, less than 0.01% of the total 9183 citations to the IS research journals.

An alternate way at looking at the data is to study the two right-hand columns. Many of the practitioner journals do not contain a large number of formal references, so these columns give an indication of how often IS journals are cited by the practitioner outlets relative to their overall citations. Thus, we can see that in 2006, CACM made 1658 citations of other work, with 38 (2.29%) being to the basket of IS journals.

<table>
<thead>
<tr>
<th>Cited by</th>
<th>Source</th>
<th>Citations of IS journals</th>
<th>All citations</th>
<th>% IS journals</th>
</tr>
</thead>
<tbody>
<tr>
<td>CACM</td>
<td>EJIS</td>
<td>2</td>
<td>1658</td>
<td>2.29</td>
</tr>
<tr>
<td></td>
<td>ISJ</td>
<td>2</td>
<td>1054</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>ISR</td>
<td>8</td>
<td>7</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>JMIS</td>
<td>7</td>
<td>50</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td>MISQ</td>
<td>19</td>
<td>1206</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>HBR</td>
<td>-</td>
<td>3224</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>IBM</td>
<td>-</td>
<td>474</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>JAMIA</td>
<td>-</td>
<td>7666</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Sloan</td>
<td>-</td>
<td>49</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Practitioner journal citations</td>
<td>2</td>
<td>80</td>
<td>0.01</td>
</tr>
<tr>
<td>All citations</td>
<td>EJIS</td>
<td>439</td>
<td>9183</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>ISJ</td>
<td>290</td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>ISR</td>
<td>1508</td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>JMIS</td>
<td>1523</td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>MISQ</td>
<td>3186</td>
<td></td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 3 Number of citations of leading IS journals (all years) by selected practitioner journals in 2006

Our analysis is not complete as it does not include all IS research articles or all practitioner journals. However, we have selected leading research outlets and practitioner journals where IS work could be expected to appear, so, if anything, this result may overstate the proportion of IS research being cited in practitioner journals, as against citations in other research outlets.

As the number of cited papers is rather small, we provide an example of which papers are cited in practitioner journals in Appendix A. This appendix shows the articles from MISQ cited in CACM in 2006 (18 articles using the Scopus database). Of these 18 articles only 4 were published after the year 2000, so they are hardly ‘cutting edge’. What is happening? It may be that the articles in CACM are
using the IS journals to provide evidence of scholarly merit rather than reporting on new and exciting work.

As we have minimal existing data to use as a benchmark, it is difficult to say if this result for IS is markedly different from that of other academic disciplines. However, Nederhof and Meijer (1995) report that in the humanities, a large part (sometimes more than 50%) of the research products is directed at external knowledge transfer, while in some science fields, such as experimental psychology, a more ‘modest’ percentage (12%) of the output is directed to a non-scientific public (Nederhof et al., 1989). Obtaining extensive data for more directly comparable analysis was beyond the scope of our study. We did, however, look further at one of the practitioner journals, CMR, which publishes across a range of disciplines. In 2007, CMR had in total 1277 citations to works in other journals. Analysis classified 242 of the citations as being to research journals, with a breakdown by journal discipline of: 162 Management; 30 Marketing; 23 Psychology, 11 Sociology; 10 Economics, 6 Information Systems, 5 Planning; 3 Law and 2 International Business. IS does not show up well in this comparison.

Given these rather disappointing results, we continued the exploration along some other lines, as shown in the next section.

4 EXTENDING OUR EXPLORATION FURTHER: SEARCHING FOR LOOSER LINKS

The analysis above indicates a rather small proportion of our research flowing directly through to practitioner outlets from our leading journals. However, we explored further, looking for looser links such as what we refer to as additional flow-on effects and non-obvious connections. By additional flow-on effects, we refer to the dispersal effects that occur when articles that appear in practitioner outlets which make reference to IS research journal output are themselves referenced by other practitioner journals. Of course, the latter may not reference those aspects of the intermediary article that discuss the original IS research, so the effect might be very limited. By non-obvious connections we refer to influential IS research that may appear in practitioner outlets without the connection to the underlying IS research being obvious: that is, it may not appear in formal citations.

Each of these aspects is explored in turn, using the same practitioner journals as included in the analysis above. CACM and Harvard Business Review (HBR) are worthy of particular note when looking for flow-on effects. Both reach many readers, as evidenced by their high circulation counts. CACM is the practitioner outlet which cites the largest number of papers in our basket of IS research journals. Polites and Watson (2008) showed in a social network analysis of CACM citations using 2003 data that CACM was the top-ranked journal in a network of 120 IS-related journals in terms of both prominence and centrality. CACM was shown to occupy a central bridging position in the network between the more technical computer science journals and the more business-oriented IS journals. HBR has a very large practitioner base and the potential to reach a significant body of managers whom we wish to influence.

Table 4 shows the result of a similar analysis performed with the practitioner journals to what was performed with the selected IS journals. It indicates that there are indeed such flow-on effects. The practitioner journals are cited comparatively more often in other practitioner journals than are the IS research journals.

CACM is noteworthy. In 2006, CACM articles were cited 130 times, mostly in CACM itself, but there were another 33 citations in other practitioner journals. Further, CACM by itself is the practitioner journal most likely to cite IS research articles (almost half of the citations to them in 2006). As suggested earlier, CACM does appear important to the practitioner influence we wield.
Table 4: Flow-on effects from practitioner journal (all years) to other practitioner journals in 2006

The effect with *HBR* is even more marked. It had a total of 4,913 citations in 2006, with 138 in practitioner journals. However, it does not itself cite other sources in its articles and we cannot, therefore, claim any obvious additional flow-on effects though there may be non-obvious connections. This might require an analysis of the content of *HBR* articles. We cannot assume, however, that any influence to practitioners is benign to IS, as the article by one of its editors (Carr, 2003) evidences. It is beyond the scope of our paper to take our analysis very much further. However, this exploration excited us to see whether there is influential IS research underlying publications in large circulation practitioner journals, which may not have not been identified in the analysis above. To do this analysis, we used Google Scholar, which has more extensive references, and more relaxed criteria for including reference material. For example, citations need not be to an *ISI* listed journal and therefore potentially include papers in the *International Conference of Information Systems* as well as *JAIS*.

Doing searches with Google Scholar and a combination of search terms including “Communications of the ACM”, “Harvard Business Review”, “information systems” and “information technology” yielded some interesting results. A number of the searches converged on the same few papers in *CACM*, *HBR* and *SMR*, all of which have very large citation counts. Table 5 shows these papers and their Google Scholar and ISI citation counts. In further research the efficacy of using Google Scholar to develop this work could be investigated. It is very likely to be most popular in the future among IS academics, if perhaps not librarians.

Table 5 shows four papers that have particularly high citation counts. Although appearing in practitioner journals, they score highly even compared with the most highly-cited IS research articles. Lowry et al. (2007) identify the Delone and McLean (1992) article on the “quest for the dependent variable” as the most highly cited article from leading IS journals, with 481 citations. Only this paper and that of Moore and Benbasat (1991), with 312 citations, are shown as having higher citation counts than the Davenport and Short (1990) article in the *Sloan Management Review*. This observation indicates that Davenport and Short not only achieved high research impact in our normal use of the term, but they are also likely to have had high effective relevance because of their reaching out to a large practitioner audience, both directly through *Sloan Management Review*’s readership, but also through flow-on effects to other practitioner outlets.

Interestingly, Davenport and Short (1990) did not cite any leading IS journals as a source of ideas. Rather they appropriated ideas from industrial engineering (citing *Industrial Engineering*) and reported directly on what had been observed in case studies in 19 companies. Earl (1994, p. 6) refers to the concept of business process redesign as belonging to the “managerial journalism domain” with Davenport and Short’s seminal article in *Sloan Management Review* as being in a journal that provides “an interface between business schools and practitioners”. In view of the impact on subsequent
streams of research on business process redesign, business process modeling and the like, and the applicability of these ideas in practice, this “managerial journalism” is perhaps not such a bad thing.

<table>
<thead>
<tr>
<th>Article</th>
<th>Citations – Google Scholar</th>
<th>Citations - ISI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information technology and business process redesign, Sloan Management</td>
<td></td>
<td></td>
</tr>
<tr>
<td>strategic management system, Harvard Business Review, 74, 1, 75-85.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Davenport, T. (1998). Putting the enterprise into the enterprise system,</td>
<td>1275</td>
<td>160</td>
</tr>
<tr>
<td>technology, Communications of the ACM, 36, 12, 66-77.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Influential IS ideas in practitioner journals

5 FURTHER REFLECTIONS

Our analysis, emphasizing the impact of IS research in terms of references to IS output reflected in our major research journals, has not shown clear links between IS research and practice, though some tentative strands have been suggested. We discuss the results in terms of our research questions, taking the second question first. We also provide suggestions for future work and action.

5.1 Can we demonstrate knowledge transfer between IS academic research and practice?

The approach in our study has not proved fruitful in showing knowledge transfer between IS academic research and practice. That is, analysis of citations of articles in leading IS journals in a sample of practitioner journals showed that a small number of IS articles were referred to the practitioner journals – less than 1% of total citations.

It is acknowledged that there are weaknesses in the approach we have used. Only a sample of practitioner journals was used and a number of the outlets we studied restrict the number of citations that authors can make. Further, by relying on citations alone we have not investigated the depth or extent to which the cited paper influences the ideas in the citing outlet. This latter weakness is one that is shared by other citation count methods that are used to assess impact, although our discipline frequently employs them as a base for assessing “research impact” and journal quality.

We suggest, however, that our study has considerable merit in that it provides a foundational step for an unexplored problem. When we began our study we could find no systematic studies of how knowledge transfer in IS could be assessed. Exploring the use of a citation count method is an obvious place to start, as this method can be applied consistently, the data is readily available and it is an approach we are familiar with for the assessment of research impact. Even the null result in our study is valuable as it means that future studies can build on it in unexplored directions.

We have made a tentative start on suggesting other avenues to explore in showing that there can be spreading effects from a publication in a practitioner journal to other practitioner journals. Further avenues to explore include more in-depth analysis of practitioner material to identify research work that has been used as a source of ideas. Further work could also widen the scope of the enquiry from practitioner publications to other transfer mechanisms such as textbooks and research monographs. Enid Mumford, for example, claimed that her work in companies could best be described in books and monographs rather than journal papers (see ISJ, 2006 for a full discussion of her work and Mumford and MacDonald, 1989, as an example). Further, articles on websites have potentially a high impact.
CIO Magazine (URL www.cio.com) is read by practitioners whom we wish to influence and so mention of our work in such sites does suggest impact. Practitioner-oriented IS publications are also crucial, such as MIS Quarterly Executive.

Other transfer mechanisms to be explored include student learning and development. Clearly IS academics influence practice through the employment of its graduates in relevant practice. However this is not evidence of our academic research making impact. Nevertheless many successful PhD students decide to continue their careers in practice rather than academia. This needs to be recognized as a positive outcome. Bjorn-Andersen (2008) argues that more than 50% of students at Copenhagen Business School not only take this route but this was their intention originally and these include some of the most talented researchers.

We now turn to our first research question.

5.2 Does IS academic research have consequences for practice?

We can give only tentative responses to this question because of our findings for the first question. We were unable to demonstrate, using the approach we adopted, that there were significant knowledge flows from IS academic research to practice. Because this result may be due to the limitations in our approach we do not want to speculate too far on the answers to the first question.

Our study did suggest, however, that ideas from IS may not be as readily passing on to a managerial audience to the same degree as in some other disciplines. A very preliminary analysis of SMR publications showed that academic works from Management, Marketing and even Psychology and Sociology were being cited to a greater degree than work from IS. This phenomena needs further study. Possible reasons include editorial policy or a lack of suitable contributions from IS researchers. A more worrying explanation would be that IS academic research does not have relevance to the SMR audience. We leave investigation of this phenomenon to further work. Concern over this issue continues in other fields and movements such as that towards “engaged scholarship” (Van de Ven 2007) and design science (Hevner et al., 2004) may go some way to addressing these concerns.

5.3 Suggested action

The finding that is clear from our study is that there is relatively low proportion of academic IS work cited in the practitioner outlets we studied, although some work appearing in practitioner outlets has itself been very well cited in absolute terms. We therefore suggest:

- **Targeting the indicators:** Authors of papers in research journals might write more articles for trade magazines and similar and communicate the research in a more easily digested form whilst referencing the original research articles. Such trade magazine output should be recognized as worthwhile in the IS research community.

- **New Magazine:** CACM has signaled a change in focus (ISWorld listserv, 2 April 2008) that may well reduce IS publications and flow-on to practice. We therefore suggest a new magazine, similar in format to the present CACM, but solely IS and perhaps produced by the Association for Information Systems (AIS). Such an outlet could, for example, include interviews with practitioners illustrating the usefulness of our research as well as evidence of it in its articles.

In conclusion, it is somewhat surprising that prior work investigating whether the impact of IS research on practice is in fact real, rather than imaginary, appears to be non-existent. Our paper makes a contribution in again calling attention to this important topic, and indicating some means by which further investigation can proceed. Our initial study is limited to bibliometric analysis but provides a base for further research.

We need to show that our research has high impact on practice, and therefore we need to work on this problem as a community to affect those indicators that show academia and the outside world that the
IS discipline has made and continues to make a positive and major impact on organizations, people and society.

References


European Union (2008). URL:


Appendix A: MISQ articles referenced in CACM articles in 2006

(Articles were found using Scopus, which yielded 18 articles, compared with the 19 articles found by Web of Science for the same source and same year, see Table 3.)


A WIKI-BASED APPROACH TO ENTERPRISE ARCHITECTURE DOCUMENTATION AND ANALYSIS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0392.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Architecture management, Collaboration, Conceptual modelling, knowledge workers</td>
</tr>
</tbody>
</table>
A WIKI-BASED APPROACH TO ENTERPRISE ARCHITECTURE DOCUMENTATION AND ANALYSIS

Sabine Buckl, Florian Matthes, Christian Neubert, Christian M. Schweda
Technische Universität München, Chair for Informatics 19 (sebis) Boltzmannstr. 3, 85748 Garching, Germany, {buckl,matthes,neubert,schweda}@in.tum.de

Abstract

Enterprise architecture (EA) management is a challenging task, modern enterprises have to face. This task is often addressed via organization-specific methodologies, which are implemented or derived from a respective EA management tool, or are at least partially aligned and supported by such tools. Nevertheless, especially when starting an EA management endeavor, the documentation of the EA is often not likely to satisfy the level of formalization, which is needed to employ an EA management tool. This paper address the issue of starting EA management, more precise EA documentation and analysis, by utilizing a wiki-based approach. From there, we discuss which functions commonly implemented in wiki-systems could be used in this context, which augmentations and extensions would be needed, and which potential impediments exist.

Keywords: Enterprise Architecture, Wikis, Information Modeling.
1 MOTIVATION

In recent years, companies had to face various challenging environmental influences, amongst others forcing them to ensure and further develop the alignment of their business and the corresponding IT support (see e.g. Laudon et al. 2006). Therefore, the management of the enterprise architecture (EA), which is commonly regarded as means to support this task, currently gains increasing attention in many medium-sized to large enterprises, reflected by a rising awareness for this field in academia (Langenberg and Wegmann 2004). Although a multitude of approaches to EA management with different backgrounds exists, ranging from practical origin (cf. Dern 2006, Niemann 2008, and Keller 2007) via standardization bodies, as e.g. the Open Group (TOGAF, cf. Open Group 2009) or the Object Management Group (CWM, cf. OMG 2001) and tool vendors (cf. Matthes et al. 2008), to academia (cf. Fischer et al. 2007, Frank 2002, Lankhorst 2005, or Ross 2006) no commonly accepted best-practice yet exists.

In spite of the number of different origins, all approaches center around the focal term architecture. According to ANSI/IEEE Std 1471-2000 architecture is defined as ‘the fundamental organization of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution.’ (IEEE 2000, S.3) Applying the aforementioned definition to the context of enterprises, the EA refers to the fundamental organization of an enterprise, embodied in its components (e.g. organizational units, stakeholders, locations, business processes), their relationships to each other (e.g. supports, hosts, is responsible for), and the principles (e.g. profit, continuity, innovation) governing its design and evolution. Additionally, the term management according to Mary Parker Follet refers to ‘the art of getting things done through people’ (van Aken 2005). Thereby, a typical management process is organized as a cycle containing the following phases: Plan –Do – Check – Act (Deming 1986, Sheward 1986). Put in the context of EA management, the Plan phase includes the development of future planning scenarios of the EA as well as the decision-making. The selected scenario is realized in the Do phase, followed by the analysis and evaluation of the implemented EA during the Check phase. The following iteration cycle is prepared in the Act phase via the identification of potential process improvements.

The goals of the ANSI/IEEE Std 1471-2000 are inter alia the documentation, communication, and analysis of architectures. Based on the definitions given above, the main tasks of EA management are:

- **Documentation:** In order to support the plan phase of the management cycle, the current (as-is) situation of the EA has to be documented. Thereby, elements on different layers ranging from business & organizational to infrastructure aspects (see Figure 1) have to be considered to provide a holistic view on the enterprise. Prior to gathering the data about the current situation, an information model\(^1\) of the architecture has to be developed, which defines the elements and relationships in-between constituting the EA. Besides the current situation, information about future states according to the plans has to be documented. Complementing the current and planned states of the EA an ideal target (to-be) state should be envisioned, which can be derived from the long-term vision of the enterprise.

- **Communication:** Gathering information from the different layers in the plan, do, and check phase, as introduced above, requires the involvement of a multitude of stakeholders, e.g. business process owners, project managers, business architects, etc. Although working for the same company, the terminologies used by these stakeholders differ widely. This communication issue is often referred to as the communication gap between business and IT (Lankes 2008, Schekkerman 2004). This gap is likely to hamper effective communication and

---

\(^1\) The information model is sometimes also referred to as meta model. We prefer the term information model, which we regard to be more unambiguously (cf. Buckl et al. 2007b).
collaboration in the EA management process. Visualizations are a commonly accepted means to bridge this gap. Another challenge is connected to the aspect of historization and traceability of management decision (Buckl et al. 2008).

- **Analysis:** Concluding the management cycle, the current, planned, and target architecture of the enterprise have to be analyzed and evaluated in the check phase to support decision making and identify potential improvements (Johnson et al. 2007). The analysis results are finally used to improve the procedure of EA management itself in the act phase.

![Figure 1: Layers and cross-functions, which form a holistic view on the enterprise](image)

The *traditional* way to address the challenges of documentation, communication, and analysis in the context of EA management, as described above, is to select an EA management tool available (cf. Matthes 2008), and to use the inbuilt information model and methodology. Due to the high investments of such an approach, we propose an alternative lightweight wiki-based approach to EA management in this article, which we deem to be suitable especially for emerging EA management endeavors. Starting with a discussion about the application of wikis in related contexts, as e.g. software engineering, in Section 2, Section 3 explicates on the state-of-the-art in starting EA management endeavors. Thereby, we emphasize on the main tasks of EA management as introduced above. Following Section 4 introduces the wiki-based approach and discusses how it can be used to facilitate the challenges arising in the context of documentation, communication, and analysis of EA related aspects. The proposed approach is thereby complemented by methods for extracting structured information from wikis, by which we hope to promote a smooth transition from beginning small-scale EA management initiatives to mature enterprise-wide endeavors through leveraging information reuse. Finally, Section 5 concludes with a short outlook on upcoming or interesting research concerning the topic and indicates points of interest, which could be subject to further research initiatives.

## 2 SELECTED WIKI USAGES AND EXTENSIONS IN LITERATURE

A number of attempts have been undertaken to use wikis in enterprises as a tool for documentation and communication. Quite a lot of these attempts have been reported in papers on knowledge management (KM). Although EA management, as introduced above, can be regarded as an issue of knowledge management, we do not detail considerations on wikis from a pure KM perspective as found in literature. Instead, we discuss usage approaches and tool extensions directly focused on the three main tasks according to the IEEE, namely documentation, communication, and analysis.

The first approach is taken from the field of software engineering, which is strongly concerned with architectural topics and focuses on the tasks of documentation and communication. In a research project executed at Carnegie Mellon University a wiki-system was used to collaboratively create and
maintain the documentation of the architecture of a single software system under development (Bachmann and Merson 2005). A key finding of research was that the structure of the documentation emerging in a wiki was different from the structure typically employed in a document-centric setting. The documentation in the wiki ended up at a different level of abstraction concerning architecture elements, i.e. complex elements were split up into single and separately maintained pages of which each documented a small portion of the element. This effect, which normally does not occur in document-centric settings, was considered helpful to reduce contention between the parties employed in documenting. This tendency of the information contributors to change the level of abstraction, if they regard an entity to be too complex to describe it in a single text, further promotes documentation maintenance. This leads to a more current documentation, although issues of assigning responsibilities for keeping pages up-to-date were experienced. The wiki-system was also found useful in promoting reuse of documentation fragments, e.g. by linking related pages or embedding visualizations such like class diagrams (Bachmann and Merson 2005). Nevertheless, the latter aspect was not free of problems, as the wiki-system did not provide means for collaboratively editing visualizations. Thus, they were created offline and embedded as images – a fact greatly hampering their maintenance.

An application showing, how wiki entries created and maintained by an appropriately large number of users, can be used as a basis for collaboratively and iteratively building a documentation schema for a field of interest is given in Hepp et al. (2007). Thereby, the authors show how basic wiki concepts as the URIs, referring to described entities, the links between the entities, and the versioning history can be used to derive a basic and simple schema of the documented concepts. While Hepp et al. outlines, that the approach taken might be greatly augmented e.g. using semantic wiki technologies, one of the core findings also holds for simple wiki-systems: increasing stability of entity descriptions indicates the evolution of, what the authors call, wiki consensus on this entity’s meaning.

In Auer and Lehmann (2007) an approach to lift hidden semantic treasures out of simple wiki-systems is presented. In the approach especially wiki templates, as commonly used throughout the wikipedia, play an important role. These templates can be seen very similar to predefined forms, which were initially intended to provide consistently formatted graphic elements for certain content in specific articles. The articles using the same template can mostly be regarded to describe entities of the same type, e.g. all articles, which use the template Settlement, describe settlements. As a consequence, Auer and Lehmann the authors of the article developed an algorithm capable of searching a wiki (in this special case the wikipedia) for the occurrences of specific templates, which were then subsequently parsed and translated to a more formal representation. Having had on the one hand astonishing success with the extraction, some caveats with the approach could be identified, especially concerning the design of the used templates. Some templates are made up in a way, that (semantic) information and presentation information are not strictly separated. Additionally, for the same concept several templates might exist. As a consequence Auer and Lehmann (2007) provide guidelines for designing semantically rich templates, which we will refer to in the presentation of our approach.

3 STARTING AN EA MANAGEMENT INITIATIVE – STATE OF THE ART

In our research project Software Cartography, we gained insights regarding the first steps taken to establish an EA management endeavor in different global acting companies, e.g. Allianz, BMW, MunichRe, Siemens, T-Com. Accordingly, a typical emerging EA management initiative typically starts with the documentation of the as-is situation of the EA. Thereby, seldom a strict and predefined information model is utilized at the beginning of an EA management initiative. Thus, the documentation of the EA is often left to text or spreadsheet documents without a formal structure. This method is often useful to gain first insights on the EA and, more important, to satisfy information needs without having to first explicate and formally describe them. Additionally, the method is commonly regarded as more appropriate for starting with EA management, as the creation or selection
of an EA information model can be considered a labor intensive task, which often employs various review-cycles.

Nevertheless, when the EA management endeavors become more mature and multi-partied, commonly an object-oriented information model for the EA concepts is used. The same is true with the rising demand for aggregated information representations such as e.g. computed metric values (see Lankes 2008), which are mostly grounded on object-oriented information models. Numerous of these models currently exist, heavily researched in the academic community (see e.g. Braun and Winter 2005 or Lankhorst 2005), by tool vendors (cf. Matthes et al. 2008) as well as standardization bodies (e.g. the Object Management Group or The Open Group), and developed by the companies practicing EA management (see e.g. Halbhuber 2004, Lauschke 2005, Brendebach 2005). This fact is further promoted, as no common and fully accepted standard information model has yet been published – with some researchers, tool vendors, and EA management practitioners claiming, that such a model suitable for any company and any way of performing EA management does not exist (Buckl et al. 2007a or Kurpjuweit and Winter 2007). Whereas Kurpjuweit and Winter propose an stakeholder-centered approach to information model creation, Buckl et al 2007a and Buckl et al. 2008 presents a concern-centered approach.

With the growing complexity of the body of management and the rising number of people to be involved as information suppliers, spreadsheets and text documents do not scale appropriately, negatively influencing the overall acceptance of the initiative, e.g. due to arising problems of maintaining the consistency (Fischer et al. 2007). A way to bring EA documentation and communication efforts to the next level is the introduction of an EA management tool, into which the previously gathered data shall be imported. This step can be considered as a turning point in two ways: Firstly, up to that point EA management and especially the documentation efforts do not comprise a major investment and are thus likely to emerge as bottom-up initiatives. In contrast switching to a tool not already present in the enterprise commonly cannot be done without having management support as this poses a great investment for the company. Secondly, the tabular and textual documentations rarely adhere to an explicaded and formal structure. Bringing this information into an EA management tool can be regarded a quite complicated task, as the majority of tools uses an object-oriented information model for the governed information (cf. Matthes 2008 et al.). Therefore, the structure behind the information, previously gathered in spreadsheets and text documents, has to be explicaded and mapped to the structure provided by the future EA management tool.

The process of mapping a grown structure, which evolved in a process potentially without centralized control, to the structure as incorporated in the tool, can be regarded to be challenging (see Matthes et al. 2008). This is on the one hand true, if the tool comes shipped with a predefined and fixed model, to which all the concepts have to be mapped. More often than not the predefined information model demands some concepts to be maintained, about which up to that point no information was gathered. On the other hand, if the tool provides adaptability of the underlying information model, the issue to tackle does not become smaller. A variety of different stakeholders throughout the enterprise wants or needs to be involved in creating and developing a company specific information model, which supports the analyses needed to perform enterprise-specific EA management tasks. Beside the different concerns raised by the different stakeholder, also definitional issue are likely to arise, because terms as application are widely used and are thus strongly overloaded concerning their meaning.

4 A WIKI-BASED APPROACH …

Based on the findings of the work of Bachmann and Merson 2005, we discuss how a wiki-based approach facilitates EA management endeavors. Thereby, we structure our discussion along the three main tasks – documentation, communication, and analysis – and show possible impediments for effectively using wikis in these areas.
4.1 Applying tags to objects is a prominent method for categorizing concepts especially in the Web 2.0 environment (cf. http://del.icio.us). Also many wiki-systems utilize this mechanism, to allow their users to contribute classification information, which constitute bottom-up taxonomies, so called folksonomies. Further, widely used wiki-systems (see e.g. http://www.wikipedia.org) employ the mechanism of templating to provide pre-formatted and structured building blocks for wiki-pages. We subsequently discuss how these two techniques can be used to facilitate the creation of an EA information model consisting of classes and properties.

4.1.1 Collaborative tagging

Golder and Huberman (2005) introduce a classification for different kinds of tags, which are used throughout a collaborative tagging environment. Based on this classification, we discuss which types of tags can be used to facilitate the collaborative creation of an EA information model. These are on the one hand the what-it-is-about tags and what-it-is, which can be regarded to be quite identical in the context of EA documentation. Such tags commonly form the overwhelming majority of tags used in collaborative tagging environments. On the other hand, the who-owns-it tags, which are used especially for tagging blogs, can also be used in EA documentation, although ownership has a different notion there. Owning a concept means to be responsible for this specific concept. Such information should better not be expressed via tags but using properties as described in the next subsection. Therefore, we do not expect ownership tags to be very frequent in EA documentation wikis. A last type of tags, the refining-categories tags are not used on their own, but together with tags of other types for refining their semantics and creating subcategories. These tags can provide information valuable to EA information modelling, although we do not expect them to be more frequent as in general tagging environments.

The tag types discussed above can be considered very useful for finding and refining classes in a similar way as e.g. in object-oriented design. There, the creation of classes is simply spoken related to grouping objects, which share common properties, and finding an abstract label for describing the group. Nevertheless, especially in complex and evolving environments, as e.g. an enterprise such a classification might not easily be found. Therefore, collaborative tagging can provide a useful contribution, when EA concepts, documented in the wiki, are applied tags, e.g. "business application" or "server". Other tags as “important”, which are widely used in tagging environments, are nevertheless not directly related to classes in the information model, but merely have a user-related semantics. They can notwithstanding be used for our approach, cf. Section 4.2.1.

4.1.2 Open templating mechanism for property modelling

Wiki templates, as discussed above, provide a mechanism for defining values for a set of fixed properties of a certain type of object, e.g. a settlement. Nevertheless, as neither a set of fixed types, which could be used for defining templates, nor a fixed set of properties exist, the templating mechanism cannot directly be applied here. Therefore, we propose an adapted form of the mechanism, which we call open templating mechanism. In the context of the mechanism the template itself is reduced to a simple table, in which the user can add rows and thereby properties as well as corresponding values. As a consequence, the user can freely introduce properties and apply values to them. Because of the nature of wiki-systems these values are only textual, i.e. no data types are

---

2 We use the terminology of the Essential Metaobject facility (EMOF), see OMG (2006). Therein, no distinct concept for modeling relationships exists; instead they are represented as connected properties.
assigned to the properties. This might cause several problems discussed below. Nevertheless, semi-complex data structures (e.g. arrays, sets) can be utilized to bundle composite values.

In order to prevent anarchic evolution of templates, a recommender mechanism is added, providing properties, which are potentially relevant to the object currently described. The relevance therein is not computed from the overall frequency of that property but from a correlation analysis between the property and the tags, which are applied to the described object. Thereby, it is determined how frequent given property and tag are used in conjunction compared to the number of usages in general. We regard this method to be appropriate for quickly and easily finding the important properties. In order to display the recommendation of the system, again a visualization resembling a tag-cloud could be utilized, in which the properties most correlated are displayed as more relevant.

4.1.3 Potential Impediments

Collaborative tagging and open templating might be well-suited for the above context; nevertheless some impediments might hamper the development of a semi-structured EA documentation. Subsequently, we outline two of them, which can negatively influence user acceptance. This acceptance is generally considered a critical success factor for tool-supported, bottom-up knowledge management endeavours.

A phenomenon repeatedly reported in the context of classification attempts, is what was called the base level in Tanaka and Taylor (1991). When asked to classify an object, people tend to use terms, which are neither strongly generalized nor specific according to their distinct knowledge backgrounds. The chosen intermediate level between generalization and specificity is called the base level of the particular person. As a consequence of the base level being specific for every person, the expertise in the domain, from which objects have to be classified, greatly influences the specificity or generality of the used classification terms. This effect might especially apply in the area of EA management, where stakeholders from many different areas are contributing to a holistic documentation of the EA. Therein, they are likely to classify according to their domain expertise, leading to issues, which potentially impede the development of a common and consistent information model.

Wiki-entries and therefore also the values contained in templates are mainly textual. While this can be seen as an important success factor for wiki-systems in general, the fact has some negative influence in information modelling. The properties' values in the templates are pure text values, even if they represent information about date or numbers. As a consequence, no type and consistency checking can be performed – endangering the overall quality of the data collected. In addition, when a full migration of the semi-structured wiki-based EA documentation to a fully object-oriented model should be performed, the data types of the properties have to be determined. While possibly some of this determination may be performed automatically, no one expects that to be possible in all cases. This is especially true, if properties representing relationships between objects are considered.

4.2 ... to EA information communication

The different stakeholders of EA management endeavors have distinct information demands and perform specific activities during the management processes. According to the information demands, these stakeholders can be assigned to groups – a functionality, which many wiki-systems provide out-of-the-box. The specific activities can be connected to roles in the wiki, such that access rights can also be granted based on the role-description taken by a person. The formation of virtual communities of practice of different stakeholders as wiki-groups can be leveraged to facilitate communication among these stakeholders by means of collaborative bookmarking or watch lists.
4.2.1 Collaborative bookmarking

In contrast to collaborative tagging, the collaborative bookmarking does not assign classifying tags to the objects, but assigns contextualizing tags, which have no general meaning, but provide sense for the user (or user group), from which the tag originates. Golder and Huberman (2005) provide a classification for these types of tags into self-referential and task-organization tags. Self reference tags are used to express a relationship of the tagged object to the user. These relationships do mostly not refer to formal and organizationally established responsibilities, but refer to intrinsic areas of interest of the specific user. Task organization tags, as e.g. "TOBEREAD" or "TODO", are commonly used in tagging environments to keep track of open activities or inform fellow users of open tasks. Both types of contextual tags do not contribute to the creation of an EA information model, but are important means for information distribution among a certain group, e.g. used as hints during the collaborative creation of an EA documentation artefact.

In order to clearly separate the different types of tags – the classifying ones needed for EA information modelling, and the contextualizing ones needed for EA information communication, we propose an additional concept, which might be helpful in this context. The user should be given the possibility to choose, whether a given tag is visible in public, to a selected user group only, or private, according the aforementioned virtual communities of practice, which the user participates in. This also gives rise to the idea to extend the tag recommender system, such that a user is recommended tags in respect to the groups, he belongs to. Such a recommender system can be helpful to prevent the evolution of numerous synonymic tags in the wiki. The tags recommended are therein sorted by relevance and displayed in an appropriate way, e.g. in a tag-cloud.

4.2.2 Watch lists and change feeds

As the documentation of the EA can be regarded an effort of strongly collaborative character, group related notification in case of specific changes and communication of ongoing documentation initiatives is often considered to be important by practitioners for a number of reasons:

- The person responsible for certain parts of the EA can get a notification, if the documentation of one of these elements is adapted, e.g. to perform quality assurance.
- The head of a group of people employed with documenting parts of the EA can keep track on the progress made in the documentation endeavour.

Watch lists and change feeds, as implemented in many different wiki- systems, may be appropriate to address the notification demands as sketched above. Nevertheless, further demands for notification may exist, e.g. a user might want to get a message on pages, which have not been changed for a certain period. These messages might give indications on parts of the EA documentation becoming outdated or parts being not well established. While such a feature is not commonly supported in simple wiki-systems, the information on the last update of a page is stored and could be easily used to satisfy the notification demand as outlined before.

4.2.3 Potential impediments

Each of the communication facilities outlined above has some subtle complexities assigned, which may lead to sub-optimal usage in an EA communication environment. The collaborative bookmarking might fall for his apparent similarity to the collaborative tagging facility. While the latter is intended to evolve a bottom-up classification schema for concepts, the former can be used as mnemonic for a user or to guide the attention of a user group. Nevertheless, one cannot expect the users of the wiki-system to keep this strict separation alive. Hence, the classification schema might get “contaminated” with contextualizing tags, which have no meaning to the public. The opposite effect might also occur – users might classify objects in the wiki private, such that the information model cannot benefit from.
The watch lists might fall for legal problems, as they could be misused as means to observe and track employee behaviour in the wiki-system. Beside this problem, the watch lists facilitate a one-way usage paradigm, where users only aggregate and consume information, but do not directly contribute to the information and documentation base. Finally, watch lists might, similar to task organization tags, be helpful for directing user attention. Nevertheless, an excessive supply of such list is most likely to distract the attention of the users.

4.3 ... to EA analysis

The documentation of the EA as incorporated in the wiki-system has to be analyzed in various ways to support the EA management process. Different wiki-systems provide functionalities, which are useful for performing EA analyses. These functionalities range from versioning, which is a common feature for most wiki-systems, to textual querying, which is supported by some semantic wiki-systems (e.g. the semantic media wiki – see http://www.semantic-mediawiki.org). By textual querying, semi-structured information from wiki-pages is queried against, and the results of the query are presented in a tabular report.

4.3.1 Versioning

Most wiki-systems provide versioning of pages, i.e. the history of past versions of the page is stored in the underlying database of the wiki. Thereby, two important aspects of EA analysis are facilitated. At first, the description of the EA or parts thereof is historized – a feature, which is often requested by EA practitioners to ensure traceability of management decisions and project changes. In contrast, many of the current tools for EA management (see Matthes et al. 2008) do only rudimentarily support versioning. Secondly, if mechanisms as collaborative tagging and open templating are used to develop and evolve an information model for EA management, versioning of wiki-pages also contributes to a historization of information models.

The possibility to distinguish between minor and normal changes of wiki-pages additionally allows to clearly separate maintenance activities in the documentation from situation, where substantial change in the underlying architecture is documented in the wiki. If the wiki-system further supports to give comments on the changes, especially for the normal ones, this functionality can be used to relate a documented architectural change to the causing EA project. This traceability of the projects, which drive the managed evolution of the EA, is often regarded to be important for EA planning and analysis (Buckl et al. 2009). Finally, the option to mark a change as minor allows reducing the amount of information, which shows up in watch lists, such that the user’s attention is not distracted by these maintenance activities, but is guided towards real architectural changes.

4.3.2 Textual querying

Textual querying allows accessing both the textual and the semi-structured parts of wiki-pages in a more convenient way. In particular, tabular views on the structured information can be created via queries, which can be used to get an overview on specific concepts, e.g. all objects tagged with “business application”. In this overview, also different kinds of sorting can be applied and comparisons of objects of the same type are performed easily.

A tabular listing of “business application” concepts might additionally be helpful in guiding the evolution of the information model, as maintained in the wiki-system. While at the beginning of a collaborative documentation endeavour, frequency counts and correlation analyses might be appropriate methods for determining the importance of a tag and the corresponding properties used in the open templates, the evolving information model may at some point become overloaded and partially anarchic. By comparing the property supply for a certain type of object, a domain expert for EA modelling can decide to make some of the properties mandatory, leading to increased consistency.
of the information model. Having taken such a decision, the tag-cloud recommender functionality as introduced above, could be adapted to incorporate the effects of the decision, e.g. by colour coding certain elements or enlarging their font-size. This could be used to leverage the documentation of specific concepts, which are especially necessary for a distinct EA management endeavour.

4.3.3 Potential impediments

The basic facilities for comparing selected versions of a wiki-page are often somewhat limited in respect to the semi-structured information supplied by open-templating mechanisms. Therefore, especially functions for executing comparisons within the templates can be considered beneficial, so that the user can see a tabular report on the values supplied for properties in the selected versions. For such a comparison, it would also make sense to compare more than two versions of the object – a feature also not commonly supported in wiki-systems.

The synopsis of different objects of the same type or versions of the same object over time can, as alluded to above, be used to facilitate decisions on the EA information model. Nevertheless, if a decision on certain properties is taken, e.g. making them mandatory, the bottom up development of the EA information model is disturbed, which might cause a decrease of user acceptance. Hence, any top-down adaptation to the community developed model should be carefully assessed to prevent a downturn in user motivation.

Textual queries are only the first step on the way to conveniently present semi-structured information to a broader user community. Going beyond simple tabular reports, graphical visualizations of the EA or parts thereof are regarded as an important instrument (Lankes et al. 2005) to both support and promote the EA management effort as well as to increase the awareness for that topic among the company's employees. Such visualizations, called Software Maps in the context of documentation of the EA, can according to the approach presented in Buckl et al. 2007b be created from information about the EA, if all concepts displayed are sufficiently documented. This should be the case, if open templating mechanisms are employed in the wiki to support semi-structured information provision.

Here, we see a potential synergy effect, if the wiki-system can provide access to automatically generated visualizations displaying concepts within the EA, which have been described sufficiently detailed. Such visualizations, if comprising which information has to be maintained for a specific concept in order to be displayed therein, could then leverage the maintenance of that data. Thereby, users of the documentation wiki are encouraged to supply additional information about the EA constituents, which they are responsible for, to make sure, that these objects are displayed in the visualizations of the EA.

5 OUTLOOK

In this article, we presented an approach to use wiki-systems for documenting, communicating, and analyzing the EA. Furthermore, we elaborated on different aspects of wiki-systems, which might be helpful in this context. Nevertheless, the approach presented has yet not been validated in practice. We assume, that especially medium-sized companies may benefit from the presented techniques for starting EA management, while larger enterprises are likely to stay to the more formal and structured tools.

Further aspects of interest regarding the approach are e.g. scenarios to use rating functions on wiki-pages to allow users to provide feedback in a structured and uniform way. Concerning these functions a multitude of implications, e.g. on user motivation would have to be considered. Additionally, these ratings could provide valuable input to the analysis of documentation templates, although a clear correlation between a low rating and the respective template used is not likely to be easily determinable.
REFERENCES


SOCIAL CAPITAL IN DISTRIBUTED SYSTEM DEVELOPMENT: A CASE OF GRID DEVELOPMENT IN PARTICLE PHYSICS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0729.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Grid Computing, Distributed IS project, Information Systems Development (ISD), Social capital theory</td>
</tr>
</tbody>
</table>
SOCIAL CAPITAL IN DISTRIBUTED SYSTEM DEVELOPMENT:
A CASE OF GRID DEVELOPMENT IN PARTICLE PHYSICS

Yingqin Zheng, Information Systems and Innovation Group, Department of Management,
London School of Economics, Houghton Street, London WC2A 2AE, UK,
y.zheng2@lse.ac.uk

Will Venters, Information Systems and Innovation Group, Department of Management,
London School of Economics, Houghton Street, London WC2A 2AE, UK,
w.venters@lse.ac.uk

Tony Cornford, Information Systems and Innovation Group, Department of Management,
London School of Economics, Houghton Street, London WC2A 2AE, UK,
t.cornford@lse.ac.uk

Abstract

This paper examines dimensions of social capital in the distributed collaborative development of the
UK particle physics Grid. It is shown that the GridPP project effectively draws upon social capital
rooted in the tradition and culture of particle physics experiments, characterized with trust, equality,
shared vision, collaboration, and pragmatism. These factors contribute to overcoming the challenges
in the creation and sharing of knowledge in the development of the Grid, a cutting-edge technology
that has to be delivered as a working system with limited time and resources. This case sheds lights
on, and provides a good example of, the importance of social capital in distributed systems
development.

Key words: Distributed System Development, Social Capital, Knowledge Processes, Grid
1 INTRODUCTION

As systems development activity becomes increasingly globalised and distributed, projects face great challenges in developing and maintaining knowledge and expertise. The need to mobilize knowledge capabilities is particularly acute when the task itself is highly complex, uncertain, and involve innovative technologies. In this paper we use the concept of social capital to examine the knowledge processes in the development of the UK particle physics Grid. Perceiving the construction of the Grid as a grand systems development challenge in technical, organizational, political and human terms, we would like to go beyond the usual software-centric view of system development, and focus on how a collaborative project, GridPP, dynamically mobilizes social capital to sustain the distributed system development process.

The UK particle physics Grid is part of an initiative which aims to produce not just a working system but a new generation of computing technology that will potentially have significant impact on scientific research, and may foreshadow the “next generation Internet” (Abbas 2004; Carr 2005). It is being developed as a large scale distributed collaboration, by particle physicists who come from a community with very distinctive work practices and culture (Knorr-Cetina 1999; Traweek 1988). They have a record of success in developing innovative computing solutions of which the World Wide Web is the most notable example (Berners-Lee 1989).

The rest of the paper starts with a literature review which connects distributed system development and social capital. The case is then introduced, followed by a brief outline of the methodology. Section five provides an analysis of the case using the three dimensions of social capital (Nahapiet and Ghoshal 1998). Section six further discusses the implications of mobilizing social capital for knowledge processes in distributed systems development, and concludes the paper.

2 DISTRIBUTED SYSTEM DEVELOPMENT AND SOCIAL CAPITAL

With the current trend of globalization where the IT industry is becoming more and more globally interconnected (Herbsleb & Moitra 2001), information systems development has increasingly become a multi-site, multi-cultural, globally distributed undertaking (Herbsleb, Paulish and Bass, 2005). Today there are more software projects running in geographically distributed environments and the so called “global software development is becoming a norm in the software industry” (Damian & Moitra 2006). One of the advantages of distributed system development is the fact that it provides opportunities for developers in dispersed locations to build and share their knowledge collectively. An example of distributed systems development is open source projects, characterized with a highly distributed environment, collaborative and rapid development among virtual teams and rapid evolution as the environment changes (Lee, Banerjee, Lim, Kumar, Hillegersberg and Wei, 2006). Similarly, the globalization’s effects on outsourcing of software production and systems development have made outsourcing take up global dimensions, and thereby become an international complex undertaking which requires a tremendous amount of support and interaction (Yalaho 2006).

As distributed system development emerges as the new paradigm in developing large-scale systems (Damian & Moitra 2006), there are still challenges and complexities involved in managing the development, such as cultural issues, communication issues and technical issues that need to be addressed (Herbsleb & Moitra 2001). Various studies have looked at aspects of managing virtual teams across time and space. Sarker and Sahay (2004) for example, identify problems experienced in distributed systems development, including those arising from geographical separation, different cultural contexts and different information systems development practices. Powell and associates (2006) similarly look at the relationship between team commitment, member effort, and trust in virtual teams as compared to collocated teams. DeLuca and Valacich (2006) examine the effectiveness of using asynchronous and synchronous media in virtual teams. In this paper, we draw upon the concept
of social capital to investigate how GridPP mobilizes its collective knowledge to overcome difficulties and challenges in its large scale distributed systems development project.

The concept of social capital has been traced back to Durkheim (Portes 1998). Bourdieu (1986) identified the value of social capital in the context of unequal power relations reproduced in societies. He defines it as “the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance or recognition” (ibid. p. 248). Coleman (1988) distinguished social capital from physical capital, embodied in material form, and human capital, embodied in the skills and knowledge acquired by an individual, and sees social capital as existing in the relations among persons. Later, Putman (1993; 2000) exported the concept out of academia and into a wider media. If Bourdieu and Coleman represent a sociological perspective on social capital, Putman took a political science perspective, whereas Fukuyama (1995) integrated the concept with trust in an economic framework, and defines it as “features of social organization such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit” (p. 57). Fukuyama’s work is considered representative of the “economic imperialist version of social capital” (Fevre 2000, p.103), which is based on a rationalistic model of individuals. In organization studies, the concept has also gained considerable currency and has been applied in a wide range of topics ranging from workers’ career advancement to firm strategies in a network (Adler & Kwon 2002). Adler and Kwon (2002) define social capital as “the goodwill available to individuals or groups. Its source lies in the structure and content of the actor’s social relations. Its effects flow from the information, influence, and solidarity it makes available to the actor.”

From the perspective of organizational advantage, Nahapiet et al. (1998) establish the linkage between social capital and collective knowledge in the organization. They argue that social capital facilitates the creation of new intellectual capital by affecting the conditions necessary for exchange and combination of existing intellectual resources, in the form of explicit and tacit knowledge and knowing capability. The linkage between knowledge processes and social capital has also been explored by information systems researchers. For example, Urquhart, Liyanage, and Kah (2008) suggest that the weaker the social capital, the harder it is for the knowledge and human capital to grow in a community. In this paper, we take the view that knowledge is socially constructed (Berger & Luckmann 1966), and that social capital draws our attention to the importance of social relationships, identities and values such as trust which facilitate the knowledge processes. Knowledge is co-constructed in a dialectic between individuals within the context of social structures (Berger et al., 1966). Individuals express themselves through language (broadly defined) which others must interpret to create new knowing (which may or may-not align with the previously known). It is for this reason that we adopt the perspective of social capital to explore practices of knowledge-construction and sharing of individuals and collectives, yet appreciating that knowledge cannot be objectified, it is always “known”. Documents and information can only be potential for knowing, rather than knowledge itself.

Nahapiet et al. (1998) conceptualize three dimensions of social capital: structural, cognitive, and relational dimensions. Structural dimension refers to the overall pattern of connections between actors – that is, who you reach and how you reach them. Drawing upon the term “relational embeddedness” by Granovetter (1992) which describes the kind of personal relationships people have developed with each other through a history of interactions, e.g. respect and friendship, the relational dimension of social capital refers to those assets created and leveraged though relationships. Key facets in this dimension include trust and trustworthiness, norms and sanctions, obligations and expectations, and identity. Finally, the cognitive dimension refers to those resources providing shared representations, interpretations, and systems of meaning among parties (Nahapiet et al. 1998, p. 244), for example, shared language and codes, and shared narratives. The distinction between the three dimensions of social capital is for analytical purposes and they are inevitably overlapping and interrelated. Within this paper we employ these social capital dimensions as analytical devices to explore the case of GridPP, and through this explore the social construction of knowledge within this community.
3 CASE DESCRIPTION

The LHC Computing Grid is building a large-scale computing infrastructure for the high energy physics community. The Large Hadron Collider (LHC) particle accelerator at CERN, the European Laboratory for Particle Physics, is designed to collide Hadron particles at energies close to those of the Big Bang in its search for the elusive ‘Higgs-Boson’ particle (believed to be responsible for matter having mass). These collisions will produce data within the LHCs four experiments (ATLAS, LHCb, ALICE and CMS). The number of collisions, and the subsequent data produced by the experiments, is vast, thus finding the Higgs-Boson has been likened to searching for “a person in a thousand world populations”, or for a “needle in twenty million haystacks. The LHC envisages producing 15 million gigabytes of data a year - equivalent to a DVD every 15 seconds or 1% of global information production (Lee et al. 2006). To store and analyze this data the LHC requires the equivalent of 100,000 PCs spread across the globe and working as a Grid (Economist 2005).

A technical perspective sees a grid as a computing platform for coordinated resource sharing and problem solving suitable for data-intensive and compute-intensive applications (Foster et al. 2001). In this way a grid connects and coordinates diverse and heterogeneous computing resources across space and different domains, presenting itself to users as though it was a single resource. A central concept for grids is that of the virtual organization (VO), and resource management is based on permissions for access to shared resources that members of a VO can make use of, disregarding actual hardware locations. Thus the four LHC experiments are examples of VO, and allow physicists from around the globe to access data and run analysis “jobs”.

The LCG has been broken down into various elements, and distributed among various countries. The UK contribution to the LCG is GridPP, a collaboration of 19 UK universities, the Rutherford Appleton Laboratory, and CERN. The GridPP project started in 2001 and has been involved in developing applications, middleware and providing technical infrastructure and storage and processing units. The LCG has a hierarchically tiered structure, with Tier 0 at CERN, Tier 1s consisting of the national IT centres in each of the major countries involved in the project, and Tier 2s being the regional centres in each country.

4 METHODOLOGY

GridPP’s unique nature provides a revelatory case of distributed systems development practice (Venters & Cornford 2006). Drawing from the interpretive research tradition in information systems the focus of this study is on sensemaking and the symbolic world of those studied (Walsham 1995). The research team includes a senior experimental particle physicist to ensure that the research is not undermined by a lack of understanding of physics.

Data collection began in August 2006, following earlier pilot work, and has included participant observations of weekly project management board meetings and deployment team meetings, quarterly GridPP collaboration meetings in the UK, international meetings of the LCG, site reviews carried out by GridPP, observation of various forums and conferences in which GridPP participates. The research team has had full access to the GridPP main documentation, and we subscribe to its main mailing list and the deployment team mailing list. At the core of this research are over fifty semi-structured qualitative interviews of between one and one-and-a-half hours, undertaken at various universities across the UK and during two week-long periods at CERN in Geneva. Table 1 provides details of the

---

1 Grids are distinguished from existing distributed computing in that scientists/users do not have to negotiate their use of different sites or resources separately and deal with security restrictions of individual sites, nor have to find out the precise location of their data. They need only one user-account to access a wide range of resources – processing and data – as permitted by being a member of a VO.
research activities undertaken. Interviews were audio-recorded, transcribed and then organized for analysis using the Atlas.Ti software. We found the software useful in handling the amount of data we had, but we didn’t want our thinking to be restricted by software. Rather than developing a theory from the data, our data analysis was closely integrated with theoretical development in an iterative and incremental process, one feeding into the other. Data analysis was closely integrated with theoretical development in an iterative process, one feeding into the other.

We can identify three stages of data analysis. The first round was open coding of the data, labelling aspects of the project, their practices, and emerging ideas from the phenomena. This is similar to axial coding in grounded theory (Glaser & Strauss 1967). But these relationships were not understood as indicating causality. In this process, some codes were merged, some became more general or more specific. Not all code families were included in our analysis, as some were considered interesting phenomena but not directly related to the key concepts. This was an iterative process until the key conceptual constructs were sufficiently refined. It should be noted that we also verify the findings from the interviews with a survey, which largely confirm all the themes that we derive from coding the interviews. In the end the analysis reported here is the result of the iterative reflections and ongoing discussions within the research team and with GridPP members, rather than a narrow machine-derived account. All the quotes given here are taken from interview transcripts, and our ideas have also been reinforced by informal conversations and participant observations.

5 SOCIAL CAPITAL & SYSTEM DEVELOPMENT IN GRIDPP

In this section we examine how collective knowledge is cultivated and mobilized in GridPP. The dimensions of social capital (Inkpen & Tsang 2005; Nahapiet et al. 1998) are drawn upon to explore various practices employed by members of GridPP in the distributed and collaborative project.

5.1 Structural dimension

5.1.1 A network structure and decentralized management

Figure 1 shows GridPP’s management structure which may be better described as a network than a hierarchy. The Project Management Board (PMB) is the heart of the network centrally coordinating the project. It provides quarterly reports to the Collaboration Board, consisting of representatives from the 19 institutes. The participating institutes enter the collaboration not under any legal obligation, but bound by a Memorandum of Understanding, which specifies the amount of resources and the level of
service that each site is committed to provide, and the funding and support they will receive from GridPP in return. This document serves more as a “gentlemen’s agreement” and thus there is no authoritative hierarchy between GridPP and the institutes, thus there are not top-down command lines between them. Decisions have to be made on a democratic basis and implemented by influence and persuasion.

The PMB, apart from the project leader, deputy project leader, and the project manager, consists of representatives from a number of internal and external committees, boards and functions. In a sense, the figure also indicates the external and internal environment in which GridPP operates. Due to role differences, some members of the PMB are more closely connected to the main GridPP activities than others. An important thing to note is that each of them serves as a node on this network, while connecting the GridPP to a variety of nodes on other networks that they represent here.

Associated with this networked structure of the project, members of GridPP are generally given huge reign of freedom to carry out their work, usually without clear instructions or strict supervision. When we asked them how they know what they need to do in their day-to-day job, most of them replied that they would look around and find out what needs to be done, as well as responding to arising problems and crises. “This environment is based on, if you want, charismatic leadership and people doing things relatively independent but also having the freedom to do them, and not having to report every two minutes on what they are doing.”

Due to the complexity and prototypical nature of the Grid, knowledge has often to be acquired in processes of exploration and experimentation. Moreover, as the technology is being developed and changing rapidly and constantly, it is difficult to keep documentation up to date. We did not observe a lot of effort in codifying and documenting of the Grid project in a structured way. As a result, most knowledge is embodied within the individual experts (Blackler 1995). Since most people in technical roles are on a contract of two to three years, there is a risk of losing expertise due to turnover. Nevertheless, the extensive communication channels within the community foster the social construction and reconstruction of knowledge through socialisation (Berger & Luckmann 1966). This focus on communication will be explored in the next subsection.

5.1.2 Extensive communication channels

Members of GridPP maintain a general understanding of the project, especially aspects related to their specific roles, through continuous and extensive communication flows in the community. Particle
physics collaborations are managed by what Knorr-Cetina (1999) refers to as “a fine grid of discourse”, channelling individual knowledge into the collaboration and providing it with a sort of “distributed cognition”. This web of communication includes a complex network of boards, committees, and working groups which are regularly holding meetings. One of the most important methods is the online virtual meeting. During these meetings wikis, webpages and blogs are used as stores of previous communication. There are also various mailing lists where there is constant exchange of questions and answers and where solutions emerge. Members of GridPP subscribe to mailing lists relevant to their own job functionalities to keep up with what’s going in the project.

More importantly, understanding and know-how of various aspects of the project are embodied in the key members of the project and carried to different clusters or groups of agents by these people sitting in various boards and a large number of meetings. Besides this formal management structure, most members of the GridPP agree that very often the more important things happen informally under face-to-face circumstances, e.g. over coffee breaks and meals, or (being a British community) “in the pub”.

Such extensive communications embody mutual monitoring and proactive sense-making. This resonates with communities of practice in which knowledge is shared through their collective communication and shared sense of identity (Brown & Duguid 1991; Wenger 1998; 2000). The structure of the project, reflective of the experimental nature of the field, provides a context in which knowledge is socially constructed (and so shared) through socialisation (Berger et al. 1966). The focus of work is thus upon exploration and sensemaking, rather than upon the following (or construction) of formalized, codified “designs” or “plans”.

5.2  Cognitive dimension

5.2.1  Shared goal

While GridPP does include some people from other fields, the majority (and all senior members) are from this “elite science” (Traweek, 1988) which is highly competitive to enter. One consequence is that members of the collaboration are strongly characterized by a shared common goal which is not to build a grid; but to enable the discovery of new physics and to probe the origins of the universe. This concept of a “shared goal” is frequently mentioned, and can be seen to bind efforts, solicit devotion, and bridge differences. A PMB member attributes GridPP’s organisation to “their history and the desire to sort of jointly achieve things”. For example, despite severe competition between experiments, they will willingly work together on grid development because it is required to do new physics: “I said I was proud of being a particle physicist, this is ‘cause particle physicists always get the job done; by and large because they are driven by one fundamental thing. They want their experiment to work when the beam gets into the accelerator, okay? And that transcends everything else they do.”

The common goal of serving Physics is an important source of motivation and commitment. As a recent physics PhD graduate pointed out: “They don’t work for money of course because particle physics doesn’t have a lot of money. They work because of their passion to do science. So they, we, strive to deliver the best result, to collaborate in the best way because we serve the same ideas and the same, the same passion.” Moreover, the shared vision provides a strong sense of direction, urgency and progress.

5.2.2  Shared culture

A shared culture in GridPP emerges from the physics background of most participants. The Grid project arose from the need for a tool to analyze data from the LHC; from the organizational memories of previous successful and innovative experiments; and from a history of cutting-edge computing. Moreover, it also emerges from a tradition and culture of strong commitment with a long term vision, pragmatic problem-solving and developing tools through improvisation and bricolage, as well as
respect for individual creativity and technical expertise. As one interviewee pointed out, “The particle physics community and the goals and the culture has had an enormous influence on GridPP... So it’s not that particle physics is telling us what to do, it’s just that we know it. It’s the culture.”

With decades of experience running experiments, it is not surprising that GridPP is set up in the structure of a particle physics experiment and largely managed in the same way. The established traditions and accumulated experience means working in large scale globally distributed collaborations is almost “second nature”. A member of the PMB commented “I think historically particle physics has this background in teamwork and this way of working you know, a very strong background in that. […] When I started off in experiments, there were about 20 people in an experiment, and now there’s 500 in the current experiment, ATLAS probably 2,000. So that […] we’ve all worked through this, of adapting to that sort of way of working, which is the sort of thing of course the Grid introduces, international projects, um, large scale resources, worldwide. So it’s sort of second nature I think…”

5.2.3 Norms

In face of the high level of complexity and uncertainty that GridPP faces, compound with limited time and resources, the way they do computing has to be very pragmatic and often on the basis of trial-and-error. In a project with such as a multitude of hardware and software elements, no one person can have a clear idea of the whole system; requirements cannot be pre-specified in detail; architectures are often conditional conjectures, and technical decision-making is emergent and empirical. As a result, GridPP has to be constantly adapting to changes, and practices in GridPP have been described as “ad hoc”. The PMB focused on accommodating change as their minimal planning process. “So we set up this project map and …the formality of change forms. So this was to formalise our freedom to change the project … yes, we had a set of milestones but you know, we had a mechanism to change them because we have to be responsive.”

Pragmatism is one the norms dominant in the particle physics community. Which is not only reflected in computing, but has roots in physics experiments as well: “not just in GridPP but in building hardware and building detectors - all sorts of setbacks occur and you have to find solutions. Certain technology doesn’t work, the company cannot provide what you want. So I think there’s this background in problem solving and project management and the sort of pragmatic approach”. This caused friction with the computer scientists working in the project, most of whom aspire to a more methodological approach. A technical expert commented, “…the people who come from a physics background are ultimately more pragmatic in computing. […]If it requires you to wrap sellotape around it to get it to work, then they will wrap sellotape around it. The people who come from a computing background tend to […] have slightly purer model of how the computing should work...whereas the physicists are happier with an ad hoc solution just to get the job done and push them through.”

5.3 Relational dimension

5.3.1 Sense of belonging

With members of the collaboration based in disparate institutes, it is important to develop an emotional bond among individuals for the project to function collectively. “We have to work very well together as a team, in order for GridPP to be successful. And because as you will appreciate it’s quite a complicated structure, there are multiple channels of communication, some of which are duplicated, some of which are contradictory, and there are all sorts of ways in which information flows. And anything that you can do to oil the cogs of the machine is going to help. And one of the things that are going on very well in GridPP is the cohesiveness of the deployment team. And I think for us to socialize together is a very important thing”. “Going to the pub” together when they meet, for
example, is one aspect of it. “It fosters a bond between people and helps... it helps a lot I think because many aspects of working in this project are frustrating because it's so large. And so if you can go out together and you can identify the problems and let out steam about them, I think that's actually a very important social function of these meetings.”

These emotional communications help to alleviate the anxiety they have to face, including the pressure of the LHC switch-on and that of showing the UK in a good light among the worldwide particle physics communities. The collaboration is committed, engaged, and is always “just about” on top of things. They seem to be constantly fire-fighting, discovering problems, managing crises, and negotiating solutions. Yet there is a high level of confidence despite the sense of urgency and disorder on the surface. Almost everybody in the collaboration that we interviewed holds a firm belief that the Grid will work; it may not work perfectly, but it will work.

5.3.2 Trust

Deriving from the shared tradition and shared goal is a high level of trust among people and institutions. One senior member of GridPP calls it “a culture of trust and equality”. The notion of trust also came out very clearly in our interviews: “it's very important for the physicists because there's so many things that they have to do in order to be able to interpret something that's been true in the data, that they have to trust what other people have done. And this is even more so when you have such big detectors as the LHC ones.”

In GridPP, trust not only lies with people’s intelligence and ability to deliver work, but more importantly, with people’s commitment to their job and with readiness to make extra effort to ensure things get done. “There will be problems coming out in all areas and you'll have to trust that people will step up to the plate and you know, do the dirty work as well as doing all the glamorous work you know, and things like that.” Trust seems to be what characterizes high energy physics as a community. “So actually the trust between the different high energy physics computing centres is much larger than what, in most of our member countries, are the legal constraints.”

6 DISCUSSION AND CONCLUSION

How does the mobilization of social capital contribute to the social construction of knowledge (Blackler & Mcdonald 2000; Berger et al 1966) within GridPP? Table 2 illustrates the implications of social capital in facilitating innovation and knowledge socialisation. A flat structure of organizing combined with minimal direct management provides a positive environment for creativity, exploration, and experimentation such that individuals can socialise knowledge. This structure similarly allows peripheral participation by non-experts so they may learn and enter the community (Wenger 1998; Lave & Wenger 1991). However, to be productive, there needs to be guidance of a clearly articulated common goal; in this case to advance the science of physics. Such a shared vision can be viewed as a bonding mechanism that helps different parts of a network integrate knowledge (Tsai & Ghoshal 1998), and provides an important sense of identity for the group (Lave et al. 1991).

Although there are extensive communication channels in GridPP, limited effort has been put on structured codification of knowing as information (Nonaka & Takeuchi 1995; Nonaka 1994). While mailing lists, wikis and blogs provide potential for knowing, they are used in a highly unstructured manner and are not seen as an important collective resource of knowledge. What compensates the lack of a codification approach to knowledge (Hansen, Nohria & Tierney, 1999) is the emphasis on socialization and personalisation within GridPP (Hansen et al. 1999). Most of the members travel frequently to attend meetings, conferences and workshops, and make deliberate efforts to talk to people face-to-face. The construction of social connections within the project provides personal directory (Wegner 1987) to knowing - i.e. “who knows what”. This process of socialization must however be sustained through the relational dimension of social capital, as illustrated by the ongoing
maintenance of trust and a sense of belonging. As the development stage of the grid draws to an end, and emphasis is shifted to maintenance and improving reliability of the Grid, such mobilization of social capital through formal and informal forms of interpersonal communication is gradually being balanced, but never replaced, with a systematic and impersonal arrangement of communication and technical support.

To conclude, the paper has drawn upon the concept of social capital to examine challenges faced in distributed systems development. The analysis on the empirical study of a complex project of system development allows us to reflect and explore dimensions of motivating, organizing, and coordinating distributed members of a systems development project and their implications for processes of collective knowing. Our findings are aligned with existing perception of social capital in the literature. Social capital has been found to enhance division of labour by reducing costs of coordination, induce and speed up sharing of tacit knowledge, and encourage co-operations, which in turn fosters firms’ innovative capability, competitiveness and ultimately economic performance (Maskell, 2000). In an era where distributed projects have become increasingly prevalent, our case generates implications for the importance of mobilizing social capital in distributed systems development processes, especially in terms of providing shared goals and culture, supporting individual creativity and building trusts in the community.

**Table 2 Social capital and knowledge processes**

<table>
<thead>
<tr>
<th>Social Capital</th>
<th>Processes of knowing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structural dimension</strong></td>
<td></td>
</tr>
<tr>
<td>Network structure</td>
<td>Enables free flow of knowledge</td>
</tr>
<tr>
<td></td>
<td>Avoids too many layers of decision making and bureaucracies</td>
</tr>
<tr>
<td></td>
<td>Allows knowing to emerge in a bottom-up manner, and circulate in the “ecosystem”</td>
</tr>
<tr>
<td>Decentralized management</td>
<td>High level of individual freedom provides incentives for sharing expertise and acquiring knowledge.</td>
</tr>
<tr>
<td></td>
<td>Encourages creativity and improvisation</td>
</tr>
<tr>
<td></td>
<td>Lack of centralized knowledge depository may lead to loss of knowledge</td>
</tr>
<tr>
<td>Extensive communication</td>
<td></td>
</tr>
<tr>
<td>channels</td>
<td>Encourages knowledge creation, especially across organizational and geographical boundaries</td>
</tr>
<tr>
<td><strong>Cognitive dimension</strong></td>
<td></td>
</tr>
<tr>
<td>Shared goals</td>
<td>Motivates knowledge acquisition of individuals</td>
</tr>
<tr>
<td></td>
<td>Provides incentives for social construction of knowledge</td>
</tr>
<tr>
<td></td>
<td>Fosters individual commitment and devotion</td>
</tr>
<tr>
<td></td>
<td>Encourages co-operation</td>
</tr>
<tr>
<td>Shared culture</td>
<td>Provides shared understanding</td>
</tr>
<tr>
<td></td>
<td>Facilitates communication</td>
</tr>
<tr>
<td></td>
<td>Lower barriers from the tacit dimension of knowing (Polanyi 1967).</td>
</tr>
<tr>
<td>Norms (of pragmatism)</td>
<td>Provides a “safety-net” which tolerates mistakes while encouraging creativity and problem solving</td>
</tr>
<tr>
<td><strong>Relational dimension</strong></td>
<td></td>
</tr>
<tr>
<td>Sense of belonging</td>
<td>Socialization facilitates sharing of tacit knowledge</td>
</tr>
<tr>
<td></td>
<td>Supports collaboration and knowledge integration</td>
</tr>
<tr>
<td>Trust</td>
<td>Promotes willingness to share knowledge</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
7 REFERENCES


Herbsleb, D. J., Paulish, J. D. and Bass, M. (2005) Global software development at siemens: Experience from nine projects. International Conference on Software Engineering (ICSE'05), ACM, St. Louis, Missouri, USA.


Proceedings ECIS 2009
USER-GENERATED CONTENT (UGC) IN TOURISM: BENEFITS AND CONCERNS OF ONLINE CONSUMERS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0121.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Web 2.0, Blogging, Business value of IT, Consumer behavior / choice / demand / empowerment / reviews / consumerism</td>
</tr>
</tbody>
</table>
Abstract

This paper examines the views of travel consumers that search for information online in relation to the specific benefits and concerns identified with user-generated content (UGC). Real contrasts in relation to views about UGC content were identified in the literature. For instance, UGC is perceived as being ‘credible’ or ‘not credible’ as an information source depending upon the view of the user. The paper reports research that examined the use of UGC by online travel consumers, sourced from an Australian tourism organisation’s online subscriber database. The study highlighted a similar range of areas that were seen as benefits and concerns in the use of UGC to support travel decisions as identified in the literature – and also similar contrasts. The most notable of these is the level of trustworthiness and reliability that is associated with the postings – they can be trusted because they are ‘real’ experiences by ‘real’ people who are independent – but can also be not trustworthy because the content that is posted may be faked by someone with a vested interest. Similar paradoxes existed elsewhere. One of the theoretical contributions of the paper is the derivation of a matrix for classifying recommendations associated with online UGC.

Keywords: tourism; websites; user-generated content; benefits; concerns; survey
USER-GENERATED CONTENT (UGC) IN TOURISM: BENEFITS AND CONCERNS OF ONLINE CONSUMERS

1 INTRODUCTION

Consumers are increasingly using different strategies to find information on the Internet about products or services they are intending to purchase. Senecal and Nantel (2004) discuss literature that relates consumers’ choices in relation to such information sources to the type of product or service being sought, noting that goods can possess either search or experience qualities. Information about goods with search qualities can be determined prior to purchase – that is, much can found out about the product or service beforehand. Information about goods with experience qualities cannot easily be determined before purchase, and it is for these types of goods that consumers will often rely heavily on product recommendations from others. Travel products and services fall directly into the category of being experience goods (Bei, Chen and Widdows, 2004). Senecal and Nantel (2004) suggest that consumers who had previously consulted a product recommendation were more likely to purchase that product than those who did not. However, in the absence of recommendations, consumers looking to purchase conducted even more searches for information (Smith, Menon & Sivakumar, 2005). Consumers were more influenced by recommendations associated with an experience product (for example, wine) than for a search product (e.g. calculator) – hence, it would be expected that this also applied to the tourism area. One of the more recent sources of online information for consumers is user-generated content (UGC), where travellers are able to examine text, images, and videos that have been posted online by fellow consumers.

The specific aim of this paper was to capture the perceptions and views of travel consumers that search for information online in relation to the specific benefits and concerns that they might have when dealing with different types of travel recommendations – more particular UGC. Hence, the paper makes a contribution to the existing literature by reporting the perceptions of online travellers and how they view websites that publish information originating with other users. A further contribution of the paper is the derivation of a matrix for classifying recommendations associated with online UGC.

2 FINDING TRAVEL INFORMATION ONLINE

High credibility is often placed on word-of-mouth (WOM) information as it is believed that the person providing the WOM information has nothing to lose and is therefore more likely to offer honest advice (Chatterjee, 2001). Furthermore, well-reasoned, logical and persuasive reviews can positively influence the likelihood of purchase by people who read them (Park, Lee & Han, 2007). However, when negative word-of-mouth is provided with the specific goal to vent frustration or anger, its influence on the receiver tends to be reduced, as the information is not perceived to be constructive or useful (Wetzer, Zeelenberg & Pieters, 2007). In a travel sense, consumers often spread WOM due to extreme feelings associated with a product ‘experience’, such as pleasure or sadness. In some instances, sharing the pleasure of the travel experience is seen as being part of the positive experience (Litvin, Goldsmith & Pan, 2008). Amongst the vast array of information sources consumers can turn to when planning travel, word-of-mouth is one of the most influential. Prospective travellers often rely heavily on advice from friends, family and other peer groups, particularly when planning travel to a destination not previously visited (Litvin et al., 2008). Research indicates that, due to the lack of commercial self-interest associated with WOM recommendations consumers tend to trust and be more influenced by this type of information than by more commercial sources such as travel agents or accommodation operators. This is because the provider of the information is not generally aiming to make a financial gain from sharing their experiences and views with others (Litvin et al., 2008). One of the aims of this article is to examine whether these circumstances translate to the online environment.
2.1 Recommendations in the Online Environment

In the online world, recommendations sources can come from other consumers, human experts or expert systems (that will recommend a particular product or service based upon a consumer’s profile – and are often known as recommender systems) (Senecal and Nantel, 2004). Consumers can use the Internet to mimic sources of information that they would have traditionally received from ‘real world’ sources. Some examples include (Peterson and Merino 2003; Litvin et al.,2008):

- Visiting websites to access information, instead relying on the traditional mass media advertising and/or information normally acquired from a salesperson
- eWOM (electronic word-of-mouth) instead of traditional WOM
- Accessing online independent sources (such as government tourism bodies) instead of their offline counterparts.
- Email- allowing the receipt of subscribed newsletters or as a communication form that facilitates globally correspondence and information exchanges

These days, many websites allow consumers to add their own content in the form of general text comments (weblogs or ‘blogs’), travel reviews, pictures and/or video. This content is known as User Generated Content (UGC). UGC sites can equate to electronic WOM marketing, whereby somebody who has an opinion about a product or service shares their views, beliefs and experiences with other people (Ahuja, Michels, Walker & Weissbuch, 2007). Fernando (2007) suggests that UGC is the opposite to traditional forms of media and marketing since content is generated by the consumer rather than by the marketer. Tools such as blogs and social networking sites (such as Facebook) have meant that consumers are better informed than ever before – not only being able to add their own comments, but also being able to find other information and articles and ‘tag’ them with their own keywords for search purposes (Buhler, 2006). Social networking sites typically operate by inviting people to join and contribute to a network. As this process snowballs the networks can grow. These communities rely upon UGC for their ongoing operation (Trusov, Bucklin and Pauwels, 2008). Dwyer (2007) suggests that websites containing these new media (such as message boards, chat rooms and now blogs) provide two types of networks – social networks and informational networks. In this article the authors adopt the more general view of UGC, that it can provide a means of social interaction for users but it is also an important source of information.

In the online context, eWOM occurs when consumers create their own information on the Internet to share their experiences and views about products they have purchased (Park et al.,2007). There is typically far more information available to the consumer in the online environment from eWOM than from traditional WOM (Chatterjee, 2001). Park et al.(2007) found that the purchasing intentions of consumers increased in line with the number of reviews that indicated that the product or service was popular.

The provision of forums that capture consumer comments potentially allow a business to receive genuine feedback on their products or services. It also provides them with another avenue in which to provide their own feedback and/or reassurance to those customers. However, there are also a number of forums that are not sponsored by businesses – such as general social networking websites like Facebook. These forums do not always attract comments from ‘typical’ consumers – in fact, as with traditional WOM, it is more likely that consumers who have had extreme (very favourable or very unfavourable) experiences are more likely to provide online comments or reviews. However, these sites could be regarded as being more neutral than those sites sponsored by businesses (Litvin et al.,2008).

In the tourism context, the Internet is an important source of information for travellers. For instance, a majority of US travellers use the medium to search for travel information (Litvin et al.,2008). As with other forms of online information, UGC related to travel can be posted on specific travel-related websites or on more generic social networking websites. This content reflects the experiences of the tourist at specific destinations (Pan, MacLaurin and Crotts 2007) or with other travel products.
2.2 Trustworthiness of user-generated content

There is no universally accepted definition of ‘trust’. Chen (2006) discusses two schools of trust. The first school regards trust as a belief or expectation about another party’s trustworthiness. The other school regards trust as a behaviour that reflects a reliance on others and some uncertainty (and vulnerability) from the person who is ‘doing the trusting’. The difference is in how trust is actually measured in a research context. Chen (2006) adopts the latter view of trust and identifies three dimensions of trust: the level of competence, the level of benevolence and the level of integrity. However, this view generally relates to the relationship between the consumer and provider – so in this instance, where we are examining the trustworthiness of UGC – it is more appropriate to adopt the first school, where we consider the trustworthiness of the party providing the UGC comment.

Criticism regarding the power of UGC to persuade travelers about travel related decisions is based on the potential for ‘fake’ content to be posted by travel operators posing as independent reviewers. This effectively defeats the purpose of enabling UGC to influence travelers in their decision making process as the content added is no longer independent, objective or credible (Bray & Schetzina, 2006). One of the concerns raised about the use of UGC sites when planning travel is how the consumer can be assured that the reviews they are viewing are in fact independent and hence trustworthy (Gretzel, 2006). One of the major concerns here is that businesses might use employees to ‘act’ as consumers to post positive comments on behalf of the business or to post negative comments about the competition (Litvin et al., 2008). Senecal and Nantel (2004) note that many consumers are sceptical about any form of communication that is perceived to be skewed towards the interests of the source of the information. Park et al. (2007) suggest that online consumer reviews are often considered more trustworthy and credible than information provided by suppliers of products and services, assumedly because consumers are considered to provide more honest information. Websites that are independent, third-party type sites tend to be considered more preferable by consumers when compared to those that are clearly operated by a business with a vested interest (Senecal & Nantel, 2004). Thus, the forum in which recommendations are presented is quite important. A possible downside of UGC, is that while traditional forms of WOM tends to come from people who are known to the consumer (i.e. friends, colleagues etc), online reviews are typically passed on by total strangers, resulting in some concern over the credibility of the source of review (Park et al., 2007, Litvin et al., 2008).

Table 1 provides a summary of the potential benefits and concerns of consumers in relation to online recommendations as identified in the literature. These are divided into three major categories; those that relate to the content of the recommendation, the source of the recommendations, and the nature in which recommendations might be encountered in the online environment. Obviously, these categories are linked (for instance, where commercial sources are seen to post fake content to bias the behaviour of potential travellers). The Content category relates to extreme or emotional postings versus well-argued postings and the possibility of fake content. The Source category relates to the notion that eWOM provides more recommendation sources and the credibility levels of independent and consumer recommendations versus those of businesses with vested interests and comments posted on social networking sites. The nature of recommendations refers to their influence due to the nature of tourism as a product and the ability of tourists to filter the wealth of information available online. The categories are groupings which have emerged from the literature and should be viewed as a starting point for this type of classification.

What we find most interesting here is that there seem to be real contrasts in relation to views about UGC content. For instance, it can be viewed as being ‘credible’ or ‘not credible’ as a source and there are both benefits and concerns about the content that is posted. This reflects the different opinions that users, and even contributors, have depending upon their different views and experiences. In a study of the content of travel blogs, Pan et al. (2007) divided a number of blog comments into ‘positive’ and ‘negative’ statements and found that there were both types of sentences across most of the categories (the ones with the most comments being attractions, amenities, history, food and beverages and natural
environment). In most cases the number of positive comments outweighed the number of negative comments (overall there were around three positive comments for every negative comment).

<table>
<thead>
<tr>
<th>Category</th>
<th>Benefits</th>
<th>Concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>Well-reasoned, logical, persuasive reviews can positively influence purchase decisions (Park et al., 2007)</td>
<td>Not useful when venting frustration or anger (Wetzer et al., 2007); Those with very positive or very negative views tend to post comments (Litvin et al., 2008)</td>
</tr>
<tr>
<td>False content</td>
<td></td>
<td>Posted by travel operators (Bray &amp; Schetzina, 2006).</td>
</tr>
<tr>
<td>Source</td>
<td>Sources more plentiful (Duhal et al., 1997; Chatterjee, 2001); Access to more reviews than with WOM (Park et al., 2007)</td>
<td>Source not known to the consumer (Duhal et al., 1997)</td>
</tr>
<tr>
<td>eWOM and ‘weak’ ties</td>
<td>Credible as source as nothing to lose by offering advice (Chatterjee, 2001; Litvin et al., 2008); Consumers and independent bodies can be neutral source (Bei et al., 2004; Park et al., 2007)</td>
<td>Businesses may appear to have vested interest (Litvin et al., 2008); Content of social networking sites is not as trustworthy as that of corporate websites (Wasserman, 2006)</td>
</tr>
<tr>
<td>Credibility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Influence</td>
<td>More likely to influence as tourism is an ‘experience’ product (Senecal and Nantel, 2004)</td>
<td></td>
</tr>
<tr>
<td>Filter</td>
<td>Can assist tourists to manage the wealth of data (Smith et al., 2005)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Benefits and Concerns of Online Travel Recommendations

3 THE STUDY

In this study we are predominantly interested in the views of consumers that use the Internet to assist with travel plans in relation to their opinions about UGC. However, linked with this is the idea that they need to place their use of UGC in context with other travel information and services that are available to them online. Thus, our more general research question relates to whether the benefits and concerns of online travel recommendations (as identified in Table 1), translate specifically to the use of UGC for travel purposes. In particular we will be looking to see if the benefits (in relation to well-reasoned reviews, plentiful sources of information and credibility) and concerns (such as extreme opinions, fake content, unknown sources and low credibility) actually do exist.

To explore consumers’ views on UGC in relation to travel planning, a quantitative study was conducted using an online survey of consumers who were known to use the Internet to gather information when planning their travel. The survey was developed based on a review of existing studies, as outlined in the previous section. It contained four key sections, the first of which gauged participants’ previous exposure to sites containing UGC related to travel. Section 2 assessed the influence of UGC sites on participants’ actual travel planning and trip behaviour. This section included two open ended questions that asked respondents to express what they liked about UGC postings regarding travel as well as any concerns they had about this type of information. Section 3 gauged the types of UGC considered most useful along with opinions about how UGC should be used.
by hospitality and tourism providers in conjunction with their existing online marketing strategies. The final section asked for the demographic characteristics of participants.

Following a pilot test of the instrument with a sample of real travelers, the final survey was conducted online in December 2007. A web-link to the survey was included in an invitation, sent via email, to participate in the research promoted through Tourism New South Wales’s database of email subscribers known as E-Scapes. At the time of dispatching the email invitation to promote the survey, there were approximately 110,000 subscribers listed on the database. An incentive prize was included with the email received by subscribers to encourage responses. Recipients of the email who chose to participate in the survey, on a voluntary and anonymous basis, simply clicked on the web-link provided and responded to the survey questions online.

Data was collected over a two-week period and the survey took approximately 10 minutes to complete. All data received was contained within a downloadable spreadsheet from the survey software that was then converted into SPSS (Statistical Package for the Social Sciences) for further analysis.

By the survey closing date, 13,281 people had participated in the study. This represents a response rate of approximately 12 percent. It should be noted that not all respondents answered every question in the survey, as they were given the option not to answer questions if they so chose. Furthermore, some questions were not asked of all respondents (e.g., names of UGC sites they had used) where their previous responses to questions indicated a question was irrelevant. A typical example of this is where respondents were asked a series of questions related to their opinions of UGC. These questions were only shown to respondents that had indicated they had used UGC websites. These factors should be taken into account when noting the total number of responses reported in the various tables in this section. Approximately 700 responses were not considered useful due to a lack of data, so in effect the usable number of responses was 12,544. The survey contained many questions and 64% of those who commenced the survey actually completed it.

3.1 Respondent Profile

A demographic profile of survey participants is provided in Table 2. The questions related to age, gender, country of origin, place of origin and income level were asked towards the end of the survey. The number of respondents that completed these questions (around two thirds of usable responses) closely matches the overall completion rate of the survey as indicated in the previous section (64%).

The age profile of participants in this study reflects the overall profile of the E-Scapes database. Approximately 51 percent of people were aged 30-49 years. A further 23 percent were 50-59 years. In terms of gender, the skew towards a higher proportion of female participants (61%) is reflective of the overall profile of the database. The profile of responses to the survey appears, therefore, to be able to be generalized to the population of travelers included in the database.

3.2 Benefits of and Concerns with UGC

In order to ensure a consistent meaning for UGC, respondents were presented with the following description at the beginning of the survey:

_A growing number of web sites are incorporating features which enable the user, such as you, to contribute their own content enabling people to communicate about special interest topics or products or services through the Internet. Such content is commonly referred to as ‘user-generated content’._

Proceedings ECIS 2009
In relation to travel and tourism, some examples of user-generated content include:

- ordinary people like yourself sharing their opinions about travel destinations, attractions and accommodation properties through blogs (weblogs) or other discussion forums
- travellers submitting photos or videos to the internet to share their travel experiences with other online users (including family, friends or total strangers who may be interested)
- consumers posting reviews of accommodation properties to sites such as tripadvisor.com
- people using social networking sites such as myspace.com, facebook.com or youtube.com to share travel information.

Respondents were asked if they had visited any Internet sites that had contained UGC (5,724 indicated that they had). To understand the issues that were of potential concern to travellers, the survey asked these respondents two open-ended questions:

- ‘What do you like about sites that contain user-generated content related to travel?’
- ‘In relation to making travel plans, is there anything about UGC that concerns you?’

Thus, these questions related to the opinions of users about websites with UGC content and did not involve any analysis of UGC content itself, such as those carried out by Dwyer (2007) and Pan et al (2007). A total of 2,546 respondents (44% of those that had visited websites with UGC) listed ‘likes’ about UGC content and 1,238 (22%) respondents indicated they held some concerns about UGC. This may relate to Pan et al’s (2007) study which found that positive comments within a blog outweighed the number of negative comments. The responses received to these questions were analysed in the qualitative software package XSight, which enables comments to be classified under key headings that reflect particular ‘likes’ and ‘concerns’ which arose.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Category</th>
<th>N</th>
<th>% responded</th>
<th>% of database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age Group</td>
<td>Under 19 years</td>
<td>39</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>20-29 years</td>
<td>893</td>
<td>10.0</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td>30-34 years</td>
<td>972</td>
<td>11.4</td>
<td>12.1</td>
</tr>
<tr>
<td></td>
<td>35-39 years</td>
<td>1091</td>
<td>12.7</td>
<td>13.3</td>
</tr>
<tr>
<td></td>
<td>40-44 years</td>
<td>1155</td>
<td>13.5</td>
<td>13.9</td>
</tr>
<tr>
<td></td>
<td>45-49 years</td>
<td>1161</td>
<td>13.6</td>
<td>14.3</td>
</tr>
<tr>
<td></td>
<td>50-54 years</td>
<td>1103</td>
<td>12.9</td>
<td>13.0</td>
</tr>
<tr>
<td></td>
<td>55-59 years</td>
<td>874</td>
<td>10.2</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td>60-69 years</td>
<td>1056</td>
<td>12.3</td>
<td>10.7</td>
</tr>
<tr>
<td></td>
<td>70 plus years</td>
<td>223</td>
<td>2.6</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>8567</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>5235</td>
<td>61.4</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>3292</td>
<td>38.6</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>8527</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Country of Origin</td>
<td>Australia</td>
<td>8273</td>
<td>97.0</td>
<td>99.0</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>259</td>
<td>3.0</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>8532</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Place of Origin</td>
<td>New South Wales</td>
<td>6210</td>
<td>72.8</td>
<td>68.1</td>
</tr>
<tr>
<td></td>
<td>Victoria</td>
<td>934</td>
<td>10.9</td>
<td>8.2</td>
</tr>
<tr>
<td></td>
<td>Queensland</td>
<td>713</td>
<td>8.4</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td>Other States</td>
<td>362</td>
<td>4.2</td>
<td>14.5</td>
</tr>
<tr>
<td></td>
<td>Overseas</td>
<td>259</td>
<td>3.0</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>Australia – not stated</td>
<td>54</td>
<td>0.7</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>8532</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Gross Household Income/Year</td>
<td>Less than $52,000</td>
<td>2178</td>
<td>27.4</td>
<td>28.8</td>
</tr>
<tr>
<td></td>
<td>$52,000 - $77,999</td>
<td>1738</td>
<td>21.8</td>
<td>24.5</td>
</tr>
<tr>
<td></td>
<td>More than $78,000</td>
<td>4041</td>
<td>50.8</td>
<td>46.7</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>7957</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 2: Profile of Survey Respondents
4 RESULTS

4.1 Benefits of UGC sites

Respondents to the survey were asked (in separate questions) to identify their ‘likes’ and ‘concerns’ in relation to UGC websites. The answers to these questions provided the authors with an opportunity to match these responses with the benefits and concerns of online travel recommendations as listed in the literature (and summarised in Table 1). As mentioned earlier, the ‘likes’ and ‘concerns’ were classified according to various headings and subheadings. Table 3 shows the ‘likes’ classified into their main heading groups.

<table>
<thead>
<tr>
<th>Likes</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trust in the source</td>
<td>1048</td>
<td>41</td>
</tr>
<tr>
<td>Traveller opinions</td>
<td>396</td>
<td>15</td>
</tr>
<tr>
<td>Relevance to user</td>
<td>225</td>
<td>9</td>
</tr>
<tr>
<td>Recommendations</td>
<td>155</td>
<td>6</td>
</tr>
<tr>
<td>Amount of information</td>
<td>137</td>
<td>5</td>
</tr>
<tr>
<td>General comments</td>
<td>124</td>
<td>5</td>
</tr>
<tr>
<td>Ease of use / speed</td>
<td>121</td>
<td>5</td>
</tr>
<tr>
<td>Specific comments</td>
<td>106</td>
<td>4</td>
</tr>
<tr>
<td>Other</td>
<td>85</td>
<td>3</td>
</tr>
<tr>
<td>Don't like / unsure</td>
<td>68</td>
<td>3</td>
</tr>
<tr>
<td>Currency</td>
<td>64</td>
<td>3</td>
</tr>
<tr>
<td>Share experiences</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>2546</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 3: Classification of ‘Likes’ in relation to UGC websites

Many of the comments that were made fitted across a number of categories list in Table 3. Where this occurred the comment was classified according to its primary emphasis, or the first occurrence of the classification in the comment. The main classification of comments occurred as follows:

- **Trust in the Source**: a number of different types of comments were classified under this heading. Typically, comments that highlighted the ‘credibility’, ‘accuracy’ or ‘authenticity’ of UGC content and the ‘balanced’ nature of UGC fitted into this category. In addition, comments that related to the type of person making UGC comments fitted here, such as UGC was generated by ‘real people’, ‘independent people’ or ‘everyday people’. Finally, there were comments that related to the authenticity of the experience of the people posting comments – they had ‘been there, done that’, they were ‘genuine experiences’ and so forth. Comments in this section were also highlighted by words such as ‘honest’, ‘candid’, ‘frank’, ‘truthful’, ‘unbiased’, ‘unsolicited’ and so forth.

- **Traveller opinions**: these were typified by comments that related to the ‘range’ or ‘variety’ of opinions available in UGC and how they could be ‘compared’. One interesting aspect of this category is that a number of respondents (59) actually suggested that the fact that there were both ‘positive and negative’ comments was a good thing as it allowed for comparisons of experiences.

- **Relevance to user**: comments were classified into this category when respondents indicated that UGC assisted in helping with their ‘own travel plans’, was ‘personalised’ to their interests or was posted by ‘like minded people’. Comments in this latter category were posted by a number of elderly respondents and also those that were looking for travel experiences suitable for families.
• Recommendations: these comments involved respondents identifying that UGC content provided specific recommendations for potentially new (sometimes ‘off beat’) experiences, or offered tips, hints or advice on what to do or expect in different situations.

• Amount of information: comments in this category referred to the quantity of information available in UGC. These comments were highlighted by words such as ‘comprehensive’, ‘lots’, ‘more’, ‘detailed’ and reference to the ‘large number of opinions’.

• General comments: no specific comments about UGC were made in this category. These were just comments by respondents that referred to UGC as being generally ‘interesting’, ‘useful’, ‘helpful’, ‘informative’ or just providing ‘information’.

• Ease and speed of Use: these comments referred to UGC websites as being ‘user friendly’, ‘easy to use’, ‘accessible’, ‘saving time’ or ‘convenient’.

• Specific comments: most of these comments related directly to the use of UGC for booking accommodation or the fact that traveller pictures are posted on UGC websites.

• Don’t like/ unsure: Interestingly, 68 respondents chose to use the ‘likes’ of UGC content to say that they did not like UGC websites or that they were unsure what UGC actually is.

• Currency: these comments related to UGC content being ‘current’ or ‘up to date’.

• Share experiences: a small number of respondents suggested that they liked UGC because they could share their experiences with each other.

• Other: these comments did not fit any of the above categories.

Table 4 shows a comparison between the benefits of online travel recommendations that were identified in Table 1 and the results of the UGC study. Overall, most of the benefits that were mentioned in the literature were well-represented in the study, the most evident being comments related to credibility. The only ‘benefit’ that did not stand out in the study was that there were not a lot of comments by respondents that suggested they were persuaded by well-reasoned or logical reviews.

<table>
<thead>
<tr>
<th>Category</th>
<th>Benefits (Literature)</th>
<th>Our study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>Argument; extreme opinions</td>
<td>Well-reasoned, logical, persuasive reviews can positively influence purchase decisions</td>
</tr>
<tr>
<td></td>
<td>Sources more plentiful; Access to more reviews than with WOM</td>
<td>Not evident in responses</td>
</tr>
<tr>
<td>Source</td>
<td>eWOM and ‘weak’ ties</td>
<td>Sources more plentiful; Access to more reviews than with WOM</td>
</tr>
<tr>
<td></td>
<td>Credibility</td>
<td>Credible as source as nothing to lose by offering advice; Consumers/ independent bodies can be a neutral source</td>
</tr>
<tr>
<td>Nature of Recommendations</td>
<td>Influence</td>
<td>More likely to influence as tourism is an ‘experience’ product</td>
</tr>
<tr>
<td></td>
<td>Filter</td>
<td>Can assist tourists to manage the wealth of data</td>
</tr>
</tbody>
</table>

Table 4: Comparison of Literature Online Travel Recommendation benefits with ‘likes’ of UGC websites from our study

4.2 Concerns about UGC sites

To understand the issues that were of potential concern to travellers, the survey asked respondents an open-ended question - ‘in relation to making travel plans, is there anything about user-generated content that concerns you?’ A total of 1,238 respondents indicated that held some concerns about UGC. Refer to Table 5 for a summary of these responses. The responses received to this question were also analysed in a qualitative software analysis package (XSight). Table 5 lists the key concerns used to classify respondents’ comments about UGC.
<table>
<thead>
<tr>
<th>Concern</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trustworthiness/ reliability</td>
<td>599</td>
<td>48</td>
</tr>
<tr>
<td>Lack of relevance to user</td>
<td>267</td>
<td>22</td>
</tr>
<tr>
<td>Extreme opinion</td>
<td>176</td>
<td>14</td>
</tr>
<tr>
<td>Other</td>
<td>132</td>
<td>11</td>
</tr>
<tr>
<td>Security/ privacy concerns</td>
<td>64</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1238</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

Table 5: Classification of ‘Likes’ in relation to UGC websites

Under the heading of trustworthiness/ reliability the key concern identified by respondents is that UGC sites provide an opportunity for businesses to masquerade as independent travellers and post their own ‘fake’ entries. The second identified concern related to the extent to which comments made on UGC sites are trustworthy, accurate and credible for reasons other than the potential for business to tamper with UGC. Other respondents noted that it can be difficult to base one’s own travel decisions on the types of comments made on UGC sites as other travellers may well have different preferences to their own. Linked to the issue of the identity of travellers, a number of people also noted that it was hard to make a judgement about the value of UGC comments when the profile of the person submitting the information is generally unknown. These comments were classified under the heading of ‘Lack of relevance to the user’. Further concerns identified related to the tendency for people to contribute commentary to UGC only when they had very positive or very negative things to say (i.e. UGC can tend to be biased rather than represent the ‘average’ travellers experience). Some respondents indicated that UGC sites can be used by ‘habitual complainers’ to vent their negative views without providing a realistic assessment of the overall travel experience. These comments were classified as ‘extreme opinions’.

Table 6 shows a comparison between the concerns of online travel recommendations that were identified in the literature (refer Table 1) and the results of the UGC study. As with the ‘benefits’, a large proportion of the comments related to the trustworthiness or reliability of UGC – in this instance it was being viewed as a concern. Within that category, there were numerous comments related to all of the areas that were identified in the literature – concerns about fake content (28% of all respondents), the source not being known to respondents (4%) and comments related to concerns about the credibility of content (14%). There were also concerns about the usefulness of comments that were classified as ‘extreme opinions’ – the main problem identified with these was that they were too biased or that the websites were just being used as a ‘complaints forum’. There was one issue of concern that emerged from the study that was not evident in the previous literature – the practice of comments being posted that were not relevant to the user. In these instances there were concerns that the content was being posted by travellers with different personal preferences, that they were only subjective or personal opinions anyway, or were just out of date. This contrasts the ‘benefit’ identified (refer Table 4) which suggested that UGC helped tourists manage the wealth of data as it was relevant to their needs.
### Table 6: Comparison of Literature Online Travel Recommendation concerns with 'concerns' of UGC websites from our study

<table>
<thead>
<tr>
<th>Category</th>
<th>Concerns</th>
<th>Our Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>Argument; extreme opinions Not useful when venting frustration or anger; Those with very positive or very negative views tend to post comments</td>
<td>Extreme opinion (14%)</td>
</tr>
<tr>
<td></td>
<td>Fake content Posted by travel operators</td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>eWOM and 'weak' ties Source not known to the consumer</td>
<td>Trustworthiness/reliability (48%)</td>
</tr>
<tr>
<td></td>
<td>Credibility Businesses may appear to have vested interest; Content of social networking sites is not as trustworthy as that of corporate websites</td>
<td></td>
</tr>
<tr>
<td>Nature of recommendations</td>
<td>Influence Not evident in literature</td>
<td>Not evident in study</td>
</tr>
<tr>
<td></td>
<td>Filter Not evident in literature</td>
<td>Lack of relevance to user (22%)</td>
</tr>
</tbody>
</table>

4.3 Summary

The literature highlighted a series of ‘paradoxes’ in relation to the benefits and concerns associated with online travel recommendations. This study was able to match and in some instances further enhance the findings from the previous literature. The paradox appears to be that for just about every benefit (‘like’) that is identified by survey respondents there is a matching ‘concern’ that is also identified. In fact, the study revealed that a number of respondents mentioned ‘extreme opinions’ as being a ‘like’ of UGC, whilst it was mentioned by others as being a concern. Of course, many respondents mentioned the credibility of UGC as being a benefit, whilst it was also mentioned as being a concern. Respondents also referred to UGC content being ‘relevant’ as a benefit, whilst others suggested that a concern was that it was not relevant to them.

The reader is referred to Table 7. Of the 1238 respondents that had listed a concern about UGC, 1179 (95%) also listed a ‘like’ about UGC. In fact, of the 1179 respondents that listed both a ‘like’ for and a ‘concern’ about UGC content:

- 274 suggested a ‘like’ related being able to ‘trust the source’ as well as a concern about the ‘trustworthiness/reliability’ of the source in the same response! In both instances (‘likes’ and ‘concerns’), the proportion of respondents that fitted into this category was higher then the proportion of general respondents that identified trust as only a like or concern. Whilst this seems quite strange, the range of comments that were made in this category for both ‘likes’ and ‘concerns’ makes it quite feasible. For instance, here are some examples of responses:
  - Like: real photographs; real experiences; Concern: fake content
  - Like: frankness of comments; Concern: commercial interest posting comments
  - Like: honesty of comments of travellers; Concern: comments might be ‘made up’ by businesses

- 28 respondents suggested that they ‘liked’ traveller opinions on UGC, but had concerns about the validity of extreme opinions.

- 16 respondents that suggested in the UGC benefits that the information in UGC was relevant for them also suggested in the concerns that it could also not be relevant!
Table 7: 'Likes' of and 'Concerns' about UGC content by respondents that expressed both

<table>
<thead>
<tr>
<th>Likes</th>
<th>Trustworthiness/reliability</th>
<th>Lack of relevance to user</th>
<th>Extreme opinion</th>
<th>Security/privacy concerns</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trust in the source</td>
<td>274</td>
<td>125</td>
<td>80</td>
<td>6</td>
<td>40</td>
<td>525</td>
</tr>
<tr>
<td>Traveller opinions</td>
<td>96</td>
<td>40</td>
<td>28</td>
<td>3</td>
<td>24</td>
<td>191</td>
</tr>
<tr>
<td>Relevance to user</td>
<td>45</td>
<td>16</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>88</td>
</tr>
<tr>
<td>Recommendations</td>
<td>31</td>
<td>16</td>
<td>12</td>
<td>5</td>
<td>4</td>
<td>68</td>
</tr>
<tr>
<td>Amount of information</td>
<td>21</td>
<td>14</td>
<td>10</td>
<td>5</td>
<td>7</td>
<td>57</td>
</tr>
<tr>
<td>Specific comments</td>
<td>23</td>
<td>9</td>
<td>5</td>
<td>7</td>
<td>11</td>
<td>55</td>
</tr>
<tr>
<td>General comments</td>
<td>19</td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>42</td>
</tr>
<tr>
<td>Ease of use / speed</td>
<td>12</td>
<td>8</td>
<td>2</td>
<td>12</td>
<td>5</td>
<td>39</td>
</tr>
<tr>
<td>Currency</td>
<td>17</td>
<td>7</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>35</td>
</tr>
<tr>
<td>Don't like / unsure</td>
<td>16</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>35</td>
</tr>
<tr>
<td>Share experiences</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Other</td>
<td>12</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>9</td>
<td>35</td>
</tr>
<tr>
<td>Total</td>
<td>569</td>
<td>259</td>
<td>167</td>
<td>58</td>
<td>126</td>
<td>1179</td>
</tr>
</tbody>
</table>

5 CONCLUSION

This study suggests that there are still many questions to be answered about the role of UGC. The literature highlighted a number of areas that were seen as strengths and concerns in the use of UGC to support the travel purchase decisions of consumers. The results of the study suggest that the most notable of these is the level of trustworthiness and reliability that can be associated with the postings – they can be trusted because they are real experiences by real people who are independent – but can also be seen as untrustworthy because the content that is posted may be faked by someone with a vested interest. Similar paradoxes exist in relation to the range of traveler opinions available (a benefit) versus the extremity of opinions that can be posted (a concern) and whether or not the content can be specifically applied to a traveller’s own situation (where those that thought it could be applied saw it as a benefit and those that thought it could not be applied viewed it as a concern). For the most part, the results of the study tended to reiterate the benefits and concerns that were highlighted in the literature. It is envisaged that until there can be improvements made in determining the reliability of the source of UGC, less use of UGC for postings involving extreme opinions and the provision of improved filtering capabilities for travelers to identify information that is relevant to their needs that these paradoxes will remain.

The authors wish to thank and acknowledge the Sustainable Tourism Cooperative Research Centre (STCRC) and Tourism New South Wales (TNSW) as the funding sponsors of this research.

6 REFERENCES


Decision Models and the Adoption of Wireless Technology

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0167.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Technology choice, Wireless technology, Diffusion Theory, Adoption</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
DECISION MODELS AND THE USE OF WIRELESS TECHNOLOGY

Gregory Gimpel, Copenhagen Business School, Center for Applied ICT, Howitzvej 60, 2000 Frederiksberg, Denmark. gg.caict@cbs.dk.

Abstract

Many nations boast high broadband internet penetration. In many of markets, consumers can choose among competing technologies to connect to the internet. Much research in the IT diffusion domain has been useful for examining the adoption of individual technologies. This paper explores technology adoption in an environment in which different technologies fill a similar need by focusing on wireless internet access. This paper uses a series of focus group interviews to analyze the use of cognitive referencing in the form of reference prices, situational (status quo) framing, and mental accounting as potential determinants of technology adoption. The findings of this study suggest that adoption decisions are reference-dependent and that researchers should consider users’ referents when studying adoption in the wireless broadband market.

Keywords: Pricing, technology choice, diffusion theory, adoption, standards, wireless, broadband, qualitative research
DECISION MODELS AND THE USE OF WIRELESS TECHNOLOGY

1.0 Introduction

In Europe, the USA, and parts of Asia, internet penetration is high, with substantial consumer use of broadband technology. A plethora of companies offers internet access through a wide variety of different technologies. Wireless data technologies for laptop computers, such as 3G and WiMAX, are new market entrants that offer a fundamentally different delivery method than the more widely adopted fixed-line access points. This paper examines influences on the consumer decision process that affect the adoption of information and communication technology (ICT) through the study of wireless internet for the computer.

Researchers often employ traditional theories of technological adoption when investigating the adoption of ICT. While these methods have proven useful when studying individual technologies, exploring different models can also enhance the understanding of new technology usage. The current broadband market offers existing technologies that compete with new technologies, all of which access the same internet. The users in this study did not evaluate choices in isolation, but weighed alternatives against their current internet technology.

This paper explores the adoption of technology in a situation in which different technologies fill a similar consumer need. More specifically, it explores the determining factors for the adoption of one internet service technology over another; the choice of fixed-line versus wireless technology in particular. Denmark, the world leader in broadband penetration according to an Economist Intelligence Unit study (2007), serves as a market exemplar for Western nations, and was consequently chosen as the location for this study.

In a market offering a plethora of technological choice, why does such a large portion of the population use such a small number of technologies? What goes on in people’s minds during the process to decide which type of information technology to buy? Wireless broadband is a relatively new offering. This paper investigates diffusion models in an environment of competing options by asking the question: what is the best model for predicting and understanding consumer adoption of high-speed wireless internet?

This paper uses focus group research to fill apparent gaps, as pointed out by the calls for alternative perspectives on technology adoption (Benbasat & Barki 2007; McMaster & Wastell 2005; Venkatesh & Davis & Morris 2007) and by the call for the exploration of the consumer decision-making process beyond construct-based research (Blechar & Constantiou & Damsgaard 2006). Following the example of Eynon (2005), this paper utilizes focus group methodology to explore user attitudes about the adoption of internet ICT. The analysis builds upon bounded rationality, prospect theory, and mental accounting. It also expands on the more recent work of Blechar et al. (2006) who applied reference pricing to wireless ICT.

The contributions of this study are two-fold. First, it adds to user behavior research in wireless services by elaborating on the process by which end users make communication technology purchase decisions. Second, it illustrates the insights gained by applying cognitive decision making models to a situation of technological choice.

This article is structured as follows: the next section provides an overview of major broadband access technology. The following section offers an overview to reference pricing and the cognitive psychology theories on which referential decision making is built, as well as the role pricing plays in ICT diffusion. The proceeding segment details the research methodology used to conduct the focus groups and analyze the data. The subsequent section presents the findings. After the presentation of data, the results segment compares research outcomes to theoretical arguments, addresses the implications for service providers, and discusses the propositions and research instrument.

Proceedings ECIS 2009
2.0 BROADBAND OVERVIEW

Many technologies enable connection to the internet at high speeds. While they all provide internet access, the various technologies deliver the service in distinctly different ways and offer different performance traits. Fixed line technology requires a user to access the internet from a specific location. Portable broadband frees users to move their computers within a limited geographic area. Mobile internet enables free movement between access points without interruption.

<table>
<thead>
<tr>
<th>TECHNOLOGY</th>
<th>CAPACITY</th>
<th>TRANSMISSION</th>
<th>PORTABILITY</th>
<th>MAXIMUM RANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>xDSL</td>
<td>1.5 Mbps (12 Mbps)</td>
<td>copper telephone lines</td>
<td>fixed location</td>
<td>5.4 km (0.3 km)</td>
</tr>
<tr>
<td>Cable</td>
<td>40 Mbps</td>
<td>coaxial cable TV lines</td>
<td>fixed location</td>
<td>1-3 km</td>
</tr>
<tr>
<td>Fiber</td>
<td>1 Gbps</td>
<td>fiber optic cable</td>
<td>fixed location</td>
<td>20 km</td>
</tr>
<tr>
<td>Powerline</td>
<td>200 Mbps</td>
<td>existing A/C power lines</td>
<td>fixed location</td>
<td>1-3 km</td>
</tr>
<tr>
<td>Satellite</td>
<td>155 Mbps</td>
<td>extraterrestrial satellite</td>
<td>fixed location / wireless</td>
<td>1000-36,000 km</td>
</tr>
<tr>
<td>Wi-Fi</td>
<td>11.3 Mbps</td>
<td>unlicensed radio band</td>
<td>portable / wireless</td>
<td>100 m</td>
</tr>
<tr>
<td>WiMax</td>
<td>2.8 Mbps (practical speed)</td>
<td>3.5 GHz radio frequency</td>
<td>portable / wireless</td>
<td>50 km</td>
</tr>
<tr>
<td>3G</td>
<td>2.0 Mbps</td>
<td>mobile telephone tower</td>
<td>mobile / wireless</td>
<td>Large: based on cellular network</td>
</tr>
</tbody>
</table>

Table 2.1 Summary of common technologies and their distinguishing features (Corning 2005)

This research makes the explicit distinction between using a Wi-Fi router to relay a fixed line connection for a few meters and using a subscription Wi-Fi hotspot service. This study considers a home Wi-Fi router as fixed-line technology because users must decide which access technology to connect to the router when making an ICT adoption decision. See Figure 2.1 for an illustration of the distinction made for Wi-Fi in this study.

3.0 THEORETICAL INSIGHTS AND PROPOSITIONS

This section begins with the call for different theoretical approaches to the study of technology adoption within Information Systems research. It provides an overview to the theoretical basis for the argument that reference prices influence the decision of whether or not to use a technology. Second, it provides an introduction to the behavioral economics literature that paves the path for the proposition that a status quo bias impacts the willingness of a person to accept a new technology. Third, it lays the theoretical grounding for user preference of simple, flat-rate payment plans that can affect the selection of one internet service over another.
Much of the recent IS adoption literature uses theories such as the Technology Acceptance Model (Davis 1989), Theory of Planned Behavior (Ajzen 1991), Unified Theory of Acceptance and Use of Technology (Venkatesh & Morris & Davis & Davis 2003) and other theories that focus on individual-level adoption. For example, researchers recently applied the Theory of Planned Behavior by expanding it to new types of IT systems (Dinev & Qing 2007) and to the acceptance of broadband among different groups (Hsieh & Rai & Keil 2008). Other studies use UTAUT to study computer applications in non-Western countries (Al-Gahtani & Hubona & Wang 2007) or to investigate the adoption of e-commerce in developing nations (Uzoka 2008).

Research continues to advance the Technology Acceptance Model, the most often employed and influential IS theory (Benbasat et al. 2007; Lee & Kozar & Larsen 2003; McMaster et al. 2005). Recent applications include wireless data networks (Yoon & Kim 2007) and mobile communications (Zhang & Mao 2008). Despite greatly advancing IS research by focusing a formerly scattered field of study (Lee et al. 2003), some argue “TAM has fulfilled its original purpose and that it is time researchers moved outside its confines” (Benbasat et al. 2007). While TAM is appropriate in an organizational context, it lacks other vital considerations necessary to explain user behavior (Lopez-Nicolas & Molina-Castillo & Bouwman 2008). Perceived usefulness and ease of use may not be sufficient criteria to study emerging services such as wireless because the impact on everyday routine may be more influential than the technology itself (Bouwman & Carlsson & Molina-Castillo & Walden 2007). TAM is binary, considering only the choice whether to adopt or not adopt one technology in isolation. It ignores context and contingency. Therefore, it does not adequately consider the complexity of the actors’ decision (Bouwman et al. 2007; McMaster et al. 2005).

This paper breaks from the study of technology in an isolated environment to address the gaps pointed out by Benbaset and Barki (2007), McMaster and Watsell (2005), and Venkatesh et al. (2007) in their calls for alternative theoretical perspectives that expand the study of technology diffusion to include contingent models. It also answers Blechar et al.’s (2006) call to explore adoption decisions beyond construct-based research.

Because people cannot gather or analyze every bit of information, they simplify and seek satisfactory, not necessarily ideal, solutions. Building on this concept of Bounded Rationality (Simon 1955), Prospect Theory, a behavioral economics approach to decision-making in the presence of choices and risks, argues that outcomes are contingent upon comparisons to a referent. People judge value by gains or losses relative to a reference point, not in terms of absolute monetary value (Kahneman & Tversky 1979). For example, people perceive a $5 price change on a $25 item as greater than a $5 change on a $500 item, even though the differences have the same financial value (Thaler 1980). Constructive preference, in which people determine their preferences when presented with a choice rather than drawing upon a priori preferences, suggests that the context and the framing of the choice affect the decision. A goal of the decision-making process is to maximize the ease of justifying a decision (Bettman & Luce & Payne 1998).

Constantiou, Damsgaard, and Knutsen (2007) argue that price is the most important attribute for basic and advanced users of advanced mobile features. Similarly, a study of Wireless Application Protocol details how users compared it against PC-based internet services (Hung 2003). Blechar et al. (2006) introduce reference pricing to the use of advanced mobile services by arguing that mobile phone users compare phone-based data services to computer-based internet services. The reference situation of the internet, perceived as nearly free of charge for similar services, made nearly any price for m-services to expensive in the consumer mind. Therefore, the redundancies between mobile services and PC-based internet could constrain mobile services adoption if users perceive the personal computer as higher in quality relative to mobile devices (Blechar et al. 2006). The importance of reference pricing in a situation of constructive preferences leads to the following proposition:

Proposition 1: the reference price of fixed-line internet creates a situation that biases the consumer against adopting wireless internet service for the laptop.
Because preference construction is contingent on the framing of the problem, the method of elicitation, and the context of the choice; contrast effects play a strong role in decision making by influencing the reference point that is used to gage value (Tversky & Simonson 1993). The perception of quality differences between a new option and the referent affect the likelihood of switching (Constantiou et al. 2006). Since losses loom larger than corresponding gains, loss aversion suggests that disadvantages are more salient than advantages (Tversky et al. 1993). In an identical situation, a different decision can be reached depending on whether the choice is framed to indicate a gain or a loss (Kahneman et al. 2003). Perception is reference-dependent and people notice and evaluate changes as percentages rather than the whole value, creating a phenomenon in which people assign a higher value to something that they view as theirs to be lost; while they assign a lower value to something they see as a gain. Take for example the person who is unwilling to pay more than $35 for a bottle of wine yet is unwilling to sell a bottle he already owns for $100. The loss aversion associated with this “endowment effect” contributes to status quo bias (Kahneman & Knetsch 1991). Loss aversion during the discernment process leads to the second proposition of this paper:

**Proposition 2:** status-quo bias adds a perceived switching cost that increases the difference between the referent and the alternative, diminishing the perceived gain from the adoption of a new broadband technology.

Consumers use “mental accounting” to frame purchasing decisions by assigning them to an expense category. They use these mental distinctions as a frame of reference to measure changes to reference prices. Decisions about which category, and whether to combine categories, impacts the perceived value of the choice. Loss aversion strongly influences how people organize financial transactions in their mind. Accordingly, “consumers don’t like the experience of ‘having the meter running.’ This contributes to what has been called the ‘flat rate bias’ in telecommunications. Most telephone customers elect a flat rate service even though paying by the call would cost them less” (Thaler 1999).

Research into the implications of the bundling wars that have begun with the telecom business explores the effect complementarity has on perceived value and how customers assess a bundle’s transaction value. Mental accounting plays a key role in assessing the transaction value and how context and perceived price affect the choice of mental “category” consumers use to budget a bundled offering (Sheng & Parker & Nakamoto 2007). The influence of mental accounting in purchase and usage decisions leads to the third proposition set in this paper:

**Proposition 3:** consumers prefer flat-rate plans over pay-per-use billing as a way to simplify mental accounting

**Figure 3.1  Theoretical model of decisions in the presence of choice and derivative research propositions**

4.0 METHODOLOGY

This section describes the research methodology used in this study. It begins by detailing the selection of the research instrument and details the study design, recruitment of participants, data collection methods, and data analysis technique.

4.1 Instrument Selection

This study explores how consumers make the decision whether or not to adopt wireless broadband for their laptops. It asks the research question: “what do people think about when deciding to buy
broadband internet?” The attitudinal-based nature of the research question and propositions makes the focus group interview the logical choice of research formats (Kitzinger 1995; Krueger & Casey 2000; Morgan 1997).

The processes of attitude formation and decision-making are inherently unobservable. When researching such topics, focus groups provide access to data that cannot be easily obtained by direct participant observation or open-ended interviews (Morgan 1997). Krueger and Casey emphasize the suitability of focus groups to “uncover factors that influence opinions, behavior, or motivation” (Krueger et al. 2000).

The group interview research technique can be a self-contained methodology that can generate principal data that can be the basis for a complete study (Morgan 1997). The data gained from a particular study provide theoretical insights which possess a sufficient degree of generality or universality to allow their projection to other contexts or situations which are comparable to that of the original study (Sim 1998).

4.2 Study Design

This study consists of multiple focus groups in order to be able to analyze data across groups to find patterns and themes (Krueger et al. 2000) and to reduce the effect of individual group dynamics on the overall dataset (Morgan 1997). The formation of three separate groups serves to achieve the data saturation recommended by Morgan (1997) and by Krueger and Casey (2000).

The number of participants represents a strategic balancing act between manageability and achieving significant idea diversity among group members (Morgan 1997). This study opted for small groups, with 6, 5, and 4 participants respectively. In order to compare and contrast data across groups, the interview questions were created prior to the focus groups (Krueger et al. 2000). The questions progressively narrow the topic of discussion specifically to gather data relevant to validating the propositions.

4.3 Participant Recruitment

Recruitment strategy and the study design encouraged openness of sharing and to facilitate interaction among participants. As recommended by Barbour, the goal was to achieve a balance between diversity and homogeneity (Barbour 2005). In order to ensure sufficient heterogeneity of ideas within each group, the selection criterion was “college-aged persons living or working in Denmark.” This basis is general enough to include a large population, yet create a sense of similarity among the group interviewees, as recommended by the literature (Kitzinger 1995; Krueger et al. 2000; Morgan 1997).

Within the mixed gender participant set, some were students, others were students with jobs, and some had completed their education and were in the Danish workforce. Thus the screening criteria also had the advantage of diversity, which according to Kitzinger (1995) maximizes the exploration of different perspectives within a group setting.

The study selected young adult participants for several reasons. The study aimed to reduce extraneous variables and prior studies show that age affects individual technology adoption (McFarland 2001; Morris & Venkatesh 2000; Yang & Jolly 2008). The selected group has lived their entire lives since the introduction of mainstream personal computing and related technologies (Prensky 2001). They make heavy use of ICT; particularly internet use, and they have therefore incorporated it into their daily lives. They consider technology to be part of the landscape (Oblinger 2003) and they consider computers as commonplace, not as technology (Frand 2000; McMahon & Pospisil 2005). Networked for most their lives (Prensky 2001), they are generally unaware of the pre-internet era (Rickard & Oblinger 2003). Equally important, they are consumers whose entrance as decision makers into the marketplace closely coincides with the wide-scale launch of wireless internet services.
4.4 Data Collection

The focus groups were convened during late April and early May 2008. After the meetings, field notes were typed recording salient points and researcher impressions of the discussions. The conversations were recorded in order to maintain accuracy during the analysis phase of the project. In keeping with the effort to make participants feel comfortable about sharing their opinions, the research plan followed the privacy guidelines recommended by Krueger and Casey (2000).

4.5 Analytical tools

The analysis used both the field notes and transcripts. The data were analyzed at the group level and on a comprehensive (study-wide) level. ATLAS.ti was used to code the data. At times many statements contained similar information; and other statements were short statements that required the context of the surrounding conversation to convey a point. The quotations included in the next section of this paper serve as demonstrative exemplars.

5.0 FINDINGS

This section details the information provided by the focus group participants. It begins with an overview and then elaborates by providing specific statements and quotations.

Based on past and present experiences, the focus group participants weigh many factors when evaluating broadband internet service. Service characteristics such as speed, security, stability, provider reputation, data transmission consistency, and convenience all play a role in the decision-making process. For all participants, price plays a pivotal role in deciding which broadband service to adopt. Rather than purely considering the monetary amount charged, they view price as the financial value derived from the amount paid for the service relative to the performance received. The overall pricing function includes the evaluation of different service characteristics such as performance metrics and provider reputation. Thus, they compare the price of one option to another, both in terms of the monetary amount and the price to performance ratio. While some consumers seek the ultimate price-performance ratio, there is a propensity for the monetary amount to supersede other attributes once the consumers’ minimum needs have been satisfied. Therefore, generally speaking, the performance portion of the function begins to weigh less heavily once a minimum threshold is satisfied.

All of the participants subscribe to fixed-line internet technology, and all use Wi-Fi access at least part of the time. The use of Wi-Fi to deliver the internet from a fixed line access point creates the reference perspective that internet access is portable. While such internet access cannot be accessed continuously over a large geographic area, it does enable access in a wide variety of locations. Some participants expressed an interest in mobile internet that they could use on moving trains, etc. Others saw little difference between mobile service and the portable service they currently experienced with a Wi-Fi link to a fixed line connection. Some subscribe to Wi-Fi hotspot services as either a paid subscriber or as part of an internet collective, but all participants primarily use Wi-Fi as router connected to the same fixed line access that the users can plug into if they choose. Therefore fixed line access serves as the reference point against which wireless alternatives are compared.

Although all participants used fixed-line for primary access, the cost basis varied widely. Some purchased internet as part of a service bundle along with VoIP or television service. Others had home internet service paid by an employer so that they experienced no out-of-pocket expense for the service. Others received internet from the building in which they lived. Often the internet was included as part of the rent, essentially creating a reference price of zero. A participant explains how this frames the decision process:

“And that might be a competition to wireless because people will have internet with their rent, as it is right now. So they don’t need that wireless or that extra internet. It’s not like
they say ‘either we’re going to have wired that we pay for or we’ve going to have wireless.’
It’s ‘we have a wired are we going to pay for wireless as well?’”

Others lived in buildings in which their choice of fixed-line provider was determined by the building, and therefore had only wireless technologies as alternatives. In a similar instance, a participant must choose between two fixed-line options available at his building. A participant whose rent payment includes internet service also belonged to the Webbies collective. Webbies members receive a wireless router that permits any member to connect to any other member. The participant states: “And this way, I also have almost internet everywhere with my cable. …It’s free and a lot of people have it.”

When asked about willingness to switch providers in order to adopt a new technology, some participants express a willingness to switch providers. Others express a status quo bias in which they are hesitant to change providers when such change was necessary to adopt a better internet technology. If participants associate inconvenience with switching, they would stay with their provider. Participants explain that this hesitation stems from a broad aversion to changing relationships with companies.

Others explain their hesitation to switch companies because of the time it takes to complete the switching process. One participant shares that uncertainty plays a part in creating his status-quo bias:

“It’s changing the paperwork and the bank details and whatever it is. But it’s also the fact that you know what you have. You know that it works or that it doesn’t work or that whatever else you don’t know about the other company.”

All of the participants’ primary internet access services employ flat-rate billing plans. The participants expressed a general preference for flat rate pricing plans and an aversion to per-use billing structures. Those who have internet-enabled phones use the data service infrequently because of the high cost of sending and receiving data. One participant uses three services with his laptop: fixed-line, Wi-Fi hotspot, and 3G. The fixed line and hotspot subscriptions are flat rate and used frequently; however, the expense of 3G causes him to use it “only as little as possible, for essential things.” Participants explained that choosing fixed rate plans simplifies their personal budgeting process and also makes it easier to decide whether or not to use the internet:

“…if there is a new opportunity to get like TV access or somebody tells you to go watch this or do this, I don’t have to worry about whether it is worth it to do that transaction; whether I want to spend the extra money on the download. I like the that flexibility of being able to do whatever I want”

Prior to broadband, internet service price varied upon use. When asked about the advantages of dial-up verses high speed fixed-line internet service, in addition to performance differences, focus group participants brought up the differences in billing strategies between the two technologies with statements like:

“When you used the modem, you know in the back of your head, that it costs for every minute. Now it doesn’t matter, really.”

### 6.0 RESULTS

This section is divided into two parts. The analysis compares the findings to the expected outcomes based on theoretical arguments and suggests a diffusion approach for wireless internet service providers based on the study findings and existing literature. A discussion of the propositions and evaluation of the focus group methodology concludes the segment.

#### 6.1 Analysis

The participating consumers view internet service options as access technologies. They have already integrated the internet with their daily lives. They are accustomed to fixed line and wireless ICT.
Therefore, people view the choice among alternative types of technology as a relatively routine purchasing process and thus employ standard cognitive choice models.

When buying broadband service, people exhibit isolation effects. In the case of the focus group data, all broadband technologies take the user to the internet, enable VoIP, data transfer, etc. As a result, the focus group participants pay attention to the factors that differentiate the choices, the most concrete of which is price. Therefore, their ISP choices are generally a function of price. The disregard for components shared by the alternatives conforms to the findings of Kahneman and Tversky (1979).

The interviewees base value on price, not just absolute prices of one offering versus another, but as an assessment of the price relative to various key features. Therefore, if one service costs twice as much, but offers performance perceived to be twice as “good” then the offers would be equal in the consumer mind. At that point, the consumer would evaluate the offering that came closest to his or her subjective performance requirements, in a rough but somewhat rational manner. As a result, the participants evaluate the offerings using their past and current experience to frame their decision, as postulated by Kahneman and Tversky (1979). They make their decisions by evaluating differences from their fixed line reference point, supporting Kahneman’s (2003) findings that people use analogical reasoning to make choices.

Focus group participants using home broadband access reimbursed by an employer or included in the rent or tenant fees perceive the reference price of broadband services as free. These consumers experience no marginal cost for adopting their current internet service, but would experience an incremental cost if they decided to use a different internet service. Therefore, any alternative choice, including wireless broadband, would be infinitely more expensive; which provides support for the argument by Blechar et al. (2006) that mobile service users make reference to existing service delivery platforms when making their usage decisions.

Participants use price analysis to view possible outcomes as a gain or loss to a reference point so they can employ a partially rational decision process. This approach to balancing accuracy and effort conforms to the findings of Bettman et al. (1998) who write that people reach a subconscious compromise between the desire to make the right decision and the desire to minimize the cognitive effort required to make the decision.

A notable situation exists when comparing current fixed line and wireless broadband service. Except for mobility and installation convenience, every metric considered by focus group participants is inferior to the fixed line frame of reference. As predicted (Kahneman 2003; Kahneman et al. 2003), the loss of functionality is weighed more strongly than the gain of another function. This also supports Constantiou et al.’s (2006) argument that quality perception may cause a status quo bias against mobile services.

The presence of an affinity for flat-rate billing and supports Thaler’s (1999) argument that “mental accounting matters.” Participants feared that pay-per-use plans could become very expensive, as postulated by Thaler’s assertion that loss aversion contributes to a flat rate bias in telecommunications decisions (Thaler 1999). Furthermore, the adoption of a flat-rate plan frees the internet user from evaluating whether or not each internet activity is worth incurring a separate expense, in accordance with Bettman et al.’s argument that one goal of the decision making process is to maximize the ease of justifying the decision (Bettman et al. 1998). In the case of broadband internet, an unlimited use plan eliminates the possibility of “buyer's remorse” after spending money to visit a website, download a file, or view a video.

6.2 Discussion

This discussion provides an overview of the findings and a summary of the contribution to theoretical literature. Second, it addresses the relevance to the business community. Finally, it reviews the methodology chosen for the study.
The data support the propositions derived from existing literature. In general, as put forth Proposition 1, participants use reference pricing as a primary decision-making tool. Although opinions vary, a general status quo bias influences purchase decisions, supporting Proposition 2. Mental accounting contributes to a flat rate payment bias in telecommunication services, as predicted by Proposition 3.

The findings contribute to the body of literature that studies technology adoption beyond individual-level technology studied in isolation. The study adds to the body of knowledge about consumer decision-making processes outside construct-based research. It adds to user behavior research in wireless services by elaborating on the process by which end users make ICT purchase decisions. It illustrates that insights can be gained when cognitive decision models are applied to technology choice.

In addition to theoretical contributions, the findings resulting from this focus group study have several business implications for companies with heavy investment in wireless internet technology. With heavy sunk-cost investments in licensing and infrastructure roll-out, much is to be lost if wireless broadband does not gain mass adoption. There are many challenges to the adoption and diffusion of wireless broadband: inferior perceived performance, high price, and reference situations that weaken the wireless value proposition. These reference situations concern the demand for mobility framed by Wi-Fi and by current usage habits.

Study data suggest wireless providers should use flat rate pricing and create an offering that disassociates mobile broadband from fixed-line internet so that users (i) assign the cost to a different mental account and (ii) view it as different enough to weaken a performance comparison between the technologies. Once customers begin using wireless internet, they may experience the “endowment effect” (Kahneman et al. 1991) which will cause the users to perceive giving up wireless internet as a loss. In such case, status quo bias will make them less likely to discontinue service. As Kahneman (2003) indicates, people opt out of the framed situation much less frequently than they opt in to an alternative.

The study creates an opportunity to review the selected methodology. In this case, the focus group format provides a method to learn the thought process of the test subjects. The interplay of the different participants during discussions provides rich data for analysis. The open-ended nature of the format allows participants to introduce new ideas into the group and therefore provides insight not attainable through surveys. During the first focus group, for example, participants introduced a topic that received significant attention and generated much discussion. This topic appears as a question within the interview guide used with subsequent groups. Additionally, a theme recurred within the first focus group that seemed a worthy topic of research. Therefore, a new question explicitly asked subsequent focus groups about their perceptions regarding the theme, enabling more elaborate data collection.

The small sample size of this study precludes developing a probabilistic model from the data; however, the focus group findings are useful for theoretical generalizability (Barbour 2005) and conceptual transferability (Krueger et al. 2000). The selection of participants with common characteristics rather than random sampling provides insights, but it limits the transferability of the findings to the general population at large.

7.0 CONCLUSION

This study uses focus group data to explore the differences as perceived by high-speed internet users that would influence the adoption of wireless broadband. By exploring what consumers consider when making technology purchase decisions, the data support the importance referencing pricing plays in technology adoption. It lends additional support to the existence of status quo biases in purchase decisions; and it strengthens the argument that consumers prefer flat-price service plans in telecommunications because of bounded rationality and the derivative concepts of prospect theory and mental accounting.
A significant implication of this paper is furthering the proposition that much consumer adoption of new technology can be explained with general decision models by providing strong evidence that these models play a primary role in adoption decisions. The implications are important both to further research as well as to industry. Focusing exclusively on technology diffusion models when conducting research or developing business plans or marketing strategy may ignore real and present influences that affect the acceptance and uptake of innovation.

While this study provides support for and insight into applying cognitive decision models to technological adoption, the study has limitations that warrant further research. For example, additional qualitative studies could enrich the findings of this study by exploring the underlying reasons why some consumers exhibit stronger status quo bias than others. Additionally, quantitative research with larger samples could validate the findings of this study. One potential study would focus on the young demographic featured herein because their technology literacy differs from the population at large. Continuing to study this demographic is relevant to adoption theory and business practitioners because the participants may view different technologies as their referents; and thus their status quo may be different from that of older individuals. At the same time, a sample of the population at large would serve to test the overall validity of the findings. Similar results between the groups would validate this study. If differences between younger and older demographics prove significant, then it might support a call for different approaches to the study of technology adoption based on generational considerations.

Acknowledgements

This work was in part supported by the DREAMS project via a grant from the Danish Agency of Science and Technology (grant number 2106-04-0007). I would like to express my gratitude for the assistance I received from the Center for Applied ICT at Copenhagen Business School.

References


## ACCEPTANCE PROBLEMS OF AMBIENT INTELLIGENCE AND MOBILE TECHNOLOGIES IN HOSPITALS IN INDIA AND GERMANY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0372.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-health, Ubiquitous systems, Partial Least Squares, Technology Acceptance Model (TAM)</td>
</tr>
</tbody>
</table>
ACCEPTANCE PROBLEMS OF AMBIENT INTELLIGENCE AND MOBILE TECHNOLOGIES IN HOSPITALS IN INDIA AND GERMANY

Kummer, Tyge-F., ESCP-EAP European School of Management, Heubnerweg 6, 14059 Berlin, Germany, tyge.kummer@escp-eap.de

Bick, Markus, ESCP-EAP European School of Management, Heubnerweg 6, 14059 Berlin, Germany, markus.bick@escp-eap.de

Gururajan, Raj, University of Southern Queensland, West Street, 4350 Toowoomba, Australia, gururaja@usq.edu.au

Abstract

Ambient intelligence systems facilitate job performance by medical staff in health care services. Several papers detail scenarios in which these technologies may support clinicians in their daily work processes, but their specific characteristics suggest such technologies also could be useful for surveillance and subsequent control of employees. Until now, only little attention has focused on resolving such issues. On the basis of 16 in-depth interviews with medical staff from three German hospitals, this study identifies a reserved attitude and several acceptance problems among interviewees. The qualitative data indicate 10 hypotheses, tested using a questionnaire study of 215 nurses in training from Germany and India. The investigation of the quantitative data relies on partial least squares modeling to identify and categorize problems with user acceptance of ambient intelligence and mobile systems on various levels. The results of this mixed methods study mainly indicate that German participants assess ambient intelligence much more critically. Consequently, specific strategies for implementing such technologies should be adopted.

Keywords: Acceptance Problems, Ambient Intelligence, Health Care, Mobile Technologies, Partial Least Squares (PLS) Modeling, User Acceptance
1 INTRODUCTION

Ambient intelligence and mobile technologies offer various possibilities for enhancing the efficiency and effectiveness of medical treatment in hospitals. Through the use of sensors for example, such systems provide context-aware support. In the health care domain, ambient systems also might support an operating team by determining whether any instruments have been left in the patient (Macario & Morris & Morris 2006). Such avoidable errors lead to approximately 17,000 deaths per year in Germany, and technologies can help reduce the risk of complications for patients.

Yet ambient intelligence also is bound to provoke public scepticism. For example, the use of radio frequency identification (RFID) tags has invoked concerns about potential invasions of customer privacy, which could result in boycotts of the devices (Spiekermann 2008). Acceptance problems also should be anticipated because an ambient system offers new possibilities for monitoring and controlling the staff in a work environment, which may create anxiety and privacy concerns. These issues also apply to medical staff in health care services. Especially in Germany, people seem quite sensitive to the potential threats of such technologies (Bick & Kummer & Rössig 2008).

Taking the specific characteristics and acceptance problems associated with ambient intelligence into account means questioning the usefulness of common acceptance models (Spiekermann 2008). New technologies imply additional risks for users and thus require adapted acceptance models, especially in the health care domain, where the traditional Technology Acceptance Model (TAM) is inadequate (e.g., Lapointe & Lamothe & Fortin 2006).

Few studies have investigated physicians’ acceptance of modern information and communication technologies in general or mobile technologies in particular (Raitoharju 2007). To bridge this gap, we investigate acceptance problems specific to ambient intelligence and mobile technologies by nurses in Germany and India, who are potential users. Germany provides a viable setting because of its negative experiences with adopting ambient technologies (Bick & Kummer & Rössig 2008); India is an emerging country that is closely connected to information technology and is culturally different from Germany. The Global Leadership and Organizational Behavior Effectiveness (GLOBE) study (House & Javidan 2004) supports this assumption. Our study considers any differences in the adoption or application of these new technologies across the two countries. To clarify acceptance problems, we follow a mixed methods approach in an effort to overcome the shortcoming of traditional acceptance models in health care settings. Specifically, we attempt to analyse hospital workers’ fears regarding their acceptance of ambient intelligence and identify any differences due to national culture.

We begin by outlining in detail the scientific background and identified research gap. We then present our research design, followed by the results of our investigation. In the conclusion, we outline some opportunities for further research.

2 STATE OF KNOWLEDGE

In this section, we define the term ambient intelligence, explain its application to the health care domain, and offer an outline of acceptance research in hospitals.

2.1 Ambient intelligence in hospitals

The term “ambient intelligence” was introduced by the Information Society Technologies Program Advisory Group (ISTAG) of the European Union to refer to environments equipped with advanced technologies and computing, which create an ergonomic space for occupants (Bohn et al. 2005, ISTAG 2003, Regmaglino et al. 2005). Weber, Rabaey, and Aarts (2005) similarly define ambient intelligence as the “vision that technology will become invisible, embedded in our natural surroundings, present whenever we need it, enabled by simple and effortless interactions, attuned to all our senses, adaptive to users and context-sensitive, and autonomous.”
As a complex service, the level of support demanded in the health care industry differs from that in other sectors, and research in this area must take specific working conditions into account. Medical staff occupy various positions, including surgeons, physicians, radiologists, and surgical and ward nurses. They often share available desks, computers, and devices and are extremely mobile. A nurse, for instance, usually walks more than a kilometre per shift (Morán et al. 2006), and they often work collaboratively in teams, rarely on individual tasks. Because their work focuses on the physical domain of the patient, digital content should provide back-up support in decision processes. Nurses also move constantly among locations and are frequently interrupted, so mobility and flexibility are necessary and fundamental elements. Ad hoc problem solving also has critical importance for the entire staff (Bardram & Baldus & Favela 2006).

Ambient intelligence might support medical services in several areas. In combination with mobile technologies, ambient intelligence can improve patient identification processes, such as by employing RFID tags that contain relevant information about the patient and thereby prevent mistreatments due to errors during treatment (Andersen & Bardram 2007).

2.2 Acceptance of mobile and ambient technologies in hospitals

Research in the field of user acceptance attempts to explain how and why users adopt new technologies. Several streams of research can be distinguished. In the domain of health care, most studies use the Technology Acceptance Model (TAM) by Davis (1989), the Theory of Planned Behaviour (TPB) by Ajzen (1991), or combinations of the TAM, TPB, and Innovation Diffusion Theory (IDT) by Agarwal and Prasad (1998). Extensions to these approaches include TAM2 (Venkatesh & Davis 2000), and combinations of various approaches have resulted in the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al. 2003). Although these various approaches offer some degree of generalisability, they generally have been developed for contexts such as desktop computers or standard applications. No studies investigate the acceptance of ambient intelligence technologies in health care, and few studies analyse the acceptance of mobile technologies in this domain (Raitoharju 2007); those that do focus on physicians rather than nurses. Only Wu, Wanga, and Lind (2007) investigate the acceptance of mobile health care systems based on an integrated model of TAM and IDT in Taiwan.

The Technology Acceptance Model establishes perceived ease of use and perceived usefulness as determinants for predicting acceptance of a technology in a given setting. In TAM2, cognitive and social influence factors also influence perceived usefulness (Venkatesh & Davis 2000). Ong, Lai, and Wang (2004) extend the TAM with perceived credibility, or the degree to which a person believes that the use of a particular system will be free of privacy and security threats. These constructs are relevant in many information systems studies, but in health care environments, the influence of perceived ease of use on other constructs is not significant (Chismar & Wiley-Patton 2003, Spil & Schuring 2006). Furthermore, perceived ease of use does not offer a significant predictor of the intention to use a technology in a clinical domain (Chau & Hu 2002, Hu et al. 1999, Jayasuriya 1998). In a study of the dynamics of IT adoption in a major change process in health delivery, the Technology Acceptance Model emerges as inadequate (Lapointe & Lamothe & Fortin 2006).

Thus, it appears that ambient intelligence features specific characteristics that enable new functions and are connected with new acceptance problems (Spiekermann 2008). For example, the possibility of surveillance, which creates suspicions of invasions of users’ privacy, leads to acceptance problems, none of which appear in common acceptance models. These acceptance difficulties have not been investigated, so established acceptance models cannot analyse ambient intelligence acceptance in health care contexts. Against this background, we endeavour to investigate which acceptance problems are most relevant for ambient intelligence and mobile technologies in hospital settings. We first investigate which acceptance problems exist in this domain and whether they influence intentions to use the technology.
3 DESIGN AND METHODOLOGY

This study uses a mixed methods design. First, we conducted in-depth interviews to identify various acceptance problems. Second, on the basis of these interviews, we developed a questionnaire that we distributed to nurses in training from Germany and India. The data collection then involved three specific stages: In the first, exploratory stage, we reviewed existing literature to identify various issues that might influence health care domains in which handheld devices could be used. The main purpose of this stage was to identify important factors and create a suitable interview instrument. The second stage involved actual data collection through interviews. Finally, in the third stage, we administered the questionnaire that we developed through the previous stages.

3.1 Explorative research design

We chose a qualitative research design to identify acceptance problems with the use of ambient intelligence technologies in hospitals. Semi-structured interviews, which are flexible and open by nature, provided a suitable means for contrasting areas of application, as described in prior literature. The study took place in the surgery departments of three German teaching hospitals. To provide a high degree of contrast, we chose hospitals that differ in size and organisation, though the number of beds in the investigated clinics is similar at approximately 30 beds each. Participants were physicians and nurses in different leadership positions. This selection criterion reflects the assumption that people in higher positions have better knowledge about relevant tasks and can better evaluate the involved staff. In addition, the interviews in every hospital involved persons in special functional capacities, so that we could understand the individual organisational circumstances. The interviews took place in the work environment, conducted by two researchers with experience in qualitative research. They lasted between 40 and 50 minutes on average and were digitally recorded, with the participants’ permission, and later fully transcribed. The transcripts also were coded separately by the two researchers. The full results of this research have been published previously (Bick & Kummer & Rössig 2008).

3.2 Quantitative research design

In this section, we detail the research components and hypotheses, research scope, and methods for analysing the data.

3.2.1 Scenario, constructs, and hypotheses

A developed scenario describes acceptance problems and benefits in three major areas. Because the scenario needed to be easy to understand, we used no special terms and confirmed its comprehensibility with a pretest. The scenario indicates:

A system automatically controls the staff of a hospital in the medication process. The patient and medicine are tagged by sensors. The medical staff uses a mobile device that provides information about the required steps for treatment and automatically documents the work. During the medication process, sensors on the medicine and patient automatically check their mutual suitability. When errors in the treatment are recognised by the system, an alarm occurs, and the error is recorded.

The benefits and potential risks (including the possibility of surveillance) of the system were emphasised. The questionnaire features three major fears, identified in the explorative study. In contrast with perceived credibility (Ong & Lai & Wang 2004), which has a mostly positive connotation, fears represent negative influence factors and have concrete negative associations with technology conditions, the work, and surveillance. The corresponding question items were developed on basis of the interviews. The questionnaire also includes key constructs from TAM, perceived usability and intention to use the system. The personal opinion of the potential user about a fictitious scenario thus represents acceptance, for which the original question items from TAM are applicable. Relevance and voluntariness also emerge as important for acceptance, according to the interviews, and
appeared in TAM2 as potential influences on perceived usefulness. To measure these constructs, we use the corresponding questions from TAM2 (Venkatesh & Davis 2000). However, the pretest indicated that perceived usefulness and relevance and intentions to use the system and voluntariness cannot be separated because they display low discriminant validity, so we merged these constructs. The final set of constructs consists of the following:

- **Fears of new technologies**: This construct consists of fears arising from the extended use of technology, such as trust in technology, the fear that technology will replace the personal component of treatment, and rejection of technology in general.
- **Fears about work**: This construct includes worries about changes to work processes. Therefore, it comprises the fear of losing one’s job and working longer, as well as the perception of the system as a burden and fears of a reduction in decision making rights.
- **Fear of surveillance**: The extended possibilities of data collection and analyzing are integrated in this construct, which consists of fears of being exposed and monitored during work. Each item entails five question items.
- **Perceived usability**: This construct measures whether participants regard the system as useful and relevant.
- **Intention to use the system**: This construct measures the intention to use technology voluntarily for work purposes.

Perceived ease of use (Venkatesh & Davis 2000) does not appear in the scenario, because it is not possible to measure and because previous studies show it is not relevant in the health care environment. However, the fear of technology construct measures, in general, whether people expect that using the system will lead to problems. Table 1 summarises the tested hypotheses.

### Table 1. Research model hypotheses.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Fears in relation to new technologies will have a significant effect on fears of surveillance.</td>
</tr>
<tr>
<td>H2</td>
<td>Fears in relation to new technologies will have a significant effect on fears about working conditions.</td>
</tr>
<tr>
<td>H3</td>
<td>Fears in relation to new technologies will have a significant effect on usability.</td>
</tr>
<tr>
<td>H4</td>
<td>Fears in relation to new technologies will have a significant effect on intentions to use.</td>
</tr>
<tr>
<td>H5</td>
<td>Fears of surveillance will have a significant effect on fears about working conditions.</td>
</tr>
<tr>
<td>H6</td>
<td>Fears of surveillance will have a significant effect on usability.</td>
</tr>
<tr>
<td>H7</td>
<td>Fears of surveillance will have a significant effect on fears about intentions to use.</td>
</tr>
<tr>
<td>H8</td>
<td>Fears of working conditions will have a significant effect on usability.</td>
</tr>
<tr>
<td>H9</td>
<td>Fears of working conditions will have a significant effect on intentions to use.</td>
</tr>
<tr>
<td>H10</td>
<td>Usability will have a significant effect on intentions to use.</td>
</tr>
</tbody>
</table>

### 3.2.2 Research scope

The objective of the quantitative research design is to evaluate the hypotheses with a paper-based questionnaire. The questions were measured on a seven-point Likert scale (Dawes 2008). A pretest of the questionnaire with professional nurses in Germany resulted in minor changes to the questions. A total of 224 participants (111 from India, 113 from Germany) completed the questionnaire in the main test, and 215 were appropriate for use in the study (9 were incomplete). The participants were undergraduate nurses in India and Germany with at least one year of experience, which ensured they were acquainted with the working environment of a hospital. These participants provide a contrast with the interviewees in leading positions in the preliminary qualitative study. Furthermore, we assumed nurses would be a relatively homogenous group compared with physicians and focus more often on medication processes as a main part of their daily work. The use of students as participants might limit the generalisability of the results, and transferring the qualitative developed items from Germany to India could limit the results as well. To reduce this concern, the items were translated into Hindi as well as English and then retranslated into German. Any differences prompted minor changes,
with the translation process repeated. A researcher also discussed the questionnaires with a small group of participants to ensure their understanding. In both countries, the data collection followed the same procedure: The researcher explained the aim of study to a class of nurses in training, emphasising the voluntariness of participating and ensuring the anonymous character of the results. The questionnaires then were distributed and collected. The average ages of the participants were 21.80 years in Germany and 24.49 years in India. The difference results from the different educational systems in these countries. Most people in India have more working experience before they go to nursing school. In Germany, 81.25% of the participants are women (18.75% males), and in India, 73.82% are women (27.18% males).

3.2.3 Data analysis method

To investigate the hypothesis, we used partial least squares (PLS) modelling. An advanced statistical method that belongs to the structural equation modelling (SEM) domain, PLS allows for the empirical analysis of a measurement model and a structural model. The structural model consists of a network that links the latent and manifest variables, and the measurement model connects the constructs with a set of indicators (Wold 1974, 1982). In contrast with traditional statistical methods such as factor analysis, regression analysis, and path analysis, PLS assesses the measurement model embedded in the structural model and thus uses an iterative algorithm that estimates indicator loadings on the construct and then among the constructs (Fornell & Larcker 1981).

We select PLS to test the hypotheses because its required sample sizes are relatively smaller than those for other SEM approaches. Moreover, PLS can test theories in an early stage of research. Various information systems studies use the same method to analyse data (Keil et al. 2000, Venkatesh 2000, Venkatesh et al. 2003). We adopt SmartPLS 2 (Ringle & Wende & Will 2005).

4 RESEARCH RESULTS

The qualitative research in Germany enabled us to identify the constructs and select the items. We next explain the measurement and structural model in detail.

4.1 Measurement model

The investigated model is reflective, so the latent variables are operated by the measurement models and explain the indicators. The strength of the measurement model is determined by its reliability and validity. We calculate the reliability of the single measure, composite reliability of constructs, variance extracted by the constructs, and discriminant validity (Fornell & Larcker 1981, Hair et al. 1998).

To assess the reliability of the single measure, we note the correlation of the indicator and the construct. The factor loading measures reliability as a score that should be greater then 0.7. Factor loadings less than 0.5 are not acceptable and therefore are excluded from the analysis (Chin 1998). If a factor is excluded for one country, we also exclude it for the other country to ensure comparability. In the final model, all the factor loadings of the indicators are greater 0.5, and most are higher than 0.7, so the constructs achieve reliability in both countries.

<table>
<thead>
<tr>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Factor Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Germany</td>
<td>India</td>
</tr>
<tr>
<td>Feas about technology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>It is always better to depend on humans rather than such systems.</td>
<td>3.321</td>
<td>3.009</td>
</tr>
<tr>
<td>I regard it as problematic to trust such a system.</td>
<td>3.532</td>
<td>3.495</td>
</tr>
<tr>
<td>Such a system contradicts ethical values.</td>
<td>3.866</td>
<td>3.058</td>
</tr>
<tr>
<td>I would be reluctant to have more to do with technological devices.</td>
<td>4.277</td>
<td>3.330</td>
</tr>
</tbody>
</table>
Table 2. Descriptive findings and factor loadings.

<table>
<thead>
<tr>
<th>Fears about work conditions</th>
<th>Mean Germany</th>
<th>Standard Deviation Germany</th>
<th>Factor Loading Germany</th>
<th>Mean India</th>
<th>Standard Deviation India</th>
<th>Factor Loading India</th>
</tr>
</thead>
<tbody>
<tr>
<td>Such a system will probably result in more errors than less.</td>
<td>3.712</td>
<td>2.124</td>
<td>0.606</td>
<td>3.107</td>
<td>1.726</td>
<td>0.826</td>
</tr>
<tr>
<td>The usage of the system will have negative consequences for me.</td>
<td>4.718</td>
<td>1.612</td>
<td>0.722</td>
<td>3.049</td>
<td>1.530</td>
<td>0.707</td>
</tr>
<tr>
<td>Such a system will lead me to work more overtime.</td>
<td>4.469</td>
<td>1.659</td>
<td>0.772</td>
<td>3.485</td>
<td>1.691</td>
<td>0.781</td>
</tr>
<tr>
<td>The system will be an additional burden.</td>
<td>3.864</td>
<td>1.732</td>
<td>0.812</td>
<td>3.252</td>
<td>1.934</td>
<td>0.812</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fear of surveillance</th>
<th>Mean Germany</th>
<th>Standard Deviation Germany</th>
<th>Factor Loading Germany</th>
<th>Mean India</th>
<th>Standard Deviation India</th>
<th>Factor Loading India</th>
</tr>
</thead>
<tbody>
<tr>
<td>The idea that I cannot avoid the surveillance of the system frightens me.</td>
<td>4.330</td>
<td>1.862</td>
<td>0.844</td>
<td>3.019</td>
<td>1.547</td>
<td>0.699</td>
</tr>
<tr>
<td>I find it objectionable when I do not know what will be recorded.</td>
<td>3.232</td>
<td>1.831</td>
<td>0.760</td>
<td>2.563</td>
<td>1.493</td>
<td>0.810</td>
</tr>
<tr>
<td>I would be afraid that I would get exposed through such a system.</td>
<td>4.402</td>
<td>2.016</td>
<td>0.801</td>
<td>2.689</td>
<td>1.553</td>
<td>0.816</td>
</tr>
<tr>
<td>I find it objectionable that I cannot change the saved data.</td>
<td>4.153</td>
<td>1.942</td>
<td>0.767</td>
<td>3.165</td>
<td>1.547</td>
<td>0.793</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Perceived usefulness</th>
<th>Mean Germany</th>
<th>Standard Deviation Germany</th>
<th>Factor Loading Germany</th>
<th>Mean India</th>
<th>Standard Deviation India</th>
<th>Factor Loading India</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using the system would improve my job performance.</td>
<td>4.376</td>
<td>1.728</td>
<td>0.872</td>
<td>2.165</td>
<td>1.276</td>
<td>0.679</td>
</tr>
<tr>
<td>In my job, usage of the system is of high relevance.</td>
<td>4.643</td>
<td>1.565</td>
<td>0.829</td>
<td>2.515</td>
<td>1.392</td>
<td>0.745</td>
</tr>
<tr>
<td>Using the system would help me do a better job.</td>
<td>4.469</td>
<td>1.825</td>
<td>0.891</td>
<td>2.418</td>
<td>1.249</td>
<td>0.712</td>
</tr>
<tr>
<td>Through the use of the system, the productivity of the hospital could be improved.</td>
<td>3.946</td>
<td>1.825</td>
<td>0.691</td>
<td>2.204</td>
<td>1.396</td>
<td>0.800</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intention to use the system</th>
<th>Mean Germany</th>
<th>Standard Deviation Germany</th>
<th>Factor Loading Germany</th>
<th>Mean India</th>
<th>Standard Deviation India</th>
<th>Factor Loading India</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under the assumption that the system would work as described, I intend to use it.</td>
<td>2.847</td>
<td>1.861</td>
<td>0.900</td>
<td>2.466</td>
<td>1.327</td>
<td>0.760</td>
</tr>
<tr>
<td>I would use such a system voluntarily.</td>
<td>3.634</td>
<td>1.870</td>
<td>0.881</td>
<td>2.204</td>
<td>1.023</td>
<td>0.647</td>
</tr>
<tr>
<td>It would not be necessary that the system be dictated by rules and regulations in order for me to use it.</td>
<td>3.846</td>
<td>1.846</td>
<td>0.591</td>
<td>2.252</td>
<td>1.405</td>
<td>0.760</td>
</tr>
<tr>
<td>If the mobile device is small, I would like to wear it.</td>
<td>3.268</td>
<td>2.127</td>
<td>0.866</td>
<td>2.233</td>
<td>1.315</td>
<td>0.767</td>
</tr>
</tbody>
</table>

Table 3. Validity findings of the measurement model.

All the Cronbach’s Alpha values are greater than 0.7. An acceptable internal consistence requires composite reliability greater than 0.7 as well. The average variance extracted (AVE) should be greater than 0.5. To assess the discriminant validity for each value of the latent construct, we determine whether the AVE is greater than the maximum squared correlation between this construct and other constructs (Chin 1998, Fornell & Larcker 1981). As we show in Table 3, all AVE values are greater than the corresponding maximum squared correlations with other constructs, so the tested models in both countries achieve this requirement, and we can assume the discriminant validity of the results.

<table>
<thead>
<tr>
<th>Fears about technology</th>
<th>Cronbach’s Alpha Germany</th>
<th>Composite Reliability Germany</th>
<th>AVE Germany</th>
<th>Max. Squared Correlation Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fears about work conditions</td>
<td>0.776</td>
<td>0.856</td>
<td>0.598</td>
<td>0.450</td>
</tr>
<tr>
<td>Fear of surveillance</td>
<td>0.714</td>
<td>0.821</td>
<td>0.536</td>
<td>0.442</td>
</tr>
<tr>
<td>Perceived usefulness</td>
<td>0.880</td>
<td>0.872</td>
<td>0.630</td>
<td>0.211</td>
</tr>
<tr>
<td>Intention to use the system</td>
<td>0.840</td>
<td>0.894</td>
<td>0.680</td>
<td>0.500</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fears about work conditions</th>
<th>Cronbach’s Alpha India</th>
<th>Composite Reliability India</th>
<th>AVE India</th>
<th>Max. Squared Correlation India</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fears about work conditions</td>
<td>0.753</td>
<td>0.845</td>
<td>0.581</td>
<td>0.578</td>
</tr>
<tr>
<td>Fear of surveillance</td>
<td>0.791</td>
<td>0.864</td>
<td>0.615</td>
<td>0.578</td>
</tr>
<tr>
<td>Perceived usefulness</td>
<td>0.786</td>
<td>0.862</td>
<td>0.610</td>
<td>0.545</td>
</tr>
<tr>
<td>Intention to use the system</td>
<td>0.722</td>
<td>0.824</td>
<td>0.541</td>
<td>0.422</td>
</tr>
</tbody>
</table>

Table 3. Validity findings of the measurement model.
Not all the facets from the model developed in Germany appear in the data from India. As a result, we deleted these indicators from both models, because the model is reflective (Chin 1998). Nevertheless, these findings are important in that they show that the constructs differ across countries. In our model, we focus on describing only the intersection of both models as a comparable core.

### 4.2 Structural model

Variance-based approaches to structural equation modelling and PLS tend to bias the results. Indicator loadings in the measurement model are estimated as too high, and the path coefficients in the structural model are estimated as too low (Chin 1998). This bias can be avoided with either an adequate sample size—at least 10 times higher than the most complex construct’s number of indicators or the largest number of paths leading to a latent construct. In the model, the most complex model consists of five indicators, and the four paths for intention to use the system are the most (Figures 1 and 2), so we require a minimum sample size of 50. Even though both sample sizes achieve this requirement, it is only a rule of thumb; critics argue that this rule does not take statistical power into account. Significance tests that lack statistical power are limited because they cannot discriminate reliably between the null hypothesis and an alternative hypothesis. Thus, the possibility exists that not all significant paths in the model will be identified. Furthermore, the statistical power should be greater than 0.8 (Cohen 1988, Hair et al. 1998). Using G*Power 3.0 for a moderate effect size of 0.15, an α of 0.05 for four predictors (latent variables), and a statistical power of 0.8, we need a sample size of at least 85 (Faul et al. 2007). Because both samples are greater than 100, the statistical power is acceptable (> 0.87), bias problems can be overcome, and the squared multiple correlation R² in the upper 13% will be detected as significant.

After computing the path coefficients in the structural model, we used bootstrapping to obtain the corresponding t-values. All hypotheses correspond to paths in the structural model. Each hypothesis test uses a path coefficient (positive or negative) and the statistical significance of the t-values. In the t-tests, we assess the 0.05, 0.1, and 0.01 significance levels (Efron & Tibshirani 1993). In Figures 1 and 2, we outline the path coefficients, t-values, and significance levels.

![Figure 1](image_url)  
*Structural model: German data.*  
(****p < 0.001, **p < 0.01, * p<0.05; hypotheses in bold are supported)*
As in the regression analysis, the $R^2$ indicates the explanatory power of the latent endogenous variables. For example, Chin (1998) regards 0.67, 0.33, and 0.19 as substantial, adequate, and low values, respectively. We use these same values to evaluate the findings. To assess the relevance of the prognoses ($Q^2$), we adopt the Stone-Geisser criteria with a blindfolding procedure (Tenenhaus et al. 2005). This approach eliminates values in the empirical data and approximates them with the resultant PLS findings. The procedure ends after all values have been replaced. Using squared errors and approximated values, $Q^2$ can be assessed; values greater than 0 suggest the constructs have relevance for the prognoses. In both models, the values for $Q^2$ are greater than 0.

![Figure 2. Structural model: Indian data.](attachment:image.png)

(Table and text continue)

5 DISCUSSION AND IMPLICATIONS

The concerns about technology, as identified in the qualitative study, have significant influences on perceived usability and intentions to use the technology in Germany. Moreover, all of the hypotheses receive support from the study findings (Figure 1). However, only four—H1, H2, H8, and H10—have significant influences in India (Figure 2). The unease associated with technology in Germany and India also affects users’ fears of surveillance and work degradation. In other words, people who do not trust technology and do not want to increase the extent to which they use it in their work also express a fear of being under surveillance and believe that their work conditions will worsen with the introduction of such technologies. This critical point shows that acceptance issues cannot be described sufficiently by the components of classic acceptance models such as TAM2 (Venkatesh & Davis 2000). New technologies such as ambient intelligence, which imply additional risks for users, require adapted acceptance models that take these risks into account. The high correlation between fears illustrates a close connection among the various risks posed by new technology. Therefore, attempts to reduce the fear of new technology may reduce other fears and perhaps increase acceptance levels.

The fear that a user’s work will be degraded because of the introduction of a technology has a significant negative influence on the perceived usefulness of ambient intelligence in both Germany and India. This negative influence confirms that people tend to consider technologies less useful when they anticipate negative consequences for themselves. In the Indian data, this fear explains only a minimal lessening of perceived usefulness ($R^2=0.084$) and has no direct affect on the intention to use the technology. However, in Germany, fears of technology and anxiety about work degradation have
considerably negative influences on both perceived usefulness and intentions to use the technology. However, the surveillance aspect has a positive effect on perceived usefulness and intentions to use a technology; that is, participants appear to understand the positive aspects of surveillance and regard them as useful. However, H1 and H5 imply that technology is very closely related to two other fears with negative implications. In another interesting finding, we note that in India, the only significant influence on the intention to use the technology is perceived usefulness. The correlation between usefulness and intention to use the technology is also much higher in India. In contrast, German respondents indicate that all their fears have an impact on their intention to use the technology. Thus, it seems relevant to take the perceived risks in Germany carefully into account.

6 CONCLUSION AND FURTHER RESEARCH

The main objective of this study is to determine whether special fears about technology influence acceptance of ambient intelligence and whether national culture also might affect attitudes of potential users. The results significantly support the hypotheses. In Germany, we identify three major concerns, each of which affects the perceived usefulness of and intentions to use the technology. However, in India, the findings suggest that only those fears related to the working environment are critical for ambient intelligence acceptance. These results imply that ambient intelligence prompts much more criticism in Germany than in other countries. It is important to allay such scepticism by supporting the implementation process better. Strong relationships among the different fears about the technology indicate it also is important to try to avoid them by adopting a holistic approach. According to the German study, surveillance can have a positive influence on the acceptance of such technologies. In contrast, the Indian findings suggest that only those fears related to work are critical for implementation, which might be enhanced through better communication with the staff.

As mentioned, our study suffers from several limitations. First, the participants are comparatively young and had minimal experience because they were still in training, so our results might not be generalisable to a wider audience. Second, our qualitative development of items in Germany and their transfer to the Indian health care system might lead to bias in the results. The transferability to other application areas (e.g., collaboration, patient monitoring) and other groups of medical staff also cannot be assumed. Third, the results are not necessarily representative of the investigated cultures. Fourth, intention to use a technology serves, in this study, as an indicator of acceptance, but a bias between intentions and actual use is likely.

In a next step, these results could provide the basis for implementing an ambient intelligence system and performing a case study. Research into the changing level of acceptance over time and comparisons across different application areas and users also would improve our understanding of related ambient intelligence acceptance problems. We suggest that the relationships between the cultural dimensions, according to the GLOBE study, and the level of acceptance should be analysed in detail. Finally, other countries could be integrated in further research projects. The acceptance of special technologies is highly relevant because of their enormous benefits. As the comparison between India and Germany demonstrates, certain countries participate in the use of ambient intelligence more readily than others, because they are more open to using new technologies in their working environments. A better understanding of the special fears and risks associated with ambient intelligence thus is required to motivate people with more critical attitudes to participate in the use of such technologies.
References


**WHY INFORMATION TECHNOLOGY IS NOT BEING USED FOR FINANCIAL ADVISORY**

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0400.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Technology adoption, IT Business Alignment / Value, Stakeholder theory, Theory Building</td>
</tr>
</tbody>
</table>
Why Information Technology Is Not Being Used for Financial Advisory

Schwabe, Gerhard, University of Zurich, Binzmuehlestrasse 14, 8050 Zurich, Switzerland, schwabe@ifi.uzh.ch
Nussbaumer, Philipp, University of Zurich, Binzmuehlestrasse 14, 8050 Zurich, Switzerland, nussbaumer@ifi.uzh.ch

Abstract

Swiss banks have returned to their roots and pay an increasing amount of attention to differentiating themselves from others through good financial advisory services. This has led to a loudly publicized standardization of IT-advisory processes, but not to an increasing use of supporting IT tools. This paper uses interviews with Swiss advisors, sales managers and IT managers, as well as focus groups of users and a survey with users to identify reasons for non-usage. The analysis is based on a framework combining principal-agent theory, IT-business alignment, technology acceptance and information behaviour. We provide evidence that the key problem explanation is the incentive system of the advisors and that poor usability of the software and lack of engagement by sales managers also contribute to the non-usage of most tools.
1 INTRODUCTION

The breaking down of the asset management market has led to a rediscovery of the traditional banking wisdom and banking business. Giving credit to companies and advising clients in their investments are increasingly seen as an attractive and sustainable source of income. As for them cost leadership is not regarded a promising strategy in the long run (Buhl, Kundisch & Steck 2002), an increasing number of financial service providers has turned to financial advisory services as a competitive differentiator. Buhl & Kaiser (2008) argue that such a differentiation may be accomplished by individualizing advisory, i.e. by aligning the advisory activities to the customer needs. Taking into consideration the complexity of advisory, they also conclude the need of extensive IT support for improving advisory services regarding efficiency and effectivity.

Switzerland is one of the leading worldwide destinations for wealth management and enjoys the reputation of being a worldwide leader in financial advisory processes. In recent years, leading Swiss banks have placed considerable effort into restructuring their advisory process in order to further improve their quality. However, studies repeatedly reveal that customers are notoriously dissatisfied with advisory services; we were able to confirm this in the context of a larger study on the investment advisory process (University of Zurich & Solution Providers 2009). In this study we also investigated the advisory's IT support, i.e. the tools advisors have available to better serve their customers, particularly in face-to-face meetings. Our observations regarding IT in advisory were based on a broad definition of IT support - we were interested in the categories of available advisory IT (apart from the widespread tools for customer relationship management (CRM) virtually every bank employs): means of IT-enabled communication (such as email, chat, instant-messaging), IT tools for information access (stock prices, ratings), tools for decision support (simulation and visualizations of scenarios) as well as IT-support for customer self-advisory (such as tax calculators).

We were puzzled to see that almost no banks used IT during the actual advisory meeting, although there were tools available. We even looked over the Swiss-German border and noticed that advisors of Deutsche Bank also did not use IT during the advisory meeting even though Deutsche Bank had won an award for its advisory software (Voss 2005). In the course of pilot interviews it became clear that this phenomenon could be explained not only by simple “usefulness” and “ease of use” as suggested by the popular Technology Acceptance Model (Davis 1989); rather, advisory IT appeared to inject itself into a complex web of interests and accepted behaviour of key bank stakeholders.

We therefore decided to look at these issues from a stakeholder’s perspective. Four groups of stakeholders were analyzed: the advisors in the banks, their clients, the sales managers responsible for the advisory process, and the IT managers responsible for building tools for supporting the advisory process. All their comments in the interviews were scanned for clues on “Why is IT not used for financial advisory services”? Consequently, the objective of this paper is to provide insight into the reasons why financial service providers fail in tapping the full potential of IT for improving advisory, especially in meetings with customers. Building on different well-established theories we thereby strive to contribute to a better understanding of the behaviours of actors and their effects in an organizational context.

The rest of the paper is organised as follows. The subsequent section reviews relevant literature and identifies useful theories to guide our search for an answer to our research question. The third section offers an overview of our research design, describing the collection and analysis of the data. The findings, presented in section 4, illuminate the problem in the eyes of each of the stakeholders. Section 5 synthesizes and critically discusses the findings, culminating in a preliminary answer to our research question. The final section closes with limitations of our research and the need for further research to support our claims and to deepen the understanding reached thus far.
2 LITERATURE REVIEW

There is no generally agreed upon definition of financial advisory. Fischer & Gerhardt (2007) provide a general definition that describes a financial advisor as “a person or organization that offers its professional financial expertise to individuals who seek assistance or want to completely delegate their investment decisions” (p. 9). From a financial service provider's perspective, the advisory process may be seen as the structured arrangement of advisory activities, with the objective to standardize and control these activities. For investment advisory, these activities may be arranged into a process with multiple phases of customer interaction, such as customer contact (collecting basic customer information), actual advisory (developing investment strategies based on the customer's situation and requirements, mostly in face-to-face meetings), realization (implementing strategies into specific product portfolios) and periodic customer support (updating the customer's requirements and optimizing strategies/portfolios). Similar definitions and arrangements of advisory activities can be found both in literature (e.g., Homburg, Schäfer & Schneider 2003, Kallhardt 1997, Lippitt & Lippitt 1984) and practice (e.g., the four steps advisory process of UBS).

We chose to base our research on financial advisory on the perspectives of the main actors involved in an advisory process. This approach reaches beyond organizational boundaries (looking at advisors and customers), organizational levels (covering both advisors and the managers responsible for the advisory process) and functions (IT and business). Accordingly, the areas traditionally covered in different IT and business subdisciplines are combined (see Figure 1). On the operational level, theories and models explaining the behaviour of the actors are selected. For clients, this is “information behaviour” and for advisors, the corresponding “advisory behaviour.” On the management level, the corresponding management disciplines provide models; for sales managers, the discipline is “marketing”; for IT management, the discussion on “technology acceptance” leads to the field of “usability.” Further, theories and models of the relationships of the actors are used: The relationships between clients and financial advisors, as well as financial advisor and the sales manager, can be analyzed by “principal-agent theory.” Any changes have to overcome the conflict using “Change Management” models. The relationship between sales managers and IT managers is covered in more detail in the “IT-business alignment” literature.

Clients: Typically, the customer’s buying cycle (Muther 2001) is not as systematic as the bank’s selling cycle. For financial advisory services, the customer’s buying cycle can be best understood using models from information behaviour (e.g., Wilson 1999). They explain how clients search, select and use information, as well as show the settings that they prefer. These models are particularly useful in explaining situations where a client has to uncover information needs s/he is not yet aware of (Belkin et al. 1982). Advisory service is typically a powerful strategy in such situations (Schwabe et al. 2008), fulfilling pragmatic information needs.

Figure 1: Integrated Non-Usage Framework

---

1 http://www.ubs.com/1/e/ubs_ch/wealth_management_switzerland/relationship/advisory_approach.html
**Client – Advisor:** The relationship between client and advisor has the features of a principal-agent conflict (Eisenhardt 1989, Golec 1992): The advisor (= agent) is paid to work for the client (= principal), but s/he may have hidden characteristics (e.g., lack of knowledge about the market) and hidden intentions (e.g., optimizing fees for the banks instead of the investment profit of the client) and may engage in hidden actions (e.g., selling assets without the real consent of the client). S/he may not want to use IT, which, for example, may make investment solutions more transparent and thus may allow the client to better monitor the advisor’s behaviour.

**Advisors:** We are not aware of any published empirical models on actual advisory behaviour. There are, however, ample normative models on how advisors should behave during the advisory process. These models originate from advisory service process design and banks promoting their advanced advisory process (see definitions above), from the finance literature, e.g., promoting a theoretically sound portfolio choice (Markowitz 1991) and from IT tool descriptions suggesting a suitable advisory approach (Buhl, Kaiser & Winkler 2007, Eberhardt & Zimmermann 2007).

**Advisor – Sales manager:** The relationship between advisor and sales manager can also be regarded as a principal-agent conflict. While the sales manager is responsible for designing an advisory process optimizing the bank’s long-term income, an advisor may focus rather on maximizing her personal short-term income through fees. Furthermore, it may be in her personal interest to bind the customer to herself as a person rather than to the bank as a whole. This may foster hidden intentions and lead to hidden actions. The advisor again may have no interest in implementing IT that could make her intentions more transparent or may prevent her from engaging in hidden actions. Swiss banks have implemented standardized advisory processes to address this known problem. Implementing supporting IT could greatly enhance their capability to enforce these new processes.

The literature on change management may enhance our understanding of why sales managers have failed in implementing advisory IT. In the context of the advisory process the sales manager owns, an understanding of conflicts of interests as well as an active design and change of incentive systems is of particular importance. As advisory technology supports client-advisory collaboration in an organizational context, the literature on groupware adoption is the most applicable regarding change management issues. In her seminal work on the failed diffusion of a Lotus Notes based knowledge management system, Orlikowski (1992) shows how the highly competitive culture of a consulting company (“up or out”) contradicts the sharing approach that lies at the heart of the knowledge management system. Top managements’ efforts to enable and persuade the consultants to use the software remained fruitless in this context. Grudin (1988) uses the example of shared calendars to demonstrate that implementation of groupware requires a critical mass of users to make them useful for everybody. Only peer pressure (Grudin & Palen 1995) may convince users to compromise their own interests (e.g., in privacy, flexibility, or, more generically, in hiding their characteristics, intentions and actions) for the sake of everybody else’s benefit (improving the coordination and increasing the efficiency). Generally, the organizational implementation of software is difficult if one group of actors has the costs and another group of stakeholders reaps the benefits (Grudin 1994). Furthermore, successful changes require a shared vision and the affected users have to be enabled to perform the change. Ward and Daniel (2006) establish the managers’ responsibility to actively manage IT benefits on the basis of a sound understanding of the stakeholders’ interests and the needed business changes.

**Sales manager:** Sales management strives for a competitive advantage (Porter 1985). The implementation of IT is interesting from the sales manager’s strategic perspective if it serves as strategic differentiator (Ward & Peppard 2002). From an operational point of view, IT can serve as a means of monitoring the advisors to avoid hidden actions and to enforce quality standards.

**Sales manager – IT Manager:** The alignment of IT to the needs of the business remains on the top of a CIO’s agenda (McGee 2008). Successful companies establish joint decision making on the business demand for IT support (Weill & Ross 2004). Ward and Peppard (2002) recommend including the management levels of IT and business, as well as the user levels in determining the demand and the steering of IT projects. IT managers are requested to actively market their products in business (OCG
2007), and sales managers and users need to actively participate in IT projects, even though it may not be their primary interest.

**IT Manager:** The IT managers strive to provide acceptable software. Davis’ (1989) model on technology acceptance (TAM) proposes that this mainly depends on usefulness (i.e. the functionality it provides) and the ease of use. Newer literature from HCI expands the idea of “ease of use” to the concept of “usability.” Apart from having pragmatic qualities of the software (e.g., whether the user finds the information s/he is searching for within reasonable time), the rising popularity of computer games has stressed the importance of the hedonic qualities of IT (Heijden 2004), i.e. how enjoyable IT use is.

The TAM has been frequently used, but also been frequently criticized and expanded upon. Newer models like the Unified Theory of Acceptance and Use of Technology (Venkatesh et al. 2003) include contextual constructs of performance expectancy, effort expectancy, social influence and facilitating conditions. The Model of Information System Success (DeLone & McLean 2003) describes the relations between information, system and service quality and use of information systems, correctly declaring that the net benefits of such systems are dependent from usage and user satisfaction. Current technology acceptance models explain the adoption behaviour of individuals, teams in special settings (Powell, Piccoli & Ives 2004) or whole organizations (Bajwa et al. 2007). They are not yet specifically tailored to explain adoption of dyads or to link individual adoption to organizational adoption.

In summary, there are considerable general building blocks that can contribute to an understanding of why IT is not used in the advisory process. However, we have also identified a few holes. What is still missing is (a) a model that is specifically tailored to financial advisory issues and (b) a combined and coherent full picture that identifies the key factors and also explains the interplay of these factors.

## 3 Research Design and Data Collection

The question of why IT is not used for financial advisory processes surfaced in the course of a larger study on the quality of financial advisory services (University of Zurich & Solution Providers 2009) in which IT use was only one question. We were thus fortunate to have a very rich set of data to contextualize the observed IT use. The larger study combined an adapted and enhanced Servqual model (Zeithaml et al. 1990) and an information behaviour model (Wilson 1997). Following the Servqual gap model, we contrasted the perceptions of bank customers, advisors and the sales managers responsible for the advisory process. Typically, these were senior executives, one or two organizational levels below the company board. We then added the IT management perspective, addressing the person responsible for advisory IT, typically a senior IT manager. The questionnaires were based on the Servqual items for the five dimensions of service quality (reliability, assurance, tangibles, empathy, responsiveness) and extended by items for the advisory process dimension (effectivity and efficiency) as well as IT-related questions. In this paper we focus on the results of the IT items and of the items on Servqual’s assurance dimension (i.e. confidential behaviour of employees), which add to the discussion on the importance of trust regarding information asymmetries.

All four stakeholder groups received the same core set of questions and each received a set of additional questions targeted at their special competence. The most extensive enhancement was on the client side where we applied a survey of their information behaviour, building on Wilson’s model (Wilson 1997) and asking which sources they preferred for their information search, how they judged them in comparison and why they chose to do so.

The data collection took place in two waves. From December 2007 to February 2008, four researchers performed a total of 21 “mystery shopping episodes” that were performed in Switzerland (16) and - for the purpose of comparison - in Austria (3) and Germany (2). The episodes entailed consulting talks with retail banks (12), private banks (6) and one provider of bank assurance. Each session typically lasted from 60 to 90 minutes. Targeting the affluent banking customer segment in our study, the researchers requested advice about the investment of sums of 50,000 to 500,000 Swiss francs, which generally classifies a customer as “affluent” by a bank. For details see Nussbaumer et al. (2009).
In the second wave of data collection, which lasted from June to October 2008, we focused on financial service providers in the Swiss banking region\(^2\) to obtain a more homogeneous sample (European banks, for example, show substantial differences regarding legal regulations on advisory). A total of 21 advisors from 19 banks, 28 responsible sales managers from 27 banks and 12 IT managers from 12 banks were interviewed. As we strove for deeper insights, the questionnaire consisted not only of quantitative questions from Servqual, but also open-ended questions for each of the dimensions. This paper draws mostly from these open-ended questions. A typical interview lasted 60 to 90 minutes. All sales manager interviews were transcribed\(^3\), and in two cases where interviews could not be conducted, the respondents answered the questions electronically. We have audio recordings plus written notes for the interviews with the other stakeholders.

For the customer perspective, we applied electronic questionnaires and involved customers in a written discussion as part of two focus group sessions. The workshops lasted 150 minutes each. A total of 28 clients participated, almost all of which belonged to the affluent customer segment. We furthermore collaborated with a very popular investment journal (Cash\(^4\)) to apply purely quantitative questionnaires on advisory quality and information behaviour in September/October 2008 (for details see University of Zurich & Solution Providers 2009). A total of 136 users responded to the survey, with the vast majority again being from the affluent customer segment. The online survey was based on refined questionnaires that were used for the focus groups. The applied triangulation of methods (Flick 2000) allowed both finding support for existing working hypotheses and seeking for a deeper understanding in areas less well understood.

The subsequent paragraphs describe in more detail the data that were actually used for explaining why IT was not used in the financial advisory services. The overall research question of “Why is IT not used for financial advisory services?” can be split into three subquestions: (1) “Is IT not used in financial advisory services?” establishes the basis for our argument, (2) “What reasons are given by the stakeholders?” establishes the multiple-perspective view for which we argued in the literature review, with (3) “Which of the reasons appear credible?” we interpret the findings.

The research on “Is IT not used in financial advisory services?” primarily draws on our observation in mystery shopping. We can further support the analysis with (a) the outcome of two student research projects (à Porta 2008, Hämmerli 2008) covering another 10 advisors, and (b) results from the interviews with advisors.

The research on “What reasons are given by the stakeholders?” draws from the interviews with the advisors, sales managers and IT managers, as well as from the workshops with clients. All the gathered data was scanned for remarks that could indicate a reason. Any applicable remark was noted and later synthesized in an overview map (see Figure 2). Most remarks came from the IT sections of the questionnaire, but a significant number was also mentioned in other sections. For a few observations we have further support from the quantitative sections of the questionnaires. Most importantly, we can give quantitative evidence on the clients’ expectation of IT and the other stakeholders’ views on this.

The research on “Which of the reasons appear credible?” is mainly interpretative. We particularly focussed on contradictory claims and their support by theory.

\(^2\) including Liechtenstein

\(^3\) Only one interview partner refused recording.

\(^4\) http://www.cash.ch
4 FINDINGS

4.1 Is IT not used in financial advisory services?

The advisors did not use IT in any of the observed 21 mystery shopping episodes. We had purposefully added Deutsche Bank to our sample because Deutsche Bank had received an award for its financial advisory solution; however, it was also not used there. Further, a student of ours interviewed 7 advisors in his Bachelor Thesis (à Porta 2008) – all of whom critically discussed IT usage in advisory meetings and emphasized customer expectations. Another student modelled the advisory process in three banks as a part of a small research project (Hämmerli 2008). He was able to identify one advisor who frequently used a computer during the actual face-to-face session. He claimed to benefit from the efficiency of the structured process and the increased shared understanding facilitated through the visualization of an investment choice on the screen in the advisory meeting. Interestingly, this advisor belonged to a retail bank (where efficiency is more an issue than in private banks). This lone exception can be seen as proof that an efficient and effective use of IT is possible in advisory meetings, but it does not change the overall picture. We therefore conclude: IT is not used during advisory meetings.

This picture changes if one includes back-office use. The advisors who prepared for the meeting (a minority) obviously had gathered information from the banking system; however, they printed the information and brought it to the meeting. Both interviewed advisors and sales managers stated that CRM systems were used to collect information on their interaction with customers and special systems were used to enforce compliance (e.g., to money laundering laws). Yet the systems appear to remain invisible during the actual advisory meeting. The following types of information systems are rarely or never used in the advisory process: decision support systems, e.g., allowing the customer to visualize and simulate investment options, and systems for self-advisory, e.g., over the Internet.

As we were not able to identify any publications on the diffusion of advisory IT into advisory meetings, we relied on our own Swiss data. We thus concluded that some IT systems are used during advisory services, but never during the actual advisory meetings. The subsequent sections provide reasons for why no systems are used during the actual advisory meeting and why only the mentioned types of systems are used in the advisory process at all.

4.2 What reasons do the stakeholders give?

Figure 2 gives an overview of the reasons for non-usage, as given by the interview partners. If a stakeholder group places the responsibility with themselves, the reason is put above the boxes. If it is in the eyes of stakeholder A and it is stakeholder B that is responsible, the reason is written on an arrow pointing from A to B.

The advisors’ perspective: Advisors state that they do not feel confident in using IT with customers. They fear the loss of competence in the eyes of the customer if they run into problems with using the IT system while the customer is present. The older generation has a lack of IT-affinity. Additionally, they are not convinced of the functionality provided by the tools and their usability. In many places, the IT-tools are poorly integrated and therefore require the advisors to switch tools during the process. In the survey they state that they have been properly trained in the usage of the tools; however, participants stated that isolated training of tool usage was insufficient for them to be applied in the advisory process, let alone in an advisory meeting, and that they should rather learn how to embed tool usage into the actual process and its diverse activities. It is thus not surprising that the benefits of using IT tools in the advisory process remain unclear to the advisor. In the interviews it became clear that the degree of standardization of established advisory processes is rather low – in most cases neither the detailed activities of the processes, nor the usage of accordingly supportive tools, is specified or enforced by management, respectively. Also, the involvement in standardizing the process is perceived as being very low, since the management specifies the processes and implements them top-down.
However, advisors in the few banks that have established rather stringent advisory processes and integrated some IT tools conceive them to be very supportive.

**Figure 2: Overview of arguments**

The advisors claim that customers do not want IT in the advisory processes. Advisory service is said to be a very personal and trust-oriented business; IT would endanger this set-up, or – as one advisor reported – it would “destroy the magic” of advisory.

**Clients:** From the clients’ perspective, the picture is clear: There is no IT provided during the advisory meeting and therefore they cannot use it. Yet, their view of the usefulness of IT for advisory processes differs from the other stakeholders’ perceptions. When asked whether they expected IT support in very good advisory processes, the clients evaluated all offered IT support tools positively, i.e. above 5 on a Likert scale of 1= “I absolutely disagree” to 7= “I absolutely agree”. IT-support for communication (E-Mail, Instant Messaging…) was rated (5,01) as lowest, and IT-support for access to latest information (e.g., stock prices, ratings, tests) as highest (5,86). IT for self-advisory was rated 5.35. Most surprisingly, even tools for decision support (simulations and visualizations of scenarios) that can realistically only be applied in a meeting, were rated as 5.68. All ratings on what clients think are significantly above the expectations of advisors and sales managers. A much lower rating of the IT support they have experienced so far indicates that there is a significant gap. Thus, we can conclude that clients are more positive on IT support than advisors and sales managers think they are.

Trust is also not the important facilitator that advisors claim it to be; rather, the opposite is true. The surveyed clients reported a rather low level of trust towards their advisors (mean agreement of 4.72 to the statement “The advisor’s behaviour is trustworthy.”). As a client noted in a focus group, there is information asymmetry in the advisor-client relationship and this asymmetry reduces trust. Thus, clients do perceive the principal-agent conflict. This explains why IT is important in the clients’ views since it reduces information asymmetry.
Sales managers: The sales managers responsible for the advisory process were mainly marketing oriented people. In their eyes, IT - e.g. for transaction processing - is not a means of differentiation. They also do not think that IT tools supporting the advisory process (like tools for CRM or decision support) may add to a competitive advantage. Thus, for them IT is rather a “hygiene” factor: Banks have a disadvantage without it (since today, for example, transactions are processed fully electronically), but they do not have an advantage if they have it (since other banks have implemented such IT as well). IT managers mostly share this view.

The interviewed sales managers pointed out that IT use might not be in the advisors’ interest, since advisors prefer to tie the customers to themselves rather than to the company (and its information system). In Switzerland, it is not unusual for advisors to take their clients along when they switch companies. The more interesting their clients and the tighter they can bind them, the higher is the advisor’s market value. They do not like the controlling function of the IT. These sales managers’ comments show that there is indeed a principal-agent conflict between advisors and their sales managers.

Sales managers confirmed that advisors are afraid of losing face with their customers when they run into trouble with IT. They are aware of the fact that advisors use only mandatory functions. They enforce the use of tools only in two areas: IT that ensures compliance (such as laws on money laundering) and the customer relationship management system, i.e. systems that indeed can be seen as hygiene factors. For other systems, they have the same laissez-faire attitude as regarding their advisory process. Managers – especially in private banks – only hire experienced advisors. As they tend to be older and have built up experience in advising without IT support, they have difficulties appropriating IT.

Both the sales managers and the IT managers agreed that IT-business alignment remains a problem. IT is rarely included in the design of advisory processes and decisions on applying tools. Thus, the developed tools do not fit newly designed advisory processes. In the eyes of the sales managers, clients do not want IT in the advisory process. In this regard, they disagreed with some IT managers.

IT Managers: Several IT managers did see that customers want to use IT in advisory, and they also saw some benefits. They were able to more precisely estimate the clients' expectation in regard to decision support systems and information systems than advisors and sales managers. They also confirmed lack of usage by the advisors and voiced frustration about the unused expensive tools. In their relationship with the sales managers, they complained about low and late involvement in decision processes regarding the design and support of advisory processes. Internally, they had a lack of roles and processes to develop and implement innovations.

5 DISCUSSION

Some of the statements of the stakeholders reinforce one another; others are obviously contradictory. In this section we discuss the credibility of the statements in the light of other statements and the literature. Further, we link arguments to identify chains of reasons and identify root causes of non-usage.

The most striking difference is the question of whether clients want visible IT support. If they didn't want IT, it would be a plausible root cause for non-usage. While advisors and their sales managers claim that customers dislike visible IT support in the advisory process, IT managers have a more positive attitude towards such IT. Most importantly, clients themselves voiced interest in support, the highest interest being in pure information access, but there is also considerable interest in decision support and self-advisory systems. Thus, we conclude that the responding clients and IT managers are more credible and that a significant number of clients would welcome visible IT support.

The most obvious cause for lack of usage would be lack of software quality. In this regard, advisors pointed at many specific quality issues where quality seemed to be a great challenge. Particularly, the widespread poor integration into processes and low usability were major issues. The IT managers cannot push process integration ahead because they are not involved in the decision processes for advisory design. The IT-business alignment is regarded as poor from both sides in this area. An indicator of
poor IT-business alignment can be found in the valuation of IT support for advisory. It is not regarded as a means of differentiation. Therefore, it is applied only in areas where it increases efficiency (CRM) or where the bank has to document compliance with regulations. Thus, we conclude that one root cause for the lack of adoption of advisory support is its low potential for strategic differentiation. This leads to a low priority of such projects, poor IT-business alignment and an isolated IT department that can neither embed advisory support sufficiently into the advisory process, nor train users adequately.

The usability problems remain in the domain of the IT managers. There is great potential for improvement, as usability is particularly important for dyad users. Only well-designed software could ease the fears of the advisors for losing control or losing face. Furthermore, including the lessons from hedonistic software design could protect or even increase “the magic” of advisory meetings. Schwabe et al. (2008) have demonstrated for the domain of tourism, how this can be done using large screens. Yet, we are not convinced that poor usability is the root cause for the low uptake of advisory software. IT managers have bought many expensive off-the-shelf tools that have not been adopted either. We have the example of Deutsche Bank that did not use their award-winning tool. Researchers have developed advisory software with an emphasis on usability (e.g., visual product editors such as the multi-touch table developed by the Swiss Design Institute for Finance and Banking5). And we have the example of the one advisor who successfully used a standard banking system in the advisory session. Thus, we conclude that poor usability contributes to non-usage and gives advisors a welcome excuse, but it is not a root cause.

This leaves us with the two principal-agent conflicts. There is an information asymmetry between customers and advisors. This asymmetry is to the advantage of advisors and therefore they will not use IT tools without other strong incentives. Especially private banking advisors have the incentive to bind the customers to themselves. Any information that the bank has on the customer, and any binding of the customer to an information system are detrimental to this interest. Since the clients rather distrust their advisors, they have a high demand for transparency, participation and control of the advisory process. Novak (2009) discusses the design of such (collaborative) approaches to better involve end-users in advisory settings. Still, the sales managers have to enforce standardized advisory processes (and not only their design) to allow establishing a mandatory usage of more tools. They would have to apply strong change management techniques to align the interests of the customers with the interests of the advisors. However, they do not see the priorities, as IT is no strategic differentiator.

Thus, our overall conclusion is that the advisors are the main obstacles to visible IT use in the advisory process. The key to the solution is, however, in the hands of the sales managers. They do not view IT for advisory as a strategic asset, and thus neither enforce the use of IT in the advisory process (with a few exceptions), nor sufficiently collaborate with IT managers.

There are two scenarios that could change this situation: An entrepreneurial bank could demonstrate that IT can serve as a differentiator and thus force the other banks to follow suit. In the discussions, particularly the IT managers saw a realistic chance of this happening. The second scenario is a consequence of the current financial crisis. After the experience with investment bankers, banks may not tolerate uncontrollable “entrepreneurs” in their own organization any longer. They may want to enforce their processes and reap efficiency and service benefits of IT tools in the advisory process.

6 LIMITATIONS AND FUTURE WORK

Our research is targeted at a deeper and multi-faceted understanding of a yet poorly understood phenomenon. We had to make some compromises in order to reach this target within reasonable time.

5 http://www.sdfb.ch/de/aktivitaeten/projekte/strukturierte Produkte
While we were able to talk to sales managers, advisors and IT managers of all major Swiss banks, we cannot be completely certain that we have reached a representative client sample. As banks are very reluctant to provide direct access to their own customer pool, we had to accept the compromise of a public survey. The survey was launched over an electronic website; thus users may have had an IT-bias. However, as the majority of the Swiss population uses the Internet, this should not be major problem. The focus groups may also have had more IT-affinity than the average affluent client, as they were recruited from our alumni organisation and from a consultancy company. We were therefore very careful with using their interpretations in this paper and for the important parts relied rather on the client survey.

A second limitation is the emerging nature of our research question: We wanted to observe the current use of IT for advisory in the mystery shopping and talked to the stakeholders about the actual use of IT and ended up with a paper on non-usage. Thus, we rarely explicitly talked to the stakeholders about non-use but rather drew conclusions from their reactions on our questions of their usage behaviour. This approach did not allow us to discuss specific non-usage issues and we therefore could not reliably count frequency of answers.

A third limitation is the diversity of the banks covered: The arguments given by a stakeholder in one bank may be specific to this bank and may not hold true for the situation in another bank. We were surprised about the coherent picture we received by adding arguments from stakeholders in different banks. But we still had to do some interpretation and sense making in order to reach the clear picture presented in this paper.

The next steps for research in this domain could go in two directions. Firstly, we need a deeper understanding of the dynamics between the stakeholders. For this purpose, observing stakeholders within one bank could be a viable approach. Secondly, there is room to validate the barriers to the adoption of advisory systems and the causal relationships proposed in the discussion section in a quantitative study.

7 ACKNOWLEDGEMENT

We would like to thank Solution Providers, particularly Ralph Mogicato, as well as Ingrid Slembek for supporting us with the interviews and providing critical input. Furthermore, we are grateful for the inputs given by the interview partners in all major Swiss banks.

References


THE IMPACT OF LEARNING CULTURE AND INFORMATION TECHNOLOGY USE ON KNOWLEDGE-SHARING: A CASE OF KFUPM

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0037.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>knowledge transfer, Learning, Shared knowledge, Education</td>
</tr>
</tbody>
</table>
THE IMPACT OF LEARNING CULTURE AND INFORMATION TECHNOLOGY USE ON KNOWLEDGE-SHARING: A CASE OF KFUPM

Eid, Mustafa, KFUPM, Dhahran 31261, KSA, meid@kfupm.edu.sa
Nuhu, Nuraddeen Abubakar, KFUPM, Dhahran 31261, KSA, nura@kfupm.edu.sa

Abstract

Nowadays Knowledge is regarded as a strategic resource in organizations, and thus the leverage of knowledge is a key managerial issue. Knowledge creation, sharing and dissemination are the main activities in knowledge management. This study examines the influence the social and technological factors such as learning culture and IT use, could have on knowledge sharing of King Fahd University of Petroleum and Minerals (KFUPM) students. A cross-sectional survey was used as a methodology for data collection and 137 valid responses were collected from all the three categories of students that include graduates, undergraduates and preparatory students. The study shows that there is a significant positive relationship between the student learning culture and IT use on student knowledge sharing. The study limitations, practical implications, along with directions for further research are discussed.

Key words: knowledge sharing, learning culture, IT use, knowledge management

1. INTRODUCTION

Knowledge is recognized not only as the most important resource in organizations (Liao et al. 2004) but as one of the primary sources of competitive advantage (Stewart 1996). Knowledge is critical to the long term sustainability and success of any organization (Nonaka and Takeuchi 1995). The importance of knowledge is eminent to both public and private organizations, particularly educational/learning institutions such as universities (John 2001). The leverage of knowledge is a managerial and strategic issue. As reservoir of knowledge, higher learning institutions are no longer just providing knowledge to students, but also managing and blending together as well as sharing such knowledge among the students. Thus, knowledge sharing is inevitably challenging and important concept in higher learning institutions.

Knowledge creation, sharing and dissemination are the main activities in knowledge management. Being part of knowledge management (KM) process (Kim and King 2004), Knowledge sharing (KS) is the exchange of experience, events, thoughts or understanding of anything. In general, people expectations from knowledge sharing are to gain more insights and understanding about concepts or practical applications, thereby improving learning and expertise. Thus, knowledge sharing can be considered as a significant ingredient for mutual learning and intellectual development to students.

To stay competitive in the education industry, institutional members must promote knowledge sharing (Kumar 2005). Conversely, the competitive nature of learning institutions, among others, may hinder the knowledge sharing among the students. Thus, studies on both barriers and enablers of knowledge sharing in learning institutions may be relevant not only for adding to the literature, but to the policy makers of such institutions. Literature search shows that limited studies where available on the factors that promote and limit knowledge sharing among students.

With the increasing investment of Information Technology (IT) in educational institutions, one expects such investments to have positive impact on the way knowledge is disseminated. Possibly, the level as well as the use of Information Technology (IT) may support knowledge sharing capabilities (Ipe 2003) in institutions of learning. However, investment in technology may not be the only factor that could enable
knowledge sharing. Other factors, social and cultural, in particular are worth considering. In this respect, support of one another, learning culture, might in a collectivistic society, like Saudi Arabia, promotes the willingness of students to share knowledge among themselves (Maccoby 2003).

Literature search shows that limited studies where available on the factors that promote and limit knowledge sharing among students. Similarly, the literature search revealed that, to date, no such study was undertaken in Saudi Arabia. Hence, the focus of this study is to examine the relationship of some dimensions of learning culture and Information Technology (IT) on students’ knowledge sharing.

This study is unique and original, not only for being one of the first in Saudi Arabia, but for being one of the few that examines how both technological and social and cultural factors together serve as antecedents of knowledge sharing. Next, the research objectives and significance of the study will be presented. Then, section two examines the relevant literature. Section three discusses the research methodology. Section four provides findings and implications. Finally, section five concludes the paper and recommends future research directions.

1.1 Objectives of the Study

In a bid to assess the impact of learning culture and use of information technology on students’ knowledge sharing in KFUPM, the study specifically aims at:

1. Determining the impact of learning culture on the students’ knowledge sharing.
2. Evaluating the role of IT on students’ knowledge sharing.
3. Assessing the level of knowledge sharing and learning culture of the KFUPM students and
4. Assessing whether knowledge sharing and learning culture differ among demographics of KFUPM students.

1.2 Significance of the Study

Literature search revealed that, to date, no research has been conducted on the role of learning culture and use of IT in promoting knowledge sharing in the Kingdom of Saudi Arabia. Thus, the result of this study may be relevant to various stakeholders in the Kingdom. To the government authorities of education in the Kingdom, the understanding of how knowledge is been shared (Brown and Duguid 2000) is important in attaining the government strategic plan of transforming the country to knowledge-based economy. Hence, the findings of the study with respect to the level of the knowledge sharing will be of great significance to this end.

The management of KFUPM may also find the results of the study of immense practical benefits because there is a need, in the first place, to know those factors that impact knowledge sharing among students before embarking on any strategy and program of supporting knowledge management in the university. From an academic perspective, this study's insights will add to the existing literature on the impact of learning culture and IT use on knowledge sharing in general; and in Saudi Arabia in particular where none exist. Therefore, the study is of significant value to practitioners and scholars alike.

2. LITERATURE REVIEW

2.1 Knowledge Sharing

Organizational success in today's dynamic and fiercely competitive environment depends largely on the ability to leverage knowledge to develop competitive capabilities to aid in developing new products, services and processes that outperform those of rivals (Kogut and Udo 1992, Nickerson and Zenger 2004,
Szulanski 1996). Knowledge is regarded as a fluid mix of framed experiences, values, contextual information, and expert insights that provide a framework for evaluating and incorporating new experiences and information (Davenport 1997). Many other definitions are abound (Davenport 1997). With respect to categorization of knowledge, there is no consensus in this regard. Researchers have identified different types of knowledge (Nonaka and Takeuchi 1995). The most common classification, however, is between explicit and tacit knowledge (Nonaka 1994).

While it is easy to transmit explicit knowledge through formal language, it is much difficult on the other hand to convey tacit knowledge (Nonaka and Takeuchi 1995). This is because explicit knowledge can be made readily available in the form of files, library collections, or databases (Nonaka and Takeuchi 1995). However, tacit knowledge is difficult to express in words or to codify in documentation. It mainly resides inside individuals’ brains (Hlupic and Rzevski 2002). It is the personal knowledge that is embedded in individual members and used by them in performing their work (Argote and Paul 2000).

It should be noted that knowledge is not an end in itself, but rather means to an end. Thus, only by harnessing and exploiting the collective wisdom and knowledge of their members can organizations adapt and develop innovative processes, products, tactics and strategies (Alavi et al. 2005/2006, Maccoby 2003). A technique widely championed by organizations in order to achieve this end is knowledge management. Knowledge management demands that knowledge should be obtained, produced, shared, regulated and leveraged by a steady conglomeration of individuals, processes and IT (Benbya and Belbaly 2005). Knowledge sharing, as a dimension of knowledge management, is in turn defined as the provision or receipt of task information, know-how and feedback regarding a product or procedure (Cummings 2004).

Knowledge sharing can also be seen as a social interaction culture, involving the exchange of employees’ knowledge, experiences, and skills through the whole department or organization. Knowledge sharing comprises a set of shared understandings related to providing employees access to relevant information and building and using knowledge networks within organizations (Hogel et al. 2003). It is the voluntary dissemination of acquired skills and experience to the rest of the organization (Ipe 2003).

At the individual level, knowledge sharing is referred to as the talking to colleagues to help one get something done better, more quickly, or more efficiently (Lin 2007). Sharing of knowledge at the individual level is the most critical to an organization, even though it may exist at other levels of an organization that include team and organizational levels (Lukas et al. 1996). Individuals can realize synergistic results greater than those achievable individually by sharing their knowledge (Cordoba and Isabel 2004).

Moreover, knowledge sharing occurs not only at the individual level, but at the organizational level as well (Lin 2007). For an organization, knowledge sharing is capturing, organizing, and transferring experience-based knowledge that resides within the organization and making that knowledge available to others in the organization (Lin 2007). A firm can successfully achieve promotion of knowledge sharing culture not only by directly incorporating knowledge in its business strategy, but also by changing employee attitudes and behaviors to promote willingness and consistent knowledge sharing (Connelly and Kelloway 2003, Lin and Lee 2004).

There are several antecedences, organizational and otherwise, to knowledge sharing. These factors include the organizational structure, organizational culture, leadership, information systems (Davenport and Prusak 1998, Bock, et al. 2005, Ardichvili et al. 2006), avoidance of embarrassment (Burgess 2005), obligation, trust, and identification (Faraj and Wasko 2002), individual ability (e.g. subject expertise, tenure) (Wasko and Faraj 2005), greed, self-efficacy (Lu et al. 2006), extrinsic rewards, fear of punishment (Burgess 2005), expected rewards, expected associations, expected contribution (Bock and Kim 2002), perceived costs, extrinsic benefits, intrinsic benefits (Kankanhalli et al., 2005), anticipated extrinsic rewards, anticipated reciprocal relationships, sense of self-worth (Bock et al. 2005) among
others. The outcome of knowledge sharing that includes productivity, task completion time, organizational learning (Argote 1999, Argote et al. 2000, Cummings 2004), enhancing innovation performance and reducing redundant learning efforts (Scarborough 2003) have been examined by a number of studies. Conversely, the absence of knowledge sharing is likely to undermine knowledge management efforts (Calantone et al. 2002).

### 2.2 Learning Culture

Both private and public organizations are increasingly recognizing the importance of culture as an essential prerequisite for readiness and willingness to learn (Calantone 2002). Simply, culture can be referred to as a system of shared values and assumptions. It influences employee interaction, organizational functioning, and even decision making in organizational settings (Lukas et al. 1996). Culture is of great importance to the learning organizations including universities (Carleton 1997). It influences or inhibits, directly, the quality of learning (Szulanski 1996), which is of utmost concern in institutions of higher education.

According to Johnston and Hawke (2002), learning culture can be defined as the existence of a set of attitudes, values and practices within an organization which support and encourage a continuing process of learning for the organization and/or its members. A learning culture is said to exist in an environment where teamwork, collaboration, creativity, and knowledge processes exist that have a collective meaning and value (Joo 2007). For an organization to improve its performance, it requires a learning culture (Kumar 2005). Hence, development of learning culture is becoming a dominant theme in the strategic plans of many organizations today (Walsham 2002).

Developing a learning culture could help in gathering, organizing, sharing, and analyzing the knowledge of individuals and groups across an institution in ways that directly affect its performance (Kumar 2005). Learning culture benefits a whole organization and certain teams within the organization and it is essential in moving an organization to a learning one (Cohen 1990) which usually support knowledge sharing. Researches that focused on factors affecting knowledge sharing have identified the relevance of learning culture, among other variables (John 2001). In higher educational institutions, in particular, learning culture is needed for the institutions to create and disseminate knowledge that is necessary for the development of such institutions. Development of such learning culture in learning institutions may also create opportunities in accessing and sharing the right knowledge at the right time and in the right location to stay competitive in the global educational environment (Kumar 2005). Hence, it can be hypothesized that:

**H1:** The level of learning culture has positive effect on students’ knowledge-sharing behavior.

### 2.3 Information Technology

The role of the information technology (IT) in sharing knowledge has been a center of debate (Maccoby 2003). While some investigators are of the opinion that knowledge management (KM) initiatives could be successful without using IT tools (Mohamed 2006, Hislop 2002), other researchers have, however, identified IT as a variable that could impact knowledge sharing (Huysman and Wulf 2006) for the fact that technology is one of the important pillars of knowledge management (Maccoby 2003). Haldin-Herrgard (2000) maintained that a great deal can be done through modern IT to diffuse explicit knowledge. It is also becoming easier nowadays to capture tacit knowledge with the aid of retrieval technologies (Kumar 2005).

A study by Pai (2006) that examined the relationship between the effectiveness of IS strategic planning (ISSP) and knowledge sharing found that top management support for ISSP has a strong significant effect
on knowledge sharing behavior. A separate study in South Korea by Kim and Lee (1996) also found, among others, that both employees’ usage of IT applications and friendliness of the IT systems significantly impact employee knowledge-sharing capabilities. It can, therefore, be expected that individuals with more usage and favorable perception of IT may demonstrate more knowledge sharing behavior (Kumar 2005). Hence, it can be hypothesized that:

**H2:** The level of students’ utilization of IT has a positive effect on students’ knowledge-sharing behavior.

Based on the study objectives which were substantiated by the reviewed literature, the hypotheses of the study are eventually developed. These hypotheses are translated into the theoretical model depicted in Figure 1. The model presumes that learning culture and IT use have an impact on students’ knowledge sharing in KFUPM.

### 3. METHODOLOGY

#### 3.1 Data Collection and sample

In a bid to examine the impact of learning culture and IT use on students’ knowledge sharing in KFUPM, the study undertook a survey questionnaire method. The survey instrument reflected the conceptual framework depicted in Figure 1.

![Conceptual Model](image)

*Figure 1. Conceptual Model*

To ensure generalization of the study findings, the questionnaires were administered based on stratified random sampling to KFUPM students. A total of 200 surveys were hand-delivered to the graduates, undergraduates and preparatory students from all the academic departments in the University. A total of 142 questionnaires were returned; of which five incompletes were discarded. The final number of usable questionnaires was 137, representing 68.5% response rate.

#### 3.2 Measures

All the constructs of the study were measured from items adapted from previous studies, with some alterations to account for the peculiarity and setting of the study. To improve the reliability and validity, multiple-item measures were used for all of the variables. Responses were recorded along a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) for all the items in the questionnaire. Learning culture was measured using 6 items adopted from Cordoba and Isabel (2004). The six items were meant to be grouped into meaningful cluster(s) by conducting factor analysis. Students’ IT usage was measured using 6 items from Kim and Lee (1996) and 1 item specifically designed for this study was added. Finally 6 items were adapted from Kim and Lee (1996) as a measure of students’ knowledge-sharing.
4. STUDY FINDINGS AND IMPLICATIONS

4.1 Profile of the Respondents

The survey questionnaires were distributed to each of the three categories of students, graduates, undergraduates, and preparatory students. Preparatory students are those students that are undergoing one year course toward starting their main bachelor degree. Undergraduates are those that are pursuing their four year bachelor degree courses. Both the preparatory and the undergraduate students are from different regions of the country, which makes the sample relatively homogeneous. On the other hand, graduate students are those students pursuing their postgraduate programs that include MBA, MS and PhD. The unique characteristic of the graduate students of the university is that they include students from different countries, not only from Saudi. The cosmopolitan nature of this category of the students may reflect the perception of students across different societies, hence enabling the study findings to be more generalized.

It can be inferred from Table 1, that 30.15% of the respondents were graduate students, 63.97% undergraduates, and the remaining 5.88% preparatory students. With respect to the GPA, it can be seen that only about 15% of the respondents have GPA of more than 3.75 and 16% with GPA of 3.51 - 3.75. On the other hand, over 50% of the respondents’ GPA is between 3.00 - 3.25 and less than 3.00. This seems to be quite representative of the population.

<table>
<thead>
<tr>
<th>Categorization</th>
<th>No of Valid Respondents</th>
<th>Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduate student</td>
<td>41</td>
<td>30.15</td>
<td>30.15</td>
</tr>
<tr>
<td>Undergraduate students</td>
<td>87</td>
<td>63.97</td>
<td>94.12</td>
</tr>
<tr>
<td>Preparatory student</td>
<td>8</td>
<td>5.88</td>
<td>100.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GPA</th>
<th>No of Respondents</th>
<th>Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.76 – 4.00</td>
<td>13</td>
<td>14.44</td>
<td>14.44</td>
</tr>
<tr>
<td>3.51 – 3.75</td>
<td>16</td>
<td>17.78</td>
<td>32.22</td>
</tr>
<tr>
<td>3.26 – 3.50</td>
<td>15</td>
<td>16.67</td>
<td>48.89</td>
</tr>
<tr>
<td>3.00 – 3.25</td>
<td>26</td>
<td>28.89</td>
<td>77.78</td>
</tr>
<tr>
<td>Less than 3</td>
<td>20</td>
<td>22.22</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Table 1. Respondents’ Profile

4.2 Reliability and Validity of Variables

Several major steps were carried out to enhance the reliability and validity of the variables. Factor analysis was first made for this purpose. Reliability was then utilized to check the internal consistency of the scales involved in the study. Coefficient alpha analysis was used to determine the extent to which items making up each measure were homogenous and loaded on the same scale (Allen and Yen 1979). Cronbach’s alpha has been suggested to be the preferable measure of index reliability. The scales used in this study were checked for their internal consistency.

Content Validity which determines the adequacy of the sample characteristics in describing the study measures (Nunnally 1978) has been established in the study. This is because the questionnaire used in this study built upon existing research where the scale items were found to be valid. For the construct validity, one technique widely used to assess such validity of an instrument is factor analysis (Kerlinger 1973). Various items that represent each dimension were analyzed to see if they are properly assigned to the appropriated scale (Carmines and Zeller 1980). Two criteria were used to identity the factor scales. First, all scale items that loaded less than 0.40 were removed. Second, a construct with the highest eigen value
would be represented by a factor. For the learning culture, all the items were found to be correlated with the factorial groups produced with the factor loading more than 0.40. Out of the 6 items, two factors emerged as can be seen in Table 2. The correlation result between these factors shows that the two factors are significantly correlated at 99% level of confidence. Thus, one factor was selected that represent the learning culture construct and named learning culture in KFUPM. This factor is the one with the highest eigen value of 2.844 and the percentage variance explained of 47.403 as depicted in Table 2.

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. KFUPM has a great deal of personal meaning to me</td>
<td>0.316</td>
<td></td>
</tr>
<tr>
<td>2. I owe a great deal to KFUPM</td>
<td>0.888</td>
<td></td>
</tr>
<tr>
<td>3. I am willing to work with one another in KFUPM</td>
<td>0.673</td>
<td></td>
</tr>
<tr>
<td>4. I am willing to ask others for assistance (e.g., information, etc.) in KFUPM</td>
<td>0.613</td>
<td></td>
</tr>
<tr>
<td>5. I feel safe sharing information/knowledge with my colleagues</td>
<td></td>
<td>0.379</td>
</tr>
<tr>
<td>6. I can trust my colleagues</td>
<td></td>
<td>0.812</td>
</tr>
<tr>
<td><strong>Eigenv Value</strong></td>
<td><strong>2.844</strong></td>
<td><strong>1.159</strong></td>
</tr>
<tr>
<td>Variance explained (%)</td>
<td><strong>47.403</strong></td>
<td><strong>19.116</strong></td>
</tr>
</tbody>
</table>

Table 2. Correlation of Learning Culture Factors

From Table 3, it is evident that the Cronbach’s alpha of the construct, learning culture in KFUM is 0.777. Since according to the guideline of Nunnally and Bernstein (1994) the value of 0.7 or above is an acceptable reliability coefficient, hence the construct has exhibited adequate reliability. With respect to the Students’ IT usage, the result of the factor analysis indicates that one item has less than 0.40 factor loading. The item was hence dropped. The remaining five items eventually yielded a single factor with 0.678 Cronbach’s alpha. Though the Cronbach’s alpha not up to 0.70, the factor could still be considered reliable since it very close to .70 (Koch et al. 2005, Graham and Nafukho 2007, Chang and Lin 2007). Some of the items used for measuring IT use are those provided by the university (KFUPM) such as the WebCT, and other systems that are specifically designed by the university authority to promote learning. Other IT items used by students for knowledge sharing and learning include the email tools available in KFUPM Intranet and the Internet. Similar to IT usage, one item of the knowledge sharing has less than 0.40 factor loading value. The item was dropped and the remaining 5 items yielded one factor. The Cronbach’s alpha reliability for these five items was 0.744, which can be considered quite reliable (Nunnally 1978).

4.3 Study Findings

Recall that the first and the second objectives of the study are to determine the impact of each of the learning culture and IT use on students’ knowledge sharing respectively. Since the appropriate analysis that assesses the influence of independent variable(s) on dependence variable is regression analysis, two simple regression analyses were run to test the two proposed hypotheses. The choice of the simple regression analysis in assessing the impact of both learning culture and IT use on the dependent variable is that each of the two variables has a single dimension. Moreover, The R² of both learning culture and IT use models explains knowledge sharing behavior of students in KFUPM. From the results of the regression in Table 4, it can be deduced that the learning culture in KFUPM, which has a positive
coefficient Beta, as hypothesized has significant relationship with students’ knowledge sharing. Hence, the first hypothesis is fully supported that learning culture has positive impact on the students’ knowledge sharing in KFUPM.

Table 3. Significant Factor Analysis and Reliability Analysis Results

<table>
<thead>
<tr>
<th>Constructs and Items</th>
<th>Standard Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Learning Culture in KFUPM Construct Items</strong></td>
<td></td>
</tr>
<tr>
<td>I owe a great deal to KFUPM</td>
<td>0.388</td>
</tr>
<tr>
<td>I am willing to work with one another in KFUPM</td>
<td>0.673</td>
</tr>
<tr>
<td>I am willing to ask others for assistance (e.g., information, etc.) in KFUPM</td>
<td>0.613</td>
</tr>
<tr>
<td>KFUPM has a great deal of personal meaning to me</td>
<td>0.816</td>
</tr>
<tr>
<td><strong>Information Technology (IT) Usage Construct Items</strong></td>
<td></td>
</tr>
<tr>
<td>I regularly use the Internet, e-mail, and electronic bulletin boards</td>
<td>0.633</td>
</tr>
<tr>
<td>I regularly use the KFUPM’s intranet</td>
<td>0.609</td>
</tr>
<tr>
<td>I regularly use the KFUPM’s content delivery system (e.g., Webct etc.)</td>
<td>0.702</td>
</tr>
<tr>
<td>In this university, information systems and software are designed to be user-friendly</td>
<td>0.703</td>
</tr>
<tr>
<td>It is easy for me to use the university’s information systems without extra training</td>
<td>0.664</td>
</tr>
<tr>
<td><strong>Knowledge-Sharing Construct Items</strong></td>
<td></td>
</tr>
<tr>
<td>I voluntarily share my know-how, information, and knowledge with other students</td>
<td>0.766</td>
</tr>
<tr>
<td>I freely share information and knowledge that will improve the academic performance of fellow students</td>
<td>0.809</td>
</tr>
<tr>
<td>I cooperate or communicate with other students in teams or groups for sharing information and knowledge</td>
<td>0.764</td>
</tr>
<tr>
<td>I discuss my academic problems with other students rather than struggling with the problems individually</td>
<td>0.541</td>
</tr>
<tr>
<td>I share information and knowledge in the class/group if I know the information/knowledge will help in understanding of other members of the class/group</td>
<td>0.611</td>
</tr>
<tr>
<td><strong>Reliability Analysis</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Construct</strong></td>
<td><strong>Beta</strong></td>
</tr>
<tr>
<td>Learning Culture in KFUPM</td>
<td>0.352</td>
</tr>
<tr>
<td>Information Technology</td>
<td>0.388</td>
</tr>
</tbody>
</table>

Table 4. Summary of Regression Analysis

Similar finding was achieved from the regression analysis result in same Table 4 of the positive influence of IT use on the students’ knowledge sharing in KFUPM. The second hypothesis is equally supported. Furthermore, the $R^2$ of the two regression models, which are 0.487 and 0.151, indicate that 48.7% and 15.1% of the overall student knowledge sharing is explained by the learning culture and IT use independent variables respectively.

Table 4. Summary of Regression Analysis

<table>
<thead>
<tr>
<th>Construct</th>
<th>Beta</th>
<th>t</th>
<th>Sig.</th>
<th>R²</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Culture in KFUPM</td>
<td>0.352</td>
<td>4.610</td>
<td>0.000</td>
<td>0.487</td>
<td>41.61</td>
<td>0.00</td>
</tr>
<tr>
<td>Information Technology</td>
<td>0.388</td>
<td>4.897</td>
<td>0.000</td>
<td>0.151</td>
<td>23.98</td>
<td>0.00</td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), Learning Culture in KFUPM, Information Technology
b. Dependent Variable: Knowledge Sharing
Furthermore, the result of regression in Table 5 shows which Information Technologies items more used for knowledge sharing. It can be inferred that the coefficient (beta value) of the item "It is easy for me to use the university's information systems without extra training" is higher showing this IT items contributes more to knowledge sharing and the least item that impact knowledge sharing is "I regularly use the KFUPM's databases (e.g., library e-database, etc)".

Table 5.  Summary of Regression Analysis

To answer the research question of the level of knowledge sharing and the learning culture of the students, one-sample t test analysis was undertaken. Since the scales for knowledge sharing and learning culture are all on 5 point Likert scales, the test value of 3 was used and the result in Table 6 shows that all the means are significantly different from neutral value 3. These findings indicate that the students are more inclined toward satisfaction than dissatisfaction with both the knowledge sharing and the learning culture.

Table 6.  One-Sample Test of Level of Knowledge Sharing and Learning Culture

The last research objective aims at examining how the learning culture and the knowledge sharing of the students differ according to the demographics. The results of the ANOVA in Table 7 below report such relationships. The results indicate that the level of the knowledge sharing among students is not significantly different with respect to both the students’ GPA and their category because the p > 0.05. This implies that the students share knowledge irrespective of their GPA or their category. However, with respect to the learning culture in KFUPM, ANOVA results in Table 6 show that the level of learning culture in KFUPM is significantly different among students with different class of GPA, though no significance difference exists between different categories of students.

<table>
<thead>
<tr>
<th>IT Items</th>
<th>Beta</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I regularly use the Internet, e-mail, and electronic bulletin boards.</td>
<td>.267</td>
<td>2.907</td>
<td>.004</td>
</tr>
<tr>
<td>I regularly use the KFUPM's internet.</td>
<td>-.112</td>
<td>-1.151</td>
<td>.232</td>
</tr>
<tr>
<td>I regularly use the KFUPM's databases (e.g., library e-database, etc).</td>
<td>.200</td>
<td>.219</td>
<td>.827</td>
</tr>
<tr>
<td>In this university, information systems and software are designed to be user-friendly.</td>
<td>.137</td>
<td>1.431</td>
<td>.155</td>
</tr>
<tr>
<td>It is easy for me to use the university's information systems without extra training.</td>
<td>.286</td>
<td>2.950</td>
<td>.004</td>
</tr>
</tbody>
</table>

a. Predictors: Information Technology Items  b. Dependent Variable: Knowledge Sharing

Test Value = 3

<table>
<thead>
<tr>
<th></th>
<th>Means</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge Sharing</td>
<td>3.76</td>
<td>13.53</td>
<td>136</td>
<td>0.00</td>
</tr>
<tr>
<td>Learning Culture in KFUPM</td>
<td>3.81</td>
<td>11.92</td>
<td>136</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note: Mean scores based on a five-point scale ranging from 1 = Strongly Disagree to 5 = Strongly Disagree.

Table 6.  One-Sample Test of Level of Knowledge Sharing and Learning Culture

ANOVA Summary Statistics

<table>
<thead>
<tr>
<th>Categorization</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge Sharing</td>
<td>2.137</td>
<td>0.122</td>
</tr>
<tr>
<td>Learning culture in KFUPM</td>
<td>2.85</td>
<td>0.06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GPA</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge Sharing</td>
<td>0.417</td>
<td>0.796</td>
</tr>
<tr>
<td>Learning culture in KFUPM</td>
<td>2.75</td>
<td>0.03</td>
</tr>
</tbody>
</table>
Notes: F-values are the result of a one-way ANOVA at $p < 0.05$ level of significance.

Table 7. ANOVA of Categorization and GPA with the Study Constructs

4.4 Discussions and Implications

The findings of the study of the impact of both learning culture and IT use on knowledge sharing among students has important implications to the management of KFUPM as well as other sister institutions. The reason is that because in their effort to enhance knowledge sharing which is inevitable in promoting learning (Graham and Nafukho 2007, Haldin-Herrgard 2000) the management of KFUPM and other higher institutions should opt to at least instill the culture of learning as well as invest more in IT and ensure more usage of IT by their students.

The finding of learning culture as a contributor to knowledge sharing shows that knowledge sharing factors do not depend solely on technology. In fact, the study finding that learning culture accounts by about 48.7% to knowledge sharing against IT use that accounts by only 15.1% is of practical implication and implies that effort toward promoting learning culture may be more viable in promoting sense of knowledge sharing between the students. To achieve this, there is need to organize some orientation courses, seminars among others to instill the etiquette of positive learning culture in the students. It can be envisaged that it may be wise to consider integrating learning culture in the course syllabi design. In the same vein, the finding with respect to the demographics that level of the knowledge sharing of the students is the same irrespective of their GPA and their category implies that, when designing a knowledge sharing program it can be standardized to all students in the university.

5. CONCLUSIONS AND FUTURE RESEARCH

Based on the study results and discussion, it could be concluded that the usage of IT and learning culture are significant variables that affect student knowledge sharing in higher institutions, and KFUPM in particular. The presence of learning culture as the major contributor shows that knowledge sharing factors do not depend on technology alone. It is suggested that in a bid to improve student knowledge sharing, the appropriate authorities and decision makers need to commit efforts and programs that could enhance learning culture and IT usage among students. In spite of the originality of this study for being one of the few that examined the influence of social and technological factors together, learning culture and IT use, on student knowledge sharing in the Kingdom of Saudi Arabia, it suffers from some limitations that are common to many other researches.

One of the chief limitations is that the sample size is relatively small and needs to be increased. To account for the sample size limitations of the study, further studies can consider respondents from several universities, possibly including female respondents, to ensure more generality of the findings. In addition to the learning culture, further studies can expand this study to include other factors that may impact students’ knowledge sharing. Also, the impact of IT on collaborative learning is another area where further research could be viable.

References


# Transfer of Technology and Knowledge – The Story of an Enterprise System Implementation

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0300.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Actor-network theory, Information Systems Development (ISD), knowledge transfer, Socio-technical theory</td>
</tr>
</tbody>
</table>
THE TRANSFER OF TECHNOLOGY AND KNOWLEDGE: A STORY OF AN ENTERPRISE SYSTEM IMPLEMENTATION

Abdul Aziz Ahmad
Manchester Business School, University of Manchester, Manchester, UK

&

Mike Newman
Manchester Business School, University of Manchester, Manchester, UK; Norwegian School of Economics and Business Administration, Bergen, Norway

Abstract

The uncertainties of information system development (ISD) provide many routes that end in poor project outcomes. These uncertainties produce opportunities for IS researchers who are continually exploring the complexities of information system development. In this paper, which is a work-in-progress, we are focusing on the implementation of enterprise systems in three universities in Malaysia.

We adopt a socio-technical process model which provides us with a vehicle to describe an in-depth analysis of the phenomena under examination. By examining the notions of technology and knowledge transfer in ISD, this paper explores the intricate relationships and interactions between its human and non-human actors in the attempts to construct various network alliances. We end with some suggestions for research and practice.

Keywords: information system development (ISD), universities, transfer of technology and knowledge, socio-technical modeling, actor-network theory (ANT)
1 INTRODUCTION

Researchers in enterprise system implementation in universities show indicates that one of the main motives for their implementation is to replace legacy systems (Wagner et al. 2006; Pollock & Cornford 2004; Wagner & Newell 2004; Scott & Wagner 2003; King et al. 2002; Oliver & Romm 2002; Allen & Kern 2001). The systems which were once the backbone of the organization have with time become obsolete. Organizations today require more robust contemporary systems that are able to cope with ever changing and intensifying market competition. Replacing a legacy system is not just a matter of changing components. It may also require a complete overhaul of the organization itself. That is why it is common to see failures in enterprise system implementation in universities due to the complexities of the specifications, functional inadequacies and lack of political support (Pollock & Cornford 2004; Wagner & Newell 2004; Scott & Wagner 2003). However, not all project failures are due to organizational complexities. Understanding organizational compatibility with project managers and vendors plays a vital role in ensuring smooth project implementation.

In general, the data analyzed in this paper was derived from three Malaysian universities where each of these organisations had technological and personnel connections with the others. While the discussion will concentrate on the enterprise system implementation project in one of the universities, some of the discussion will apply to the other two. Our study was guided by the following research questions:

1. What are the key actor networks in the development process and how do the networks develop and interact diachronically?
2. What are the implications for research and practice resulting from our findings, particularly for similar institutions?

The structure of the paper is as follows. Part two provides a summary of the main features of actor-network theory. In part three, an overview of the research methodology which includes the case description, data collection method and data analysis is presented. This is followed by the presentation of the findings in part four. Part five revisits the research questions in an attempt to better understand the phenomena. The paper ends with conclusions that have emerged from the study.

2 PREVIOUS RESEARCH

2.1 Actor-network Theory (ANT) and the process of translation

According to Latour (1993) the world is full of hybrid entities which contain human and non-human elements (Tatnall 2003). The development of these hybrid entities is facilitated by the heterogeneous nature within the information system environment. ANT is unique as it is engaged in both sides: ANT attempts to discuss the heterogeneity of elements in forming a network while also examining the asymmetry of its elements. In relation to generalized asymmetry, Latour (1993, p. 95) has described it as a rationale of symmetric treatment in observation on the relationship between technical computer systems and social organizations where all human and non-human elements are treated equally and are labelled as actors or actants.

---

1 In this paper, implementation refers to an on-going process which includes the entire development of the systems from original suggestion through to the installation of the system (Lucas 1981) and development refers to a stage within the implementation process.
The next concept we will use is the actor-network. This is considered as the most crucial part of ANT and looks at the process of building a network of actors. This process involves several main activities including interactions, negotiations and translations. In our research we will focus on the process of translation and in this paper we will concentrate on Callon’s four phases of the translation process (Callon 1986). The first phase is the problematization. In this phase an issue or a problem will be identified and the role of each actors affected by the issue is defined. In this phase, some actors will establish themselves as indispensable to the solution of the problem (McMaster et al. 1997). They consider themselves to be the ‘obligatory passage point’ (OPP).

This process of translation is continued with the phase of interessement. According to Cordella and Shaikh (2003), translation is a circular process of interpretation of interest. Interessement is a group of actions by which an entity attempts to impose and stabilize the identity of the other actors it defines through its problematization (Callon, 1986). The moment that an actor accepts the interest defined by another actor is called enrolment (Callon, 1986). This point of enrolment can only be observed through time and it is crucial for a researcher to identify each of the different stages or phases of network building. According to Hanseth et al. (2004), during this process, programs-of-actions are inscribed where roles, competencies and responsibilities are delegated between the actors within the network. This is further elaborated through the concept of circularity of the translation process which, according to Cordella and Shaikh (2003), ensures that the interest of all actors within the future network is properly inscribed within the network. According to Latour (1992), inscription (the fourth phase) is a process of creating technical artefacts that ensures the protection of an actor’s interests (Sarker et al. 2006). In layman’s terms, this process can be considered as the signing of a memorandum of understanding (MoU) that binds the party to a specific objective or in ANT terms, interest. For example, a centralized enterprise system can be seen as an inscription of the interests of management who commissioned the project. However the system can become subverted to the interests of other groups in the process of its development.

3 RESEARCH METHOD

3.1 Case descriptions

This study takes place in three universities in Malaysia that are similar, yet possess unique features. It is their similarities and uniqueness that have made this research challenging and produced a number of surprises.

![Figure 1 Inter-organization transfer of technology and knowledge](image)

*Figure 1 Inter-organization transfer of technology and knowledge*
Case 1

This is the youngest of the three universities and was established in 2002 when they implemented the integrated management system (IMS). With a current enrolment of 3,000 students and 1,000 staff, the need to maintain and further improve the IMS was essential and ongoing. The implementation of the enterprise system was an integral part of the university’s strategic planning and was initiated during the university’s infancy. This has ensured the smooth internalization of an electronic working culture. The head of ICT in Case 1 was later employed as the project manager in Case 3.

Case 2

Established in 1922, this is the oldest education institution in Malaysia and under the mandate of the university act, it became a full university in 1997. The problems that arose from the islands of legacy systems created the need for a more integrated information system. Today, this dynamic and mature university comprises 16,000 students in addition to a compliment of 580 faculty and 764 management staff. The experiences and the expertise over the years have produced more stable and fully tested business processes. The implementation of such a large enterprise systems in this culturally-developed organization was completed seamlessly in 2005.

Case 3

This university was also established in 2002, and represents an amalgamation of nine institutes which were previously managed individually. The diverse operational processes among these previously independent institutes brought about a call for an integrated management system (IMS) similar to that adopted in Case 2 to enhance their operational efficiency. With 13,000 students supported by 2,000 de-centralized staff members, the establishment of a single integrated database was thought to produce a more efficient management of resources. This two-year project launched in February 2007 has experienced the greatest number of problems encountered in any of the above noted IS development projects. For our study we will focus on one component of an integrated management system, an integrated finance module.

While the above description has emphasized the unique nature of each individual case, the one element that is common among all three cases is the ‘system’ (Figure 1) which has been developed on the same base system (V0 above). Depending on the timing of their implementation, any new system enhancements or modifications are integrated into the latest version. This refers to V1 and V2 in Figure 1. The system for Case 2 was developed based on Case 1’s latest version. The system for Case 3 was built upon the latest version of Case 2. Whilst Case 2 is more mature in terms of its operation compared with Case 1, the base system for Case 3 was deemed to be more up-to-date and more comprehensive. However, the project manager in Case 3 resigned in August 2008, just seven months after her appointment. The project is currently being managed by the information technology department (ITD) of Case 3. As of the last visit management has no intention of appointing a new project manager.

3.2 Data collection

<table>
<thead>
<tr>
<th>Interviewee</th>
<th>Min.</th>
<th>Interviewee</th>
<th>Min.</th>
<th>Interviewee</th>
<th>Min.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td></td>
<td>Case 2</td>
<td></td>
<td>Case 3</td>
<td></td>
</tr>
<tr>
<td>Bursar</td>
<td>50</td>
<td>Bursar</td>
<td>NA</td>
<td>Senior General Manager (Fin.)</td>
<td>70</td>
</tr>
<tr>
<td>Finance Officer 1</td>
<td>44</td>
<td>Deputy Bursar</td>
<td>86</td>
<td>Head of Unit – Finance (HQ)</td>
<td>44</td>
</tr>
<tr>
<td>Finance Officer 2</td>
<td>75</td>
<td>Head of Unit</td>
<td>NA</td>
<td>Head of Unit – Finance (Branch)</td>
<td>70</td>
</tr>
<tr>
<td>Head of ICT (ex)</td>
<td>59</td>
<td>System analyst (ICT)</td>
<td>53</td>
<td>Software Developer</td>
<td>72</td>
</tr>
<tr>
<td>Vice Chancellor</td>
<td>30</td>
<td>Software Developer</td>
<td>72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software Developer</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. The interview schedule – Cases 1, 2 and 3

Proceedings ECIS 2009
Table 1 represents the list of interviews conducted during the field work in July 2008 but during the second round of interviews in March 2009, we have managed to gain responses from other actors related to the cases especially with the ex-project manager of Case 3.

The research will attempt to analyze the socio-technical relationship during the process of ERP system implementation. Therefore, the interviews were conducted with as many ERP system stakeholders as possible in order to gain an in-depth understanding of the process. Interviews were tape-recorded and verbatim transcripts were produced.

### 3.3 Data analysis

Since this paper is a work-in-progress, the analyses of the data are only based on the available transcripts to date. Further interviews and data gathering will be undertaken in due course over the next year. The analysis began with developing a project timeline analysis for all the three cases. Initial analysis of the project timelines shows interesting yet diverse issues between the three cases. Upon reading and re-reading of the interview transcripts, a matrix of project critical events together with direct quotations was produced. During this process it became apparent that the relationship between the cases was also significant in regard to the personnel involved (their knowledge) and the vendor and its technology (the base system). Hence, this is the focus of the paper.

Next, we focused on identifying critical events that occurred during the relationship. Since the focus of the paper will be on Case 3, attention was paid to the antecedent conditions of the other two cases in relation to the third one. Based on these historical experiences, a more in-depth understanding was achieved. Finally, we applied the notion of translation within ANT to explain the dynamics of the interactions and relations during the enterprise system development. According to Monteiro and Sahay (2000), applying ANT in IS research requires the identification of key actors, interests and scenarios and to follow them through time and this is what we are attempting in this paper.

### 4 FINDINGS

Generally, the analysis of the findings in this paper will be based on the socio-technical process model which links antecedents, event processes and context to outcomes (Lyytinen & Newman, 2008).

#### 4.1 The story of an enterprise system implementation through a socio-technical process account – Antecedent conditions

At the time of writing, Case 3 was already one and a half years into the two-year project period and approximately only thirty percent of the finance system has been developed and deployed and as of March 2009 the project period has lapsed. But still the finance system is only thirty percent developed and used. Why have they not been able to complete the other seventy percent of the system? What went wrong in this project? This paper is not an attempt to solve the problem but to better understand the phenomena of the case under study. The notion of transfer of technology will be used to further examine this case study. By adopting a process study model, this paper will use four sections: antecedent conditions, the process, context and the outcomes.

The complexities of an information systems development are not only due to the technicalities of the system itself but most of the time are also due to the intricate relationship between its socio-technical elements. In this paper, the understanding of the antecedent condition of the related actors is crucial. But understanding the history of the organization alone will not be sufficient. We assert that due to the complexities of IS development, an in-depth understanding of the history and background of the vendors (software developers) and the project managers are also crucial to our story.
4.1.1 The university as a unique organization.

The main characteristic of the organization that should be considered in IS development is the establishment of the university which is itself an amalgamation of nine institutes. But as a result of being "nine under one roof", many management problems surfaced. One of the most important issues within the organization was the streamlining of the business processes:

"...the implementation of the finance system is to make sure that the processes are standardized throughout the campuses..."

Senior General Manager (Finance) – Case 3 (17/07/2008, p. 1)

The legacy system of the university did not provide solutions to their problems. With more dispersed campuses and a higher number of students, having a non-integrated system created chaos. Within these campuses also, there were different ways of doing things with very different management styles. In addition, different work styles between the top management and the middle management posed challenges to achieving a successful project.

4.1.2 Vendor (system developer) as one of the important components of IS development

The process of identifying a suitable vendor to supply and deliver an enterprise system is one of the most crucial parts of any project. Usually, the assurance that is given to the clients in providing a complete and usable system develops high expectations among the clients. In this particular case, the assurance or guarantee from the vendor to provide a workable system stems from their previous experiences in the same industry. A successful implementation of prior projects (Cases 1 and 2) provided that confidence. Based on the previous projects, there are several criteria that needed to be fulfilled to ensure project success. The most critical was the transfer of knowledge between the vendor (system developer) and the technical team of the organization. A dedicated technical member attached to a module ensured knowledge was captured and utilized. As the prior client in Case 2 mentioned:

"...the good thing about the project is that our ICT team have efficiently captured the development knowledge that enable us to stand on our own... even our ICT team have understood the debit and credits of accounting transaction and the links in the systems..."

Project Manager (Finance) – Case 2 (08/07/2008, p. 10)

He adds;

"...even after the UAT [User Acceptance Test] there are still minor changes to ensure the completeness of the system and we have to thank the ICT team... they are able to adopt the knowledge and settle some of the problems themselves."

Project Manager (Finance) – Case 2 (08/07/2008, p. 10)

A successful transfer of knowledge ensures an uninterrupted progress of system development. Another important criterion is the user’s capabilities to provide complete and stable business processes. A well-developed business process that is derived from years of operational experiences ensures a smooth development for the vendors. The only development issues would be the integration aspect of the systems which was resolved through continuous coordination and integration meetings and discussions. But in contrast to these positive aspects, the inability of the vendor to comprehend or adapt to different organizational cultures has created some uncertainty towards the current project.

4.1.3 Project manager as the key player in IS development

The project manager often plays a very important role in IS implementation. Due to their complexities, managing such a vast enterprise-wide project requires the multi-tasking capabilities of a project manager. Wide knowledge and experiences in both technical and human aspect of the projects are vital
in ensuring project success. In this case, the 20-years of experience of the project manager in IS implementation gave confidence to the clients. In addition, most of her experiences were based within the educational environment. Prior to Case 3, the project manager handled a similar implementation (i.e. Case 1) as well as a more recent one which was an e-management system. Holding the post of Head of ICT, she initiated the implementation of an enterprise system using a rapid application development (RAD) strategy.

“...so far [an] IT project fails because it takes too long to finish the development.”

Ex Head of ICT – Case 1 and Project Manager - Case 3 (24/07/2008, p. 1)

She further explains:

“...to be successful, this system owner... should understand the nature of system development and rapid approach because they can’t expect the software to be perfect at the 1st time when it is installed. Because of the rapid development we just deploy the 1st draft so that the user can start using it and start improving it based on their actual requirement... the system owner must be aware that they are evolving in the system development, it is part of the software development, it is not like having a complete software, this is a process of software development that the system owner must be involved 100% in the development and improvement of the software...”

Ex Head of ICT – Case 1 and Project Manager - Case 3 (24/07/2008, p.1)

According to her, there are several pre-requisites for a successful RAD implementation. The skills and knowledge together with the commitment of the system developer (vendor) are essential. However, to date, she was not able to replicate her previous success in Case 3 by applying the same concept and strategies acquired in Case 1.

4.2 The story of an enterprise system implementation through a socio-technical process account – The crisis identification

![Figure 2 The 'intention baggage’ or project actors antecedents](image)

Figure 2 provides a summary of the historical background of the projects. Each of the actors identified in the projects have brought their own intention baggage or expectations to the projects. We have

---

2 Because of space considerations we were not able to include the full process model description based on Lyytinen and Newman (2008). This will be explored in a future paper.
already noted that after two years of the development project only thirty percent of the finance system is developed and deployed. What has gone wrong? In the first part of this paper we have discussed the antecedent condition of the situation under study. It is pertinent to understand the history of the case before any further analysis is made. There are several important issues that need to be explored.

The need for such a system was identified due to the complexities of the new organizational structure. The implementation of an enterprise system will create a more streamlined and standardized business processes. The vendor (software developer)’s prior development with other clients provides a basis for an educational system’s best practice. This project tried to build upon the existing best practice and further improve the systems. For the project manager, the rapid application development (RAD) approach, proven to be successful in previous projects, was to be replicated in this project with the same intention to reduce the development time of the project.

While the intention of the clients remains intact, the actors within the project organization have caused the crisis in the IS implementation.

4.2.1 Misunderstanding of project complexities

The adoption of rapid application development (RAD) approach by the project manager together with the ‘adapt and adopt’ method by the vendor have caused chaos within the project development team. The overall project period was two years. It seemed appropriate initially. But when the project manager came in and tried to develop and use the student intake modules within six months, the user went ‘haywire’.

Although the approach was agreed unanimously in the steering committee meeting, it created doubts in the mind of the user. This was further worsened by the fact that the project manager worked on the project on a part time basis and not full time as in her previous projects. Crucially, her control over the project was further impeded because she was not able to create a coalition with the chairman of the steering committee and the IT director. Since there was no alignment of interests between the project manager and the top management, the pre-requisites for a successful RAD approach were missing. The project manager failed to create a coalition with the chairman of the steering committee and the IT director in contrast to her previous experience. This could be due to the fact that the top management did not see the importance of the approach itself or the top management had a ‘laid back’ working style or they were not powerful enough to handle the dominant middle manager.

“…the chairman of the steering committee is not very powerful due to the fact that he is not able to control the group by giving leeway to them.”

Vendor (system developer) – Case 3 (01/08/2008, p. 3)

The intention of the vendor (software developer) to just employ the ‘adapt and adopt’ approach was also scrapped when fifty percent customization of the so called base ‘best practice’ system was agreed upon. In an unfavourable comparison with the vendor’s other clients that manage to adopt the approach, the non-standardized business processes in Case 3 created the need for such major customizations. Currently, the reengineering of the business process is being carried out in parallel with the business requirement study (BRS) sessions.

4.2.2 Misalignment of project expectations

The project manager’s critical misinterpretation was the simplification of the business process. While mainly concentrating on ensuring a smooth flow during the student intake, no attention was given to the more critical back-end process of the related modules. This was critical for the finance modules when her attention was given only to the process of issuing invoices to students rather than looking at the overall accounts receivable modules. This surfaced when the user (finance) was asked by the project manager to sign off the user acceptance test (UAT) for accounts receivable only after the student registration.
The failure of the project manager to assess the organization capabilities to comprehend such a tight deadline created major issues. This meant that the users were not able to develop complete and stable business processes through a formal process of re-engineering. As agreed by the users, a thorough re-engineering exercise of the current business processes seemed crucial due to the fact that Case 3 was an amalgamation of nine institutes with different organizational processes. To make things worst, the scepticism over the system escalated when the user’s requirements did not match the vendor’s existing ‘best practice’ systems. This has instantiated the need for a customization and to date, according to the vendor, fifty percent customization has been carried out.

The departure of the project manager from office after only seven months has strained the relationship between the user and the vendor. There are many reasons speculated for her leaving. According to the vendor, the dominance of the user over the project steering committee is crucial to understand her departure. Lack of support from the chairman of the steering committee and the IT director can also be seen as a motive for leaving the project. Primarily, the task of a project manager is a boundary person or a ‘buffer’ between the user and the vendor. Thus, the absence of the project manager has caused a direct impact on the vendor. Any system modifications have to be completed by the vendor. Unlike the vendor’s previous clients, the successful transfer of knowledge from the vendor to the IT team has enabled the clients to modify the used systems rather than relying to the vendor alone.

5 DISCUSSION

We begin by revisiting in turn the two research questions raised at the beginning of the paper:

1. What are the key actor networks in the development process and how do the networks develop and interact diachronically?

In this particular IS implementation case study, it can be clearly identified that there were attempts in creating alliances or networks to achieve the project’s objectives. In this case both the intended network failed due to the involvement of the various actors and consequently the pre-existing organizational network became stronger and more robust.

The problematization phase was seamlessly completed by the project manager in that she was able to create the need for a rapid system development methodology for the Case 3 IS implementation project.

Top management, the users, the IT department, the vendors and the project itself were identified as the key actors within the IS implementation. The project manager established herself as the obligatory passage point (OPP) for the project implementation. Everyone agreed on the need to develop and use the related modules within six months of the implementation process and according to the project manager this could only be achieved through RAD methodology. The intended networks are portrayed in figure 3.

Over time, the consensus that was achieved during the steering committee meeting on the application of RAD seems to have deteriorated. This was due to the project manager’s failure to understand the complexities of the organization and its unique history and features. Through her experiences she recognized the pre-requisites of a RAD approach but failed to acknowledge the organizational resource capabilities that have created chaos in the project. All her planned devices in ensuring interessement within the built network collapsed. Her plan to create a coalition with top management to ensure control over the project was diminished by the user/ middle manager’s dominance over the chairman of the steering committee. In addition there is no evidence showing she had achieved a good relationship with the IT director, further reducing her control over the project.

The project manager’s lack of judgment over the organizational readiness towards new approaches severely impacted the overall project management thus resulting in her failure to create the network, leading to her departure from the project. The project manager, during the interview session, foresaw that the project would eventually fail. This resulted from her failure to create a partnership with the steering committee that could have then supported her.
Further attempts to form a network between the vendor and the user within the organization took place. The vendor faced no issues during the problematization phase since the users identified the challenges of business process disintegration and non-alignment due to the merger. The latest version of the base system, the vendor, the user, the steering committee and the project manager were identified as the actors related to the problems and the vendor with its prior experiences and expertise established themselves as indispensable to the solution.

The base system, which was initially seen by the user as a perfect match to what was missing in their daily operations, was later viewed by the actors as incomplete and not full-proof. Negotiations and re-negotiations over the need to conduct a comprehensive business process re-engineering or major customizations over the based systems created tensions between the vendors and the users. Furthermore, the absence of the project manager in the complex relationship has reduced the indispensable position of the vendor in favour of a user-dominant network. The legacy left by the project manager on rapid development methodology has also impacted the overall development plans negatively.

The users were now more focused on modifying and enhancing the current system rather than concentrating on the undeveloped modules (Figure 4). The users established themselves as indispensable to the need for creating a more complete, full-proof and fully tested system.
Imposing themselves (i.e. the customer is always right) over the vendor to ensure compliance has to a
certain degree achieved enrolment. With their strong control over the steering committee and lack of
intention to appoint a project manager, the users are confident that they will complete the system
although it will take more time.

In summary, both failures to complete the process of translation are due to the strength of the existing
networks formed within the organization between the top management, the middle managers and the
IT director. This created a problem for the vendor and the project manager when they tried to impose
their wills in order to establish alliances.

2. What are the implications for research and practice resulting from our findings?

This work-in-progress paper provides multiple pictures of actors in action. As discussed earlier, each
of these actors had their own implicit intention or expectations towards the project. These intentions
and expectations were being translated into actions during the project executions. However, the
process of translating intentions and expectations into executable actions must be done with extreme
care and justification. In our case, there are several instances where the converted intentions are not
doable and are thus rejected. The IS development project site resembled a war zone where all the
players had strategies and an action plan. Therefore, when the project manager came into the project
site, based on her previous experiences, she had strategies that she thought no one could challenge.
She possessed technical skills as well as soft-skills. But everything was based on her experiences.
When she laid out her plans to the steering committee, they were identical to her previous projects and
did not take into consideration the uniqueness of the clients and her own capability as the project
manager. As clients they agreed with a strategy that provided them with the most up-to-date system
that can be developed and used as fast as possible.

Due to her inadequate judgement over the client capabilities, frustrations were revealed in all the key
players. The project manager also became frustrated since she was not able to realize her plans and left
the project. According to the project manager, the failure to create relationship with the steering
committee was due to the fact that they (the steering committee) saw her as an outside person (part of
the vendor) rather than a team member within the project. The users were frustrated due to the fact that
they could see what is going wrong but were not able to communicate with the project manager who
was too preoccupied with her own strategies.

Lastly, the vendor, who was frustrated as a result of the project manager’s concept of ‘ongoing
modifications’ that cause them to waste most of their time modifying old systems rather than
developing new systems. From the recent interviews conducted with other actors who dealt with the
project manager, she was not comfortable in her role. Clearly, the research will find resonance with
those studying enterprise system implementations and practitioners will recognise the problems arising
from a failure to establish usable networks with the client. In our further work we will explore these
issues in greater depth.

6 CONCLUSION

This paper is a research-in-progress and the dynamic of the enterprise system implementation project
in Case 3 has produced the continuous emergence of major events. To date, the two-year project
period has lapsed with little improvements to the percentage of completion for the finance system.

At thirty percent completion, the vendor pulled out of the project site with the argument that the
project period had come to an end. Prior experiences of both the vendor and the project manager are of
little help in this case scenario since they are not able to replicate their history. i.e. What worked in
Cases 1 and 2 seemed of little help in Case 3. The dominant network that existed before the project has
inevitably been reinforced by events. We will continue to monitor the major events of case three when
we analyse the interviews from the March 2009 visit and will provide updates to our story of
technology and knowledge transfer in this university setting.
References


DIGITAL DIFFERENTIATION, SOFTWARE PRODUCT LINES, AND THE CHALLENGE OF ISOMORPHISM IN INNOVATION: A CASE STUDY

Lena Andreasson, Viktoria Institute, Hörselgången 4, 417 56 Gothenburg, Sweden, lena.andreasson@viktoria.se
Ola Henfridsson, Viktoria Institute, Hörselgången 4, 417 56 Gothenburg, Sweden, ola.henfridsson@viktoria.se

Abstract
This paper examines the adoption of software product line engineering to implement digital differentiation of physical products. The introduction of such software-based variety can typically be challenging for firms innovating within the realm of a manufacturing paradigm. In particular, the mutual dependency between the organization design and product design of new product developing firms may counteract attempts to induce change through software product line engineering. On the basis of innovation theory and the notion of isomorphism, the paper presents a case study of digital differentiation at one of the world’s largest automakers, GlobalCarCorp. Relating to the literatures of software product lines and product families, the contribution of the paper is a lens through which to understand the role of isomorphism in implementing digital differentiation in new product development. In addition, practical implications are derived from this in-depth study.

Keywords: digital differentiation, IS implementation, software product lines, innovation, innovation management, isomorphism.
1 INTRODUCTION

The digitization of physical products is radically challenging the innovation processes of established firms. On one hand, it multiplies the space of digital options for not only incrementally improving existing offers but also to launching radically new services on the market (Jonsson et al. 2007). On the other hand, seizing emergent digital options is difficult because current product innovation practices may not involve necessary IT capabilities and organizational agility (Sambamurthy et al. 2003).

In the innovation literature, the tension between options provided by a new technology and institutionalized practices established over longer periods of incremental innovation has received significant attention (Abernathy 1978; Abernathy and Utterback 1978; Anderson and Tushman 1990; Hargadon and Douglas 2001; Tushman and Anderson 1986). In particular, the notion of dominant design captures how institutionalized innovation practices typically congeal over time as a template for product innovation within an industry (Teece 1986). Dominant designs help firms organize their innovation processes so that they capitalize on their intellectual, relational, and technical resources. In particular, successful firms orchestrate a reciprocal relationship between organization design and product design. Baldwin and Clark (2000) refer to this relationship as the fundamental isomorphism between task structure and design structure.

While the fundamental isomorphism between organization design and product design is important to exploit a dominant design, it lowers a firm’s capability to respond to technological discontinuities (Anderson and Tushman 1990). The introduction of digital technologies in physical products represents such a discontinuity for manufacturing firms. While existing processes embed an innovation logic fine-tuned for a tangible, hardware-based business, embedding software into existing product architectures introduces an alien innovation logic requiring new architectural knowledge (Andersson et al. 2008). A pressing issue is therefore how to handle these parallel logics when organizing innovation processes.

This paper focuses on the adoption of software product line engineering for implementing digital differentiation of physical products. This focus directs the attention to two literatures on modularity, currently divided along disciplinary lines. First, the body of literature on software product lines is concerned with strategies for managing software families in a way that facilitate flexibility and modularity (Clements and Northrop 2001; Pohl et al. 2005). This software engineering-literature has its origins in Parnas’ (1976) work on software families, and has evolved as a result of the growing complexity of handling and exploiting software innovations. Second, the product family literature is concerned with platforms as a means for achieving product differentiation without compromising mass production advantages. This body of literature has emerged within the context of the product innovation literature, addressing the concerns of manufacturing companies in areas such as the automobile industry (Karlsson and Sköld 2007), camera industry (Robertson and Ulrich 1998), and office material industry (Cooper et al. 2001).

The contribution of this paper is a lens through which to understand the mutual dependency between organization design and product design in implementing digital differentiation in new product development. Drawing on innovation theory and the notion of isomorphism (Baldwin and Clark 2000), we present a case study of such differentiation through software product line engineering at one of the world’s largest car manufacturers: GlobalCarCorp. The following research question is addressed: How can software product lines be implemented to leverage digital differentiation in firms innovating in a manufacturing paradigm?

The remainder of the paper is structured as follows. Section two reviews the extant literatures on product families and software product lines. Section three outlines the theoretical framework including the notion of isomorphism. While section four presents the research methodology, section five outlines the case of software product line engineering at GlobalCarCorp. Analyzing the case, section six outlines implications for theory and practice, while the last section concludes the paper.
2 RELATED LITERATURE

Throughout this paper, we refer to digital differentiation as a software-enabled process that identifies, implements, evaluates, and maintains distinctive characteristics of a physical artifact relative to other artifacts of the same class. For the purposes of this paper, we review two streams of literature that relate to digital differentiation but yet are divided along disciplinary lines. First, reviewing the innovation literature, product differentiation has been an established topic for long. The research on mass customization (Pine II 1993) as well as product platforms and families (Halman et al. 2003) are of specific relevance in the context of this paper. Second, research on software-based differentiation has emerged within the software engineering community, typically oriented towards methodologies for managing variety in specific software families (Clements and Northrop 2001; Clements and Northrop 2003; Pohl et al. 2005).

2.1 Product Family Research

Product differentiation has been addressed quite extensively in the innovation literature (Aaker 2003; MacMillan and McGrath 1997; Pine II 1993; Robertson and Ulrich 1998). Conceptualized as product family, platform, or line research (Halman et al. 2003; Robertson and Ulrich 1998), one of the main themes of this literature has been the challenge of achieving a differentiated offer to customers without increasing cost and development time proportionally. The implementation of differentiation attributes on a common architecture promises to maintain or increase the market share and keep customer attention (Aaker 2003). The typical approach to this challenge is the development of a common architecture, which allows for the delivery of well-adapted products while maintaining a high degree of commonality of components (Lundbäck and Karlsson 2005). In this regard, mass-scale advantages can be combined with customization (Pine II 1993).

Managing product families can be challenging. As an example, Karlsson and Sköld (2007) note how multi-branded strategies typically involve counteracting forces linked to commonality and differentiation. On one hand, the commonality force directs attention to identifying a common solution for a set of brands identified as a product family. On the other hand, differentiation within this family is important as to avoid brand “cannibalization”. It has been observed that careless search for commonality opportunities through common architectures may cause negative revenue effects (Kim and Chhajed 2000).

2.2 Software Product Lines

In the software engineering community, software product lines have emerged as an increasingly important research theme (Clements and Northrop 2001; Pohl et al. 2005). The introduction of digital technology in physical products enables new forms of differentiation that do not rely on hardware changes. In this regard, costly development of new tools and models for each product change is virtually avoided. Firms introducing software product line engineering are promised to achieve product differentiation in a more cost effective way (Kreuger 2002). Indeed, it is argued that the use of software product line engineering not only results in lower costs, shorter lead times, and increased variety, but also in higher quality of software products belonging to the same family (Ereño et al. 2006).

Reflecting its software engineering origin, the literature on software product lines is design-oriented in its emphasis on better methodologies for managing software families. As outlined by Clements and Northrop (2003, p23) “a software product line is a set of software intensive systems sharing a common, managed set of features that satisfy the specific needs or mission and that are developed from a common set of core assets in a prescribed way”. To this end, domain engineering (Pohl et al. 2005), or core asset development (Clements and Northrop 2001), is the process where identification of common and variable features of a product family is accomplished.
Pohl et al. (2005) distinguish five iterative steps of domain engineering to ensure variability of the common assets: product management, domain requirements engineering, domain design, domain realization, and domain testing. First, product management sets the direction of the project and manages consumer goals throughout the software development process. The objective is to develop the product roadmap to the extent it is possible in a given point in time. The second step is to determine common requirements of the domain. The following two steps involve designing a solution, which is later implemented. The last step of this iterative process is to test the domain solution. While these steps resemble any single system development process, the emphasis on identifying commonalities within the software family and the options of variability makes software product line engineering promising for accomplishing software-based differentiation.

3 THEORETICAL FRAMEWORK

The literatures on product families and software product line engineering are important backdrops in our quest of understanding digital differentiation of physical products. However, in these streams of literature, little attention is paid to the relation between organization design and product design. Given the received innovation literature on dominant design (Anderson and Tushman 1990; Teece 1986; Tushman and Anderson 1986) it would be useful to further investigate the institutional challenges evoked in introducing digital differentiation in the new product developing process in firms designing and manufacturing physical products.

Isomorphism is an established concept within institutional theory (DiMaggio and Powell 1983; Meyer and Rowan 1977; Scott 1995; Zucker 1977). In this paper however, we draw on Baldwin and Clark’s research (2000) on design and industry evolution. In particular, we apply their layers of structure (LoS) model and its relation to fundamental isomorphism as a lens with which to understand the role of institutional arrangements when transforming innovation processes assembled for hardware-based differentiation into processes suited for digital differentiation. The LoS model distinguishes three different layers of structure important in product innovation: artefact, design, and task structures. On the most basic level, the artifact structure refers to the tangible instantiation of a particular design. In other words, the artifact structure is what can be seen, heard, touched, and used. It simply performs functions that create value for its user. Encompassing its architecture and functions, the design structure is a description of the artefact. In this regard, it is not only the technical blueprint of the artefact but also defines the adaptability of the design to changing circumstances and markets. Typically, physical products embed a modular architecture, enabling flexibility because of its implementation of independence between structural elements of the design (Baldwin and Clark 1997; Sanchez and Mahoney 1997; Ulrich 1995). Lastly, task structure refers to the set of activities needed to be performed to realize a design. The task structure implies organizational design elements including communication channels, information filters, and standard operating procedures (Henderson and Clark 1990).

In this paper, we are particularly interested in the relation between task structure and design structure. The received innovation literature describes alignment of these structures as imperative to enable successful new product development (Henderson and Clark 1990; Sosa et al. 2004). Referring to this relationship as the fundamental isomorphism, Baldwin and Clark (2000) not only underline the tight coupling between the two structures but also suggest the challenges involved when initiating change in either structure. In what follows, we apply the LoS model in general and the concept of isomorphism between design structure and task structure in particular for understanding digital differentiation through software product lines at a global automobile manufacturer. The next section describes our methodology.
4 RESEARCH METHODOLOGY

4.1 Research Setting and Design

Our case study research was conducted at one of the world’s largest automakers, referred to as GlobalCarCorp throughout this paper. At GlobalCarCorp, R&D operations serve many brands. The multiplicity of brands creates incentives to identify and develop product platforms for making differentiation economically feasible.

In 2005, one subdivision of GlobalCarCorp in Sweden received global responsibility for R&D on instrument clusters. This responsibility came in a transition period of instrument cluster design. Digital technologies were introduced into the design of instrument clusters including speedometer, tachometer, and other types of driver information during this time. This change in technological basis introduced a range of digital options for new types of differentiation within product families. Given its traditional hardware basis, however, digital differentiation also breaks with extant task and design structures.

Designed as case study research (Eisenhardt 1989), our 12-month study was initiated in November 2007 and was specifically focusing the 3-year process by which digital differentiation was implemented through software product lines at GlobalCarCorp. Exemplifying process research rather than variance research, the intention has been to identify and predict “patterned regularities over time” (Markus and Robey 1988).

4.2 Data Collection and Analysis

The data collection can be described as an iterative process. It consisted of three phases, as described below. Concurring with the typical case study, data collection also included multiple data sources including semi-structured interviews, meeting notes, workshop documentation, and email correspondence.

The first data collection phase included meeting attendance as to frame the area of concern. Analysis of meeting notes generated five areas of significant interest (development methods, organization, digitization, differentiation, and architecture), which, together with relevant literature, guided the remainder of the data collection process.

The data collection of the second phase mainly included recorded and transcribed semi-structured interviews. All in all, 25 interviews, ranging between 40 minutes and two hours, were conducted. The interviews were based on an interview template developed on the basis of the themes identified in the first phase. Respondents ranged from managers to developers, and they covered expertise such as software, architecture, graphical interfaces, ergonomics, design, and market. The semi-structured interviews improved the understanding of the organization’s task and design structures and how these changed with the introduction of the SoftCluster tool. Statements from respondents concerning one or many of the five significant areas of interest were especially focused on.

The third phase was confirmatory in character. After completing the first two phases, process charts of the software development and differentiation process were developed. These process charts were discussed with GlobalCarCorp employees as to verify and extend our preliminary understanding of the case. The findings were assessed in view of relevant literature. Then, the theoretical framework was applied to extend our ideographic understanding beyond the specific case setting.

Exemplifying engaged scholarship (Van de Ven 2007), workshops were organized with the project team over the last two phases to feedback new perspectives on their work practice based on prior research and empirical findings.
5 SOFTWARE PRODUCT LINES AT GLOBALCARCORP

In 2005, GlobalCarCorp designers working with instrument clusters introduced the idea of implementing software product line engineering. There were three main reasons for doing this:

- Controlling and managing software variants became significant as GlobalCarCorp decided to leverage from its portfolio of multiple brands.
- GlobalCarCorp initiated an economics of scale strategy with the intention to use the same hardware for many products, while accomplishing differentiation through software.
- The digitization of the car necessitated increased software development professionalism.

While the instrument cluster traditionally consists of hardware with little, or no, modification possibilities, the digitization of interfaces in the instrument cluster opened up new possibilities to present information in flexible and differentiated ways (see Figure 1). For instance, context-aware information presentation can be used to reflect the current condition of the car or immediate traffic situations. It also allows the introduction of new functionality and differentiation features in the instrument cluster.

Figure 1. Instrument cluster with fully reconfigurable display

GlobalCarCorp identified five product families to manage instrument cluster variants for over 80 car models, amounting to approximately 3.8 million cars each year until 2012. In addition, the new strategy involved separating hardware and software development. GlobalCarCorp intended to increase flexibility in the selection of hardware and software suppliers.

Different hardware requirements and cost requirements guided the process. Each cluster family was divided in different levels: base, mid, and high, representing levels of appearance and features. In other words, the complexity of handling software variants was considerable. Because of this complexity, GlobalCarCorp decided to adopt a tool that would facilitate the process.

5.1 The SoftClusterTool

The SoftClusterTool is a comprehensive XML-based tool for differentiation of the software architecture of instrument clusters. The tool enables the development of one basic software module for all types of displays in a specific product family. By using an editor, a HMI (Human Machine Interaction) engineer can differentiate the common cluster through changes to the XML-files controlling static data such as font, language, colors, layout, and graphics, or the dynamic data produced by functional units of the car. Examples of functional units are the FM tuner, phonebook, and thermometer, as well as units related to the condition of the car.

5.2 Implementing the New Process

Before adopting the SoftClusterTool, HMI engineers were responsible for the HMI portion of the instrument cluster on the basis of input from the design department and managers of functional units. HMI requirements and the overall cluster design would later be implemented by the selected instrument cluster supplier (see Figure 2). In this regard, the supplier was responsible for implementing and delivering the integrated cluster, including both hardware and software. Once implemented, GlobalCarCorp tested and evaluated the cluster. Changes and updates were then
collected and submitted to the supplier as change requests. This process was unique, although similar, for all brands within GlobalCarCorp.

A common problem of the traditional design process was differences in requirements interpretation between GlobalCarCorp designers and the supplier. In other words, there were often a gap between the final implementation of the supplier and GlobalCarCorp’s original intention. In addition, new requirements emerged over time, including seemingly minor issues such as change of a word, bitmap or a color. Since every update required contacts with the supplier, the process was both time consuming and costly. In fact, as a designers noted, late change requests were anticipated by the supplier and essentially an important element in their business model.

![Figure 2. GlobalCarCorp’s old task structure for instrument cluster design](image)

In view of the problem with change requests, the tool was adopted in 2005. The intention was to in-house the software for realizing digital differentiation, also called HMI software, side of realizing instrument clusters. Rather than having the supplier implementing change requests related to HMI software updates, the new process, enabled by the tool, attributed the instrument cluster suppliers the role of a component supplier only. Without prior experience of software implementation, new roles and tasks had to be defined at CarCorp.

Given that the software implementation previously was done by the supplier, cognitive ergonomics and interaction design competences had been prioritized within CarCorp’s HMI design group. In view of the new tool, people with programming skills in general and XML competence in particular were necessary. In addition, a new role was defined to cater for the idea of using the tool and the new process for accomplishing digital differentiation between brands in GlobalCarCorp’s product families. Before separating software and hardware, as was done in the new process (see Figure 3), each brand-specific cluster had a responsible manager.
Due to the significant effort put into developing the SoftClusterTool, however, little focus was nevertheless on adjusting the organization and process to harvest the anticipated benefits of the tool. The traditional way of developing information cluster with significant focus on hardware components and their design structure and coupled task structure remained at GlobalCarCorp.

One example of this prevalence of institutionalized practices was the lack of standard operating procedures for gathering and communicating software differentiation requirements from GlobalCarCorp stakeholders. In this regard, programmers were required to work proactively to acquire relevant information. This was both time-consuming and challenging because the programmers had to know exactly who to contact and what information to ask for. As a result, programmers often had to make decisions outside their obligations to meet deadlines and the overall productivity pressure. Essentially, they had to develop informal ways of gathering requirements, drawing on requests from people they already knew or were referred to by other employees. Ideally, such requests would involve perspectives from employees working on branding, interaction design, functional units (e.g., radio, antenna, or navigation) or cluster hardware. Without any general process in place, however, programmers of the HMI group often faced conflicting requirements and had to rely on individual judgment.

Before adopting the SoftClusterTool, the task of a HMI engineer at GlobalCarCorp was separated from other modules of the instrument cluster. Many decisions were made by the supplier, including the important task of pooling and integrating different requirements. While programmers face the consequences of the new tasks that the SoftClusterTool implies, other stakeholders at GlobalCarCorp were not fully aware of the new way of gathering requirements. Consequently, they drew on work processes already established before introducing software product lines. The anticipated benefits of the SoftClusterTool therefore remained to be seen.

6 DISCUSSION

Operating in a manufacturing paradigm, product innovation in the automotive industry is typically geared towards integrating different hardware components, or modules, into a fully functional product. As an important element in such integration, requirements engineering with regard to the different
modules and their interfaces is a core competence (cf. Sanchez and Mahoney 1997). Following the
digitization of the car, however, automotive design increasingly includes separating the software
elements from hardware ones. Following this design change, the relationship between the supplier and
the automotive firm is changing. Using our theoretical framework (Baldwin and Clark 2000), the
design change occasioned a change in the layers of structure of designing instrument clusters at
GlobalCarCorp. The digital options generated were challenging in view of the established
isomorphism between the established design and task structures.

The case presented in this paper illustrates the adoption of software product line engineering in
instrument cluster design at a global automaker. In particular, we examine the introduction of a tool
that was designed not only to enable flexible catering for change requests related to software but also
to allow for digital differentiation over product families. The SoftClusterTool was an attempt to
harvest commonality benefits on the hardware side of instrument cluster design, while achieving
differentiation between GlobalCarCorp brands with software. In this regard, the adoption of software
product line engineering was an attempt to implement a vision of digital differentiation.

Referring to task structure as the set of activities needed to be performed to realize a design (Baldwin
and Clark 2000), it is clear the adoption of the tool at GlobalCarCorp rendered some important
implications associated with the established isomorphism between design structure and task structures
in instrument cluster design: the redefined supplier role and the decoupled development of hardware
and software. In what follows, we explore these implications in view of the LoS model in order to
understand the innovation challenges faced by producers of physical products when attempting to
seize digital options.

6.1  Isomorphism and Parallel Task Structures

As noted in our case, the old task structure of instrument cluster design was highly dependent on
suppliers. While GlobalCarCorp specified, tested, and evaluated the implemented product before
assembly into the car, some core activities in the task structure were performed at the supplier
including hardware-software implementation and the coordination of hardware and software
requirements. Given that this division of work tasks was not only institutionalized in GlobalCarCorp’s
supplier relationships but also reflected in the dominant industry structure, the new design structure
occasioned by software product line engineering was challenging. Indeed, the debundling of hardware
and software design instantiated through the SoftClusterTool created a new situation for both
GlobalCarCorp and suppliers. At GlobalCarCorp, however, the task structure changes needed for
mirroring the design structure changes have not yet been fully completed. At the end of the study, it
could be noted that no real isomorphism existed between the design structure and task structure. As a
tangible example of this, the final products, the instrument clusters, currently lack digital
differentiation characteristics between brands (which was one of the main objectives at the outset). In
other words, despite a change in the design structure and the implementation of the SoftClusterTool,
prevailing task structures counteracted these initiatives.

There were three important aspects of digital differentiation and the established isomorphism between
task structure and design structure at GlobalCarCorp. First, one tangible change amounted to the
newly employed programmers, who were supposed to implement requirements submitted by
stakeholders throughout GlobalCarCorp. Because stakeholders such as brand managers, interaction
designers, and functional unit managers were unable to systemically supply such software-specific
requirements to SoftClusterTool users, however, they had to take on tasks for which they were not
responsible. This caused frustration among a key group of people involved in the digital differentiation
initiative.

Second, as a multi-branded firm, GlobalCarCorp viewed digital differentiation and software product
lines as elements in a strategy intended to combine commonality and differentiation. Given that such
strategies traditionally applies to the hardware side of the car (cf. Sköld and Karlsson 2007), the
corresponding strategy on the software side was completely new. It shifted major tasks of the domain
engineering process (Pohl et al. 2005) to GlobalCarCorp from suppliers. In this regard, complexity has increased and new patterns of requirements coordination and integration were needed. Even though GlobalCarCorp recruited programmers with XML knowledge, the new engineers could not fully accommodate the responsible task of handling the domain engineering process. Because GlobalCarCorp had no prior competence in the area, the new employees could not foster new requirement gathering processes with other stakeholders in the firm who still relied on the established hardware-related isomorphism between design structure and task structure. As an illustration, the differentiation plan (Sköld and Karlsson 2007) for the five different cluster families only included strategies for hardware differentiation.

Third, the institutionalized boundary relationships between suppliers and GlobalCarCorp were blurred. The old task structure had been instantiated and fine-tuned over many years. With the new initiative, a portion of that process remained, namely, the hardware and pre-installed software, while the differentiation software responsibility was taken on by GlobalCarCorp. This required new boundary-spanning behavior, where the old task structure needed to be split into two parallel ones (cf. Figure 3). In essence, the hardware set the rules and the software had to follow. As illustrated in Figure 3, the hardware was first developed and then software was added. This resulted in an inflexible way of handling software updates. While software is more flexible than hardware with regard to updates, adoptions, and adjustments, this sequencing downplayed the flexibility introduced by the SoftClusterTool. This advantage dissolved in the already introduced hardware constraints. It simply was not possible to make software updates as frequently and swiftly as the tool rendered possible due to lack of supporting task structures.

7 IMPLICATIONS

We report a case study of GlobalCarCorp’s proactive attempts to introduce software product line engineering for combining commonality and differentiation in instrument cluster design. Using Baldwin and Clark’s (2000) LoS model and the notion of isomorphism, we analyze this case for understanding the mutual dependency between organization design and product design in implementing digital differentiation in new product development. Our study has significant implications for the innovation literature on product families (e.g., Halman et al. 2003; Karlsson and Sköld 2007; Robertson and Ulrich 1998) and software product line engineering (Clements and Northrop 2001; Pohl et al. 2005).

In sum, in spite of attempts to introduce a new task structure around digital differentiation software in cluster design, our evidence pinpoints the difficulty for a manufacturing organization to establish software product line engineering processes that break with the fundamental isomorphism between task structure and design structure (cf. Baldwin and Clark 2000). Because the established isomorphism is based on a hardware paradigm that will remain in firms producing physical products, attempts to establish new design structures and task structures for the software portion of the sub-system are challenging. As illustrated in the requirement elicitation problems at GlobalCarCorp, the new task structure will be severely challenged by the dominant design and task structures related to hardware. This insight is a contribution to the software product line literature (Clements and Northrop 2001; Pohl et al. 2005), which tends to disregard surrounding processes, including hardware-related ones, when outlining their recommendations on how to accomplish software variety. In a similar vein, in its orientation on physical products only, the product family literature (Halman et al. 2003; Karlsson and Sköld 2007; Robertson and Ulrich 1998) cannot either cater for this observation. The innovation literature on product families tends to oversee the increasing digitization of physical products, thereby disregarding the new and relevant relationships emerging between differentiation across the digital and physical realms.

We believe that more research is needed in this domain in order to handle the innovation challenges and opportunities that manufacturing organizations face in a time when their manufactured products...
are becoming digitized. After all, a new set of core competences seems to be needed when seizing emergent digital options.

8 REFERENCES


TOWARDS A RESEARCH FRAMEWORK FOR A HUMAN DEVELOPMENT-BASED “BOTTOM OF THE PYRAMID” ICT DEVELOPMENT STRATEGY IN SOUTH AFRICA

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0326.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Developing countries, Digital Divide, E-inclusion / exclusion, Action research</td>
</tr>
</tbody>
</table>
TOWARDS A RESEARCH FRAMEWORK FOR A HUMAN DEVELOPMENT-BASED “BOTTOM OF THE PYRAMID” ICT DEVELOPMENT STRATEGY IN SOUTH AFRICA

Walter Brown, Monash SA, 144 Peter Road, Ruimsig, South Africa, walbrown@mail.ngo.za
Irwin Brown, University of Cape Town, Private Bag, Rondebosch, 7701, Cape Town, South Africa, Irwin.Brown@uct.ac.za

Abstract

There is a wealth of knowledge linking Information and Communications Technologies (ICTs) to human development, but Africa’s leaders have not used it well to develop its human capital. The continent has the lowest global levels of ICT diffusion and human development. This paper proposes a way of developing Africa’s vital access to knowledge through ICT, by reviewing and using the human development record of South Africa to develop a research framework for a human development – based ICT development strategy. The framework comprises a research-based bottom-up participative action program that responds directly to the specific information needs of the country’s socio-economically marginalized communities, the so-called “Bottom of the Pyramid” (BOP), by ensuring their maximum participation in the research and development process. The proposed framework draws liberally from the vast accumulated knowledge about human development and its principal tool, the ICTs, especially from Asia where similar development activities are delivering impressive results. The multi-cultural and multi-lingual characteristic of South African society, including its recent socio-political history, presents numerous challenges for balanced human development, but also unique opportunities to build an effective model for improving the usefulness of ICTs for human development in the country and by extension, in Africa as a whole.

Keywords: Human Development; ICT for Development, South Africa, Bottom of the Pyramid
1 INTRODUCTION

Soweto, South Africa, 2008: “Bra” (brother) Peter Wembe is a 64 year-old Soweto based jazz musician known as “Piwe” by his friends. He obtained a Bachelor of Music degree from a leading South African university in 1985. He lives with his wife, two unemployed daughters and a musically talented 15 year-old grandson Nelson. Nelson’s father, said to have been a budding musical genius himself, died of HIV/AIDS soon after Nelson’s birth. Piwe’s main source of income is an informal “Spaza” shop, a modified shipping container out of which the family sells basic groceries, cigarettes, soft-drinks and pre-paid telephone cards to the local community. The total family income from this informal business ranges from R2000 to R3000 per month (approximately US$2 per person per day). This income is supplemented by a few local jazz concerts each year, a few “Corporate Gigs” during end of year festivities, occasional international musical tours arranged by international jazz music celebrities who appreciate Piwe’s style and talent, and a small income from royalties on past recordings. Nelson attends an “informal” school near his home, but has largely lost hope of academic achievement. He rehearses diligently in the hope of emigrating to Europe or the USA as a professional jazz artist. Nelson has appeared on several local television talk-shows and documentaries promoting local young talent, but his contribution to the family’s income is minimal. The Wembe family have basic ICT literacy, an old personal computer with dial-up connection to the Internet, and they dream of establishing a small music studio from which they can produce and market their own commercial recordings. They also worry constantly about security, and their foreign-sounding surname which may provoke xenophobic violence against them, as has happened to many similar families in other informal residential enclaves throughout the country in May 2008 (HSRC 2008, UN 2008). Piwe is a third generation South African, his grandfather was an early migrant mine-worker from the then Belgian Congo.

The semi-fictitious story of the Wembe family is based loosely on a living family residing in Soweto at the time of writing this paper. Their names, ages and other identification features have been changed, but their lifestyle, their daily challenges and their hopes and fears are representative of the majority of South Africans residing at the bottom of South Africa’s social pyramid (BOP) (IFC, 2007), also called the “Second Economy”, a description coined by South Africa’s former State President Thabo Mbeki (du Toit and Neves, 2007). Their story puts a human face to this development discourse, and provides a useful foundation for addressing their needs to access information and knowledge via the global ICT infrastructure that will enable them to help themselves out of the vicious poverty cycle that afflicts so many South Africans.

The Human Development concept was introduced by the United Nations Development Programme (UNDP) in 1990 to replace GDP per Capita as the main measure of national development. UNDP publishes annual Human Development Reports (HDR), with each year focussing on a different theme of concern to humanity as a whole. The theme for HDR 2004 was “Cultural liberty in today’s diverse world”, a theme that harmonizes well with South Africa’s currently fragmented multi-cultural society. This paper draws extensively from the most recent HDR 2007/2008 (HDR 2007) which focuses on climate change and the search for human solidarity to combat the global threat, but also provides an extensive database of human development indicators. A full history and description of the concept is beyond the scope of this paper, however an excellent description of the evolution of the concept has been prepared by Stanton (2007). A broad definition of human development favoured by the OECD is “the process of enlarging people’s choices. Their three essential choices are to lead a long and healthy life, to acquire knowledge and to have access to the resources needed for a decent standard of living” (OECD, 2003). The human development index (HDI) is a composite index used to gauge development across these three choices (HDR, 2007).

The authors of this paper share the opinion that the most important human development choice is the choice to acquire knowledge. Without this choice, all the other human development choices will be undermined by the lack of historical and current knowledge of how they can best be achieved.
Information and Communications Technologies (ICT) are the most efficient knowledge dissemination tools today. They enable unrestricted information flows that support the creation and use of global knowledge for the full range of human development, bridging cultural, gender, geographic and even knowledge divides. ICTs have been closely linked to all the Millennium Development Goals (MDG), (Gilhooly, 2005), and their importance to global development has been stressed by Payaril (2005) who observes that “the divide between the income rich and the income poor, the technology have-s and the technology have-nots, the information rich and the information poor, has become the most serious political economic problem facing the world today”. This paper draws on the accumulated knowledge of both human development and ICTs, with specific reference to the South African experience, typified by the semi-fictitious Wembe family. In the next section, the human development challenges in South Africa are discussed. This is followed by discussions of ICT development in South Africa. Finally a brief outline of the proposed framework for ICT development at the BOP is provided in Section 4. The conclusion provides a speculative discussion of how the Wembe family and BOP compatriots could benefit from the proposed research.

2 HUMAN DEVELOPMENT IN SOUTH AFRICA

South Africa has the largest economy on the African continent with purchasing power parity gross domestic product (GDP PPP) of nearly US$470 billion in 2007, closely followed by Egypt with US$404 billion, and nearly twice that of Nigeria, the most populous and third largest economy on the continent. Despite this, South Africa faces one of the severest human development challenges on the continent. The human development statistics that follow quantify South Africa’s human development challenges, drawing statistical data from the UNDP Human Development Report (HDR 2007), the IMF World Economic Outlook for 2008 (IMF, 2008), Statistics South Africa (Stats SA, 2007), the South African Development Indicators (2008), and ICT-specific data from the International Telecommunication Union’s (ITU) World Telecommunications ICT Indicators 2007 (ITU, 2007).

2.1 HDI Indicators

South Africa (HDI rank 121 in 2005) was one of 17 countries that suffered a reversal in human development between 1980 and 2005. After steady growth from 1980 to a peak of 0.745 circa 1995, a major decline ensued which resulted in the HDI reversing to 0.674 by 2005, the third steepest global decline after Zimbabwe and Swaziland. The country’s HDI global rank fell 36 places from 1990 to 121st in 2005 (HDR, 2007). Figure 1 compares South Africa’s HDI to the averages of each major development region.

![HDI Trends: South Africa compared to major world regions](image)

*Figure 1 HDI Trends: South Africa compared to major world regions
Source: UNDP HDR (2007/2008)*
2.2 The Choice of a Long and Healthy Life

The choice to lead a long and healthy life is one of three main indicators making up the HDI (HDR, 2007). A key indicator is life expectancy at birth amongst citizens of a country, with subordinate indicators that include the impacts of specific disease pandemics. The key contributor to South Africa’s steep HDI decline shown in Figure 1 was the nation’s failure to control the spread of HIV/AIDS. In 2006, South Africa had the 6th highest prevalence with an infection rate of 18.8%, outranked only by the country’s neighbours Swaziland (33%), Botswana (24%), Lesotho (23%), Zimbabwe (20%), and Namibia (20%). As a consequence of the HIV/AIDS pandemic, South Africa’s life expectancy at birth declined to 53.4 years by 2005, a rank of 145 out of 173 countries, and the national probability at birth of surviving to age 65 years fell to 33.9% for South African males, placing South Africa 160th out of 173 countries in this longevity indicator (UNDP, 2007). For greater clarity on these vital HDIs, Figure 2 compares South Africa with Thailand, Colombia, Turkey and South Korea, selected for their demographic and economic similarities, and for their geographic representation.

![Life expectancy at birth; Probability of surviving beyond age 65 (males)](image)

*Figure 2* Life expectancy at birth; Probability of surviving beyond age 65 (males)


2.3 The Choice of Knowledge Acquisition

The choice of knowledge acquisition is another key indicator that makes up the HDI. It is assessed with regards to literacy rates and school enrolment ratios (HDR, 2007). South Africa’s national educational achievements have been disappointing. The PIRLS (2006) and TIMSS (2003) international education assessments comprising 40 and 46 countries respectively, ranked South Africa last amongst the participating countries with average achievement scores 61%, 57% and 51% below the global averages for reading literacy (PIRLS 2006), and mathematics and science (TIMMS 2003) respectively. Compounding these disappointing results is their racial characteristic. Data derived from South Africa’s 2001 census shows that 16.5% of Black South Africans (pre-colonial indigenous South Africans) had no formal education, compared to 3.1% White South Africans (of European only ancestry), 6% Indian South Africans (ancestry traceable to the Indian sub-continent) and 8.7% Coloured South Africans (descendants of mixed race unions irrespective of ancestral origin) (Stats SA, 2007). At the higher end of education, 12.8% of White South Africans had a diploma or first degree as their highest educational achievement, compared to 1.8% of Black South Africans, 5.9% Indian South Africans, and 1.7% Coloured South Africans. Figure 3 illustrates the educational divides between South Africa’s official race groups. The post-apartheid South African government has retained the
apartheid legacy of race-based classification as a means of monitoring progress in reversing apartheid era racial segregation and discrimination.

![Figure 3](image_url)  
**Figure 3.** Highest levels of educational achievement within South Africa’s race groups. Data source: Stats SA (2007) population census 2001.

2.4 Access to Resources

The choice of having access to the resources needed for a decent standard of living is the third indicator that makes up the HDI (HDR, 2007). National poverty and wealth, the distribution of both, and inequality extremes are the popular measures of this choice. South Africa has the highest GDP in Africa, but also one of the highest levels of economic inequalities in the continent and the world. A popular measure of economic inequality is the Gini coefficient. A nation or community that shares its wealth equally amongst its members has a Gini of 0, while one in which all wealth is consumed by one person has a Gini of 100. UNDP HDR 2007/2008 reports a Gini of 57.8 derived from year 2000 surveys, while the South African government reports a Gini of 69 for year 2005. The rising level of economic inequality is recognised by the South African government and its development partners as one of the most challenging aspects of human development. Due to historical reasons, the country’s inequalities have a strong racial characteristic. Statistics derived from South Africa’s Development Indicators 2008 (PCAS, 2008) show that Black South Africans who constitute 80% of the population live off an average gross national income (PPP) of approximately US$2500, equivalent to that of Bangladesh and Haiti, while their White compatriots enjoy an equivalent income ten times higher, equivalent to high income economies like New Zealand (HDI rank 19) and South Korea (HDI rank 26).

South Africa’s unemployment levels are similarly skewed along racial lines, with 30.5% unemployment amongst Black South Africans, compared to 4.5% for their White compatriots (Stats SA, 2007). Inequalities of this magnitude are a source of conflict and social instability. Stewart (2000) concludes that in every major conflict, group perceptions and identities are enhanced by sharp differentiation in political participation, economic assets and income, and social access and well-being. These inequalities present a strong environment for conflict. Cramer’s (2005) analyses of conflict and inequality quotes Plato (circa 400 BC): “We maintain that if a state is to avoid the greatest plague of all - I mean civil war, though civil disintegration would be a better term - extreme poverty and wealth must not be allowed to arise in any section of the citizen-body, because both lead to both these disasters” to argue that the debate is as old as mankind’s strive for peaceful societies. By relating Cramer’s (2005) and Stewart’s (2000) analyses to South Africa’s inequalities and various media reports of historical and current conflicts throughout the world, troubling parallels emerge.

As, Cramer (2005) and Stewart (2000) argue, South Africa’s deep inequalities may not provoke internal instability, but conditions exist for this to happen. The outbreak of xenophobic violence in
May 2008 (HSRC, 2008) represents a disturbing harbinger of this possibility. In a related development, an important outcome of the World Summit on the Information Society (WSIS) was the establishment of a global research unit and think-tank of how ICTs can be used prevent conflict, and to help societies to recover from it after an outbreak (Mansell and Nodenstreng, 2007). The Switzerland based ICT4Peace Foundation is the UN specialized organ coordinating the global search for ICT applications that prevent and reduce the threats of conflict (Stauffacher et al., 2005).

### 3 TECHNOLOGY DIFFUSION - ICT

In this section, South Africa’s ICT development is compared to those of Colombia, South Korea, Thailand and Turkey, four countries with demographic and economic similarities with South Africa over the last 20 years. Table 1 illustrates these similarities, and shows that South Africa led the group in both economic and ICT indicators in 1980 and 1970 respectively, but lagged in both by 2005. South Korea’s impressive growth record has been attributed to aggressive mass education and rollout of ICTs, typified by the country’s early 1980’s policy to provide a telephone line to every family in the country (Reynolds et al. 2005). This policy has served South Korea well. The fixed line telephony infrastructure was the key to the country’s recent dominance in broadband access and Internet usage, the provision of cost-effective xDSL services on existing copper cable telephony infrastructure. South Africa has been slow to leverage the installed telephone cable capacity for broadband and other ICT services derived from it, even after studying and supporting the concept of Local Loop Unbundling (LLU) that is critical to its success. (DoC 2007).

<table>
<thead>
<tr>
<th>Country</th>
<th>Population (Millions)</th>
<th>% Urban Population</th>
<th>GDP/Capita (US$ Current Prices)</th>
<th>Telephone Density (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. Korea</td>
<td>48</td>
<td>81</td>
<td>1679 16444 10% 1.5 49.5</td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td>45</td>
<td>73</td>
<td>1368 2669 3% 2.8 16.8</td>
<td></td>
</tr>
<tr>
<td>Thailand</td>
<td>63</td>
<td>32</td>
<td>696 2710 6% 0.3 11.0</td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td>73</td>
<td>67</td>
<td>2235 7110 5% 1.0 25.9</td>
<td></td>
</tr>
<tr>
<td>S. Africa</td>
<td>48</td>
<td>59</td>
<td>2764 5166 3% 4.1 10.0</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Key data similarities between the benchmark countries. GDP/Capita expressed in current US$ prices; ICT data from ITU World Telecommunications ICT Indicators 2007 (ITU, 2007); demographic data from HDR (2007/2008)

<table>
<thead>
<tr>
<th>Country</th>
<th>Internet Users (%)</th>
<th>Mobile phones (%)</th>
<th>Fixed phones (%)</th>
<th>PCs (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colombia</td>
<td>0.2 10</td>
<td>0.7 48</td>
<td>10 17</td>
<td>2 4</td>
</tr>
<tr>
<td>S. Korea</td>
<td>0.8 68</td>
<td>3.7 79</td>
<td>42 49</td>
<td>11 53</td>
</tr>
<tr>
<td>S. Africa</td>
<td>0.7 8</td>
<td>1.3 72</td>
<td>10 10</td>
<td>3 8</td>
</tr>
<tr>
<td>Thailand</td>
<td>0.1 11</td>
<td>2.2 48</td>
<td>6 11</td>
<td>1 7</td>
</tr>
<tr>
<td>Turkey</td>
<td>0.1 15</td>
<td>0.7 60</td>
<td>21 26</td>
<td>1 6</td>
</tr>
</tbody>
</table>

Table 2: Basic ICT Statistics in the four benchmark countries. Data derived from ITU ICT database 2007 (ITU, 2007).

Table 2 provides basic ICT data for the four benchmark countries in years 1995 and 2005. Except for mobile telephony, South Africa’s ICT growth was lowest in the group. One disturbing feature of the data in Table 2 is that South Africa’s penetration of Internet users, fixed line telephones, and personal...
computers aligns with the country’s racial demographics. It is mainly White South Africans and a small but growing number of Black “First Economy” South Africans who enjoy the services of these knowledge-rich ICTs. The “Second Economy” population depends almost entirely on costly mobile telephone services for their information needs. Recent research by Skuse and Cousins (2007) concluded that the success of mobile telephony in South Africa had introduced an unexpected negative outcome: chronically poor mainly rural communities developed dependencies on the mobile telephone services, while their high usage fees effectively deepen chronic poverty.

High ICT user prices have been recognised by all South Africans and their international development partners as significant barriers to national development, as noted by South Africa’s Minister of Communications, Dr. Ivy Matsepe-Casaburri in her budget vote speech (Budget Speech, 2008). Figure 3 quantifies the depth of this barrier to development, and verifies the concern expressed by Parayil (2005).

<table>
<thead>
<tr>
<th>Country</th>
<th>1st Economy %</th>
<th>2nd Economy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Korea</td>
<td>51.4%</td>
<td>13.3%</td>
</tr>
<tr>
<td>Thailand</td>
<td>28.2%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Turkey</td>
<td>3.2%</td>
<td>2.2%</td>
</tr>
<tr>
<td>South Africa</td>
<td>5.2%</td>
<td>13.3%</td>
</tr>
<tr>
<td>Colombia</td>
<td>3.2%</td>
<td>5.2%</td>
</tr>
</tbody>
</table>

Figure 3: Mobile Telephony (2006) and Internet Access Price Baskets (2005).

3.1 South African ICT Policy and Regulatory Environment

Historically inherited race-based inequalities clearly dominate nearly all South African discourse on human development and its complex multiple facets. The national ICT policy and regulatory processes have not been immune. Costly and often damaging compromises between the need to ensure Black South African ownership and management of the critical ICT industry within the framework of national Black Economic Empowerment (BEE) policies, and the need to develop the nation as a whole have had to be made. A recent study by the Centre for International Development at Harvard (CID), a partnership that includes the National Treasury of the Republic of South Africa, poses the research question: Is Black Economic Empowerment a South African Growth Catalyst? (Or Could it Be…) (Andrews 2008). Andrews’ conclusions note numerous disappointing outcomes of the BEE policy, but remain positive that it could be a growth catalyst. One key conclusion proposed is that government should focus its BEE policy on the middle and bottom of the economy, the primary focus and helm of this paper.

A classic example of the conflict between implementation of the BEE policy and national poverty alleviation through ICT development is the recent case of broadband wireless spectrum allocation. On 17th June 2008, South Africa’s state regulator ICASA published a strict BEE policy for awarding the valuable 2-3 Ghz spectrum (ICASA, 2008). This required applicants to prove 51% Black South African ownership with an emphasis on women, unspecified Black South African participation in management and control, and skills development in line with national BEE policies (ICASA, 2008).
Just 3 days later on 20th June 2008, the USA Federal Communications Commission (FCC, 2008) published its approach to allocating this same spectrum. The FCC planned to license a single nationwide operator that would be obliged by regulation to reserve 25% of its network capacity for free pro-poor broadband access, 50% coverage of needy Americans in the first 4 years, rising to 95% by the time the license expires after 10 years (FCC 2008). The contrasts are stark. South Africa, a developing country chose to use invaluable broadband wireless spectrum as a political tool to reverse historical injustices, while the United States of America, arguably the biggest economy on earth, chose to use the same technology to bridge its socio-economic divide. A useful research question could be: of the two approaches, which one was more likely to reduce South Africa’s severe socio-economic inequalities and promote Black South African participation in the economy?

The complex uniquely South African dilemma of balancing historical injustices with future growth is beyond the scope of this discussion; however, valuable insights can be obtained from numerous studies and analyses. For example, Hodge (2008) provides a detailed fifteen-year review of South Africa’s telecommunications policy, observing the numerous failures to solve the pro-poor access and price conundrums. Horwitz and Currie (2007) provide a critical appraisal of the country’s ICT policy dynamics, highlighting in some detail the conflicts between liberalization, privatization and BEE ownership, and the costly errors that resulted from these conflicts. Fleming (2002) explores the role of ICT in developing the democratic processes in South Africa, and Brown and Brown (2008) propose a radical shift in South Africa’s definitions and perceptions of ICTs so as to enable policy-makers to focus more on their human development utility.

4 A RESEARCH FRAMEWORK FOR BOTTOM OF THE PYRAMID ICT DEVELOPMENT IN SOUTH AFRICA

The brief quantitative review of human development in South Africa presented in preceding sections of this paper justifies an urgency to take corrective action. Action research has been recognized as a means of ensuring relevance and applicability of ICT research to practice (Baskerville and Myers, 2004). A research framework for carrying out an action-based study is outlined in this section. The components of the framework are neither original nor unique. They draw liberally from the vast quantity of knowledge about ICT and development that has been accumulated throughout the nearly 250-year evolutionary history of electronic ICTs. They also benefit from the experiences of fully developed countries, newly industrialised countries, developing countries, and failing or failed states. An extensive database of knowledge about ICTs for development, and the successes and failures of different nations has been compiled to support the research framework. The major departures from traditional approaches to ICT research and development are:

- The primary focus is on the complex components of human development: the information and knowledge needs of South Africa’s socio-economically marginalized communities, and of the people tasked with supporting their development efforts. Prior research has shown the weaknesses of approaches that are techno-centric, where the focus is on technology rather than people (Roode et al., 2004).

- The complexity of human development demands that multidisciplinary and multi-institutional approaches be adopted. Each human development challenge, be it economic deprivation, knowledge acquisition, health related challenges, or inequitable distribution and high prices of ICTs are intricately intertwined. Fragmented approaches to reducing these challenges have generally been total or partial failures in most African countries. The low comparative levels of human development on the continent are clear proof of these failures. An alternative integrated approach is proposed in the framework, one in which research and development in all human development challenges are conducted simultaneously, and the ICT support solutions built from the results of this research, rather than taking a technological determinist view of ICT (Licker, 2001);
A balanced “Evolutionary” and “Clean Slate” approach to technological innovation for BOP-ICTs will be used. The challenges facing human development at the BOP have an uncanny resemblance to the challenges facing the evolution of the Internet (Schwarz da Silva 2007, Sung-Su Kim et al, 2008) - evolutionary tweaking versus radical redesign to overcome significant current problems. In this research framework, a balance between using traditional ICT networks, evolving networks in the form of Next Generation Networks (NGN) that tend to cater exclusively for the needs of the economically endowed, and totally new innovations for pro-poor ICT applications will be developed. The research emphasis will be on the information and knowledge needs, not on the technological means of delivering them, as advocated in techno-centric approaches (Roode et al., 2004);

A “bottom-up” approach will be used. The research focus will be on people residing at the BOP, the full range of sociological factors that influence their human development, their capacities to adopt, afford and use advanced ICTs, will be the innovation drivers. A strategy to integrate the “bottom-up” solutions with current and planned “top-down” development solutions will be a key factor of the framework. Purely top-down approaches to ICT strategy development are known to have limited success (Segars and Grover, 1999).

In developing the research framework, the complexities of creating and managing a multidisciplinary multi-institutional research and development team have not been underestimated. They have been accommodated in the research schedule with allowances for continuous review and adjustment. Thus far, the research process has been initiated, the initial proposal has been drafted and initial communications and canvassing support from potential researchers and funding agencies has begun. Intricate details of the emerging framework are beyond the scope of this paper, however, an outline is provided in Annexure 1. It is anticipated that a Steering Committee will have been formed by mid June 2009, and will be tasked with setting up the required management and research teams, their operational structures, and initiating the four phases of research activities over the planned duration of three years. A key outcome of the research will be to chart the path for continuous ICT for BOP development in South Africa, accounting for changing socio-economic environments and technologies (Avgerou, 2008).

5 CONCLUSION

This paper provides a short background on the key factors that should guide a human development-oriented ICT development strategy in South Africa. The challenges are immense, but they can be eased significantly by focusing attention on South Africa’s BOP market, or the “Second Economy” as it is popularly known locally. A concerted single-minded search for ICT solutions for this population group that are both effective and affordable is not beyond the national capability, especially with the support of modern ICTs. A detailed description and discussion of the proposed research framework aimed at addressing these challenges is beyond the scope of this paper, but expected and speculative outcomes can be discussed in relation to the fictitious Wembe family used to introduce the paper:

Soweto, South Africa, mid-2012: “Bra” Peter Wembe is now 69 years old. He travels less now, but enjoys the growing success of his grandson Nelson, whose growing international reputation as an avant-garde jazz saxophonist with a distinct African flavour reminds many global jazz fans of jazz legends like Charlie Parker and John Coltrane. The family has a more comfortable income: they extended their “Spaza” shop to include a jazz café with live music at the weekends, which is becoming a popular tourist destination. The café is managed by Peter’s daughters, and doubles as an Internet café during the day. They use a low-cost broadband service to source goods and services for their business. This service helped launch Nelson’s international career: he promotes his and Granddad’s original music compositions and performances through online streaming to the global jazz marketplace, and receives a growing number of lucrative live concert contracts as a result. They still dream of a recording studio, but this is getting more remote as real-time online opportunities emerge. Their
primary broadband access is derived from the excess capacity of the neighbouring supermarket’s fibre optical connections and servers. The supermarket was allowed by new regulatory provisions to follow the UK’s “Tesco Telecoms” model of reselling excess capacity to the local community (Tesco Telecoms, 2003). Peter’s Internet café servers provide fixed VoIP services to the homes of twenty neighbours who pay a fixed monthly fee of R50 for unlimited national voice services. An extended family relative maintains these and other similar services in the community. Peter thinks he could derive more revenues by reselling broadband, but that is in the pipeline waiting for his granddaughter to graduate with an IT diploma from the local polytechnic. Life for the Wembes has improved since the arrival of fast internet.

6 ANNEXURE 1: THE RESEARCH FRAMEWORK OUTLINE

<table>
<thead>
<tr>
<th>Project Title</th>
<th>Promoting the use of Information and Communications Technologies (ICTs) for Human Development from South Africa’s Bottom of the Pyramid (BOP).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lead implementing institution</td>
<td>Monash University’s South African Campus</td>
</tr>
<tr>
<td>Narrative Summary</td>
<td>Goal: To develop ICTs that support Human Development at South Africa’s BOP.</td>
</tr>
<tr>
<td></td>
<td>Purpose: To add new knowledge of ICT for HD and introduce BOP-friendly ICT networks and supporting policies and regulations.</td>
</tr>
<tr>
<td></td>
<td>Outputs:</td>
</tr>
<tr>
<td></td>
<td>1. Reports and publications.</td>
</tr>
<tr>
<td></td>
<td>2. BOP ICT infrastructure and services</td>
</tr>
<tr>
<td></td>
<td>3. Pilot and small scale ICT projects</td>
</tr>
<tr>
<td></td>
<td>4. Equipment and network innovations for capex &amp; price reductions.</td>
</tr>
<tr>
<td></td>
<td>5. Business models developed for BOP SMMEs, in ICT</td>
</tr>
<tr>
<td></td>
<td>6. Macro and micro economic and business models</td>
</tr>
<tr>
<td></td>
<td>7. ICT policy &amp; regulatory development</td>
</tr>
<tr>
<td></td>
<td>8. International R&amp;D partnerships established.</td>
</tr>
<tr>
<td>Objectively Verifiable Indicators</td>
<td>Increasing use of ICT for HD by BOP citizens.</td>
</tr>
<tr>
<td></td>
<td>Published scientific papers; increased use of ICT for HD.</td>
</tr>
<tr>
<td></td>
<td>Functional BOP ICT networks, services, applications &amp; content</td>
</tr>
<tr>
<td></td>
<td>Functional BOP ICT projects</td>
</tr>
<tr>
<td></td>
<td>ICT services affordable by BOP citizens</td>
</tr>
<tr>
<td></td>
<td>New entries by SMMEs to ICT business sector</td>
</tr>
<tr>
<td></td>
<td>Reports, publications, SMME ICT businesses</td>
</tr>
<tr>
<td></td>
<td>Reports published and submitted to relevant authorities.</td>
</tr>
<tr>
<td></td>
<td>Number of international R&amp;D partnerships</td>
</tr>
<tr>
<td>Means of Verification</td>
<td>ICT growth indicators, e.g. ICT Development Index (IDI).</td>
</tr>
<tr>
<td></td>
<td>Number of published papers; statistical reports</td>
</tr>
<tr>
<td></td>
<td>Final Project Report, sectorial reports and IDI indicators</td>
</tr>
<tr>
<td></td>
<td>Project and sector reports by users &amp; developers</td>
</tr>
<tr>
<td></td>
<td>Project reports and market growth.</td>
</tr>
<tr>
<td></td>
<td>Company registrations, reports.</td>
</tr>
<tr>
<td></td>
<td>Final project reports and development proposals.</td>
</tr>
<tr>
<td></td>
<td>Final project reports.</td>
</tr>
<tr>
<td></td>
<td>Project reports.</td>
</tr>
<tr>
<td></td>
<td>Correspondence &amp; reports</td>
</tr>
<tr>
<td></td>
<td>Accepted research proposal, equipment audits, communication systems tested.</td>
</tr>
<tr>
<td></td>
<td>Project reports</td>
</tr>
<tr>
<td></td>
<td>Evaluation workshops and reports.</td>
</tr>
<tr>
<td>Assumptions and Risks</td>
<td>Access to local and international databases and statistics.</td>
</tr>
<tr>
<td></td>
<td>Speed of ICT adoption at the BOP relative to R&amp;D schedule, levels of investments in Bop-ICTs</td>
</tr>
<tr>
<td></td>
<td>Policy, regulatory &amp; commercial interest barriers</td>
</tr>
<tr>
<td></td>
<td>Policy and regulatory entry barriers, financing</td>
</tr>
<tr>
<td></td>
<td>Effective R&amp;D partnerships, entry barriers</td>
</tr>
<tr>
<td></td>
<td>Business entry, regulatory barriers, costs</td>
</tr>
<tr>
<td></td>
<td>Policy &amp; regulatory barriers</td>
</tr>
<tr>
<td></td>
<td>Political willingness to promote BOP market development.</td>
</tr>
<tr>
<td></td>
<td>Budgetary resources, international collaboration.</td>
</tr>
</tbody>
</table>

Activities:

1. Phase 1: Finalise R&D proposal & first planning workshops |
   Working draft research project proposal, Steering Committee functional |
   Correspondence & reports |
   Budgetary provision for phase 1 activities |

2. Phase 2: Establish R&D teams, management structures; detailed project planning, equipment procurement. Set up communications systems |
   Final research plans; equipment delivery; communications networks and website tested. |
   Accepted research proposal, equipment audits, communication systems tested. |
   Project funding. |

3. Phase 3: Implementation |
   Research teams dispatched and research in progress |
   Project reports |
   Project funding. |

4. Phase 4. Project completion and follow-up strategy development |
   Evaluation workshops and reports. |
   Final Project Reports. |
   Project funding. |
References


Budget Vote Speech (2008). The Minister of Communications Dr. Ivy Matsepe-Casaburri delivered at the National Assembly on 3 June 2008. The South African Department of Communications.


DoC: Department of Communications (2007). Local loop unbundling: A way forward for South Africa. The South African Department of Communications.


RECOGNIZING WORK PRIORITIES AND TASKS IN INCOMING MESSAGES THROUGH PERSONAL ONTOLOGIES SUPPLEMENTED BY LEXICAL CLUES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0340.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>End user, Information overload, Ontology, Work performance</td>
</tr>
</tbody>
</table>
RECOGNIZING WORK PRIORITIES AND TASKS IN INCOMING MESSAGES THROUGH PERSONAL ONTOLOGIES SUPPLEMENTED BY LEXICAL CLUES

Tagg, Roger, Gandhi, Prashant and Srinivasan Kumaar, Raaj, University of South Australia, School of Computer and Information Science, GPO Box 2471, Adelaide SA 5001, Australia, roger.tagg@unisa.edu.au, prashant.gandhi@gmail.com, skraaj@gmail.com

Abstract

Email has been recognized as one of the chief sources of the current epidemic of information overload. The work described in this paper is part of an ongoing project to re-focus the attention of information workers on “what needs to be done” rather than “what someone wants us to read”. The approach is to base a software agent that pre-processes incoming email on the user’s “personal ontology”. The ontology used in this work contained two facets, namely Work Topics and Task Types, and also included lexical “clues” whereby, if such a word or phrase is found in the text of a message, the system can infer with a certain probability that a particular ontology concept (class or instance) is relevant. The paper describes the prototype tools developed and some initial trials. Further trials and improvements are planned.

Keywords: Information Overload, Task Orientation, Ontology, Lexicon.
1 INTRODUCTION

The concept of information overload has been recognized for more than a decade (e.g. Whittaker et al 1996). Suggested causes range from trends in the organizational climate towards mistrust, risk aversion and micro-management, e.g. (Allen and Wilson 2003) to the availability of too much information and of the IT facilities themselves, e.g. (Alberts 1996).

Many authors, e.g. (Mulder et al 2006) have stated that email is the biggest contributor to information overload. A workshop discussion in (Whittaker and Sidner 2005) reported, for example, that in the previous three years, the number of email messages has increased eight fold. Some organizations have instituted email-free days (Wakefield 2008) and some individuals have given up email as a main channel of communication, e.g. (Suarez 2008).

The search to find ways to use IT to relieve information overload is becoming more urgent with every year that passes. Without relief, information workers will find that they have less and less time to carry out the functions for which they are actually employed, and decisions they make may be adversely affected (Sutcliffe 2006). The authors of the current paper, following a series of earlier experiments (Tagg 2007) under the theme “Virtual Private Secretary”, have over the year 2008 developed a prototype that aims to address individual overload, primarily that arising from email. The principles of our prototype are that:

- It should be activity-oriented, i.e. it should be concerned with what we as individuals most need to do, rather than what email senders want us to read.
- It should reflect the collaborative nature of most information work, but from a federal, bottom-up viewpoint rather than a top-down “one group, many members” perspective.
- It should make use of the information that a computer system can provide about an individual’s work topics, contacts, priorities etc.
- It should not be so complex that an ordinary user will be deterred by the effort of setting up and maintaining it.

The remainder of this paper is organized as follows. We first describe our rationale for the approach we have taken and the stages of the work we undertook. Next, we give the details of our prototype system for recognizing work priorities and tasks, and some comments on the evaluation we performed. Following this, we discuss our approach for users to maintain their personal ontology including the lexical clues. We then outline some of the related work that is addressing similar objectives to ours, and this is followed by some discussion and comparison. We then conclude with some suggestions for future work.

2 RESEARCH DESIGN

2.1 Justification for using a prototyping methodology

According to Clay Shirky, reported in (Wagner 2008), Information Overload has been with us for so long that it should be seen as a fact of life, rather than a new problem; instead, the real problem is that our filters are no longer adequate. The work described in this paper, therefore, concentrates on one activity where, in the authors’ experience, better filtering seems to be particularly needed; namely that of incoming email messages.

The literature shows a wide range of theories for reducing information overload, ranging from intelligent software agents to organizational restructuring, personal goal re-appraisal and even speed reading. However so far there has been little evidence of success. One symptom is the lack of major improvements (Hall 2004) in task management facilities in everyday tools such as MS Outlook, or in
web browsers. We have therefore taken the approach of building a number of prototypes, based on relatively heuristic algorithms, to see what working with such prototypes feels like to users. A similar approach has been taken by other researchers in this field, for example the group at Xerox PARC (Bellotti et al 2007).

2.2 Review of alternative approaches to a prototype email pre-processor

Current commercial tools, such as Outlook, typically offer an email filtering facility using rules, with a wizard that helps the user express the desired rule. However Ducheneaut & Bellotti (2001) suggest that rule-based filters are not very widely used, suggesting that this feature is either not useful or is difficult to use. The same authors argue that filtering is mainly useful for filing schemes based on the identity of the sender, and offers weaker performance when required to deal with more complex categorization. Crawford, Kay and McCreath (2002) suggest that the process of composing rules is, for many users, cognitively demanding.

Another approach has been to extend the idea of spam filtering, where an approach based on Bayesian classifiers and machine learning is usually used. The application of this approach for categorizing emails into multiple categories (rather than just spam / not spam) can be found in studies conducted by Maes (1994), Segal and Kephart (1999), Mock (2001), and Crawford, Kay and McCreath (2002), amongst others.

Some other approaches address the recognition of tasks and priorities, rather than the categorization of emails. In current practice, the receivers of emails are highly dependent on the senders. Outlook, for example offers “follow up” and “high priority” flags; the sender can also send a formatted “task” as a message – though this does not seem widely used. Bellotti and her colleagues (2007) have approached tasks partly through recognizing threads of emails, coining the term “thrasks”. The more mundane reality, though, is that most users have to fill a form manually in order to turn a message into a task that would appear in a “to do” list.

Although historically, many proposed methods for email pre-processing have been based on Machine Learning, we have not followed this approach. Our reasons were:

- Machine Learning techniques depend on acquiring sufficient training data in order to build a competent classification system. The “slow start” problem, by which classifiers slowly build-up their competence as more examples are provided, may be more intense when a user has, say, tens of different categories rather than “spam/not spam”. A significant amount of training data, with user supervision, would therefore be needed in order to acquire strong classification results from the beginning.

- Making use of a personal ontology takes account of what is known in advance about the user’s work structure. Ontologies have recently gained increased attention within other research areas, and we believe that their application to email pre-processing is worth further exploration. However we recognize that, because user’s work structures change over time, a hybrid approach will be the most effective in the longer term.

2.3 History of previous prototypes developed within the authors’ research group

VPS (Virtual Private Secretary) is an ongoing project within this university’s Information Systems Laboratory, having started in 2005. Some of the earlier efforts on this project are described in (Tagg 2007). An overall view of the architecture of the current and planned components of the VPS project is shown in Figure 1. The upwards blue arrow represents the direction in which we think the user’s task interface ought to develop.
The numbers in the diagram refer to some of the prototypes developed within the VPS project up to the end of 2007:

1. An email pre-processor “TaskMail” that reformats emails, which contain one or more key words in a defined set, as tasks in a “to-do” list with links to the message body and any attachments.
2. A middleware architecture for aggregating tasks from different source applications such as Workflow, Project Management and ERP (Enterprise Resource Planning).
3. A “myTasks” portal that gives a unified to-do interface to tasks that come from different sources (e.g. workflow, project management, ERP system) via a unified TaskXML (Tagg 2007).
4. An Outlook add-on for importing tasks formatted in TaskXML into the Outlook task facility.
5. A “drag and drop” ontology editor called EzOntoEdit that allows users to import ontologies (OWL or CSV formats) into a personal OWL file and amend it graphically wherever possible (Einig et al 2006).
6. The VPS email categorization agent that is described in this paper, which uses the personal ontology maintained in 5 above.

In addition to the above, we also attempted in 2005 to support “sender-assisted” categorization, by means of a generator of “contact us” pages based on the personal ontology of the recipient, senders of messages being required to enter certain fields that give message context.

A further project in 2007 added into EzOntoEdit the facility for including “clues” in the user’s ontology. A clue is defined as a lexical string which, if detected in an incoming document, indicates (with a certain probability) that the document is relevant to a particular ontology category.

### 2.4 Attempting to ease the burden on the user in maintaining a personal ontology

However initial evaluation of the clue-enhanced EzOntoEdit ontology editor suggested that users would probably be discouraged by the effort required to fill in the forms needed to specify these clues. We were also concerned about the user’s workload in reacting to changes in the pattern of his or her work.
We therefore investigated the use of an automated text analysis tool, Leximancer (Smith 2003) which has been used in a number of domains to suggest both “Themes” (what a document is about) and “Concepts” (words or phrases most strongly associated with a Theme).

We applied Leximancer to collections of emails in two different ways:

- Documents created by concatenating all the emails - received or sent - that had been already archived by a user, into each of around 20 major categories;
- A single concatenated document of all the emails in the same user’s inbox that had not yet been archived.

The first approach aimed to detect significant differences in the frequency of words or phrases in the different categories, while the second approach tried to deduce themes *ab initio*; the desired number of themes being set as a parameter. While the results made some sense, we judged that:

- The influence of the chosen stop word list was very high.
- The effect of email senders including the text of previous emails in a thread was variable and unpredictable.
- Files of concatenated emails behaved quite differently from the single longer documents for which Leximancer had been designed.

We therefore decided to abandon this approach for the time being. This research is described in more detail in (Tagg, Lalwani and Srinivasan Kumaar 2008).

### 2.5 Research design approach eventually taken

We therefore split our work for the remainder of 2008 into two parts:

- To build a new prototype email pre-processor, this time one that separated the ontology classes and instances (in two dimensions, “work topics” and “task types”) from the lexical clues used; and
- To develop a better, more user-friendly approach to capturing the lexical clues from the user’s normal activities.

We would later bring these two parts together and test their effectiveness in predicting the work topic and “taskiness” – following a term coined in (Macbeth 2008) - against the manual judgment of users.

### 3 THE VPS CATEGORIZATION AGENT

#### 3.1 Mode of operation of the prototype

The mode of operation of our prototype is described in Figure 2 below. The letters show the high-level sequence of operation.

- Ontology maintenance (A1-A3) is discussed in section 4 below.
- The OWL ontology is converted to Oracle (B) for the agent’s internal purpose, although the definitive version is maintained in OWL XML format.
- Messages arrive in the user’s mailbox asynchronously (C), however …
- The email agent only reads the inbox at set intervals (D), during our trials, this was 5 minutes.
- Pre-processing (E) involves stripping off message bodies and attachments into separate web-accessible files, and tokenizing the message body and other fields.
- The context recognition stage (F) uses the ontology to tag the messages in one or more “work topic” or “task type” categories, the default being “uncategorized”.
- The user views the categorized task list, derived from a Tasks database (G1) in a “MyTasks” portal, with links to the original message text bodies and attachments on a web server (G2).
Our approach to categorization is a very simple and heuristic one.
1. Scan an incoming document for occurrences of any of the clue strings (words or phrases) in the user’s ontology.
2. Apply the concept of Indication Strength, which is the subjective probability that presence of the clue’s string indicates relevance to an ontology concept.
3. The Weighted Indication Strength (WIS) for each matched clue is then calculated based on multipliers that reflect the number of occurrences of the clue. These range from 1.2 (or a 20% raise of the indication strength) for 2 occurrences of the clue, to 1.75 for 10 occurrences or more.
4. Aggregate all the clues that refer to the same ontology category (class or instance) and add their WIS values.
5. If the total of the WIS values matches or exceeds the threshold value (default is 1.0), insert a row in the “categorized” table, and include URI links to the original message and any attachments.
6. Repeat for each category that is indicated by clues in this document.
7. Repeat for all documents received since the last clock event (currently, at 5 minute intervals).

The procedure is performed twice; the first time with respect to Work Topics, the second with respect to Task Types. The Task Types we used were as in Table 1 below. In our trials, we removed the middle “invitation” priority level for simplicity, and reallocated the Task Types as shown by the arrows.

<table>
<thead>
<tr>
<th>Priority</th>
<th>High</th>
<th>Invitation</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“Definite task”</td>
<td>“Conditional task”</td>
<td>“For information only”</td>
</tr>
<tr>
<td></td>
<td>“Reply only” (explicit request)</td>
<td>“Questionnaire” -&gt;</td>
<td>“Private communication”</td>
</tr>
<tr>
<td></td>
<td>“Vote”</td>
<td>“Conference invitation” --&gt;</td>
<td>“Uncategorized”</td>
</tr>
</tbody>
</table>

Table 1. Task Types used in the prototype.
The user’s ontology can be amended between clock events, and if it has changed, the new ontology will automatically be used the next time the in-box is processed.

The myTasks portal displays a scrollable table of inferred tasks, sorted either by Task Type within Work Topic, or vice versa. By default, only the high-priority tasks are displayed, but the user can choose to look at the low-priority list if he or she has time to do so. There is also an option to display the original email message body for a task on a table that has been highlighted by the user.

3.2 Evaluation of the prototype

In the time available we were only able to test the system against one user’s manual categorization. However this was done twice, once for an early version of the prototype and again for a later version when the task inference had been improved (by adding to the ontology some clues indicating template-based task-requesting emails that were repeatedly received). Both trials involved around 100 emails forwarded from the user’s live inbox. For Work Topics, performance was consistent in both versions with precision over 80% and recall 67%. Task Type categorization was less successful – precision was only 43% in the first trial and 62% in the second; recall was around 55% in both.

On reflection we felt that we still need to improve the ability to accurately infer tasks from the text of emails. Three particular problems appeared to contribute to these results

- a proportion of definite tasks were not recognized as such, due to the seemingly inexhaustible circumlocutions in the English language of asking for something to be done
- we did not have in the prototype a means of recognizing dates, which appear to be a good indicator of taskiness
- our “Conditional Task” category was too wide and did not distinguish between conditions that were quite likely to be true and those that were extremely speculative.

3.3 Results of an informal questionnaire carried out during the project

Halfway through the project, we demonstrated the prototype at a “project fair” to a mixture of students, academic staff and project sponsors from industry, and asked them to complete a questionnaire; 43 people responded, out of 70-80 people attending. We do not claim that this is a representative sample of users in the wider world, but we were interested to gauge initial reactions from people who saw our prototype working.

Over half the respondents reported that they regularly juggled 6 or more work streams. Deadlines were the chief driver to carry out tasks. People received 75% of their tasks through email, and they frequently “batch” tasks so that they get a run of doing similar tasks together. 80% declared an interest in trying a tool such as our prototype, although most said that they would like to keep the ability to look at the traditional email inbox view. The feedback suggested that we need to look more closely at peoples’ personal information habits and preferences.

4 GENERATING A PERSONAL ONTOLOGY AND CAPTURING LEXICAL CLUES IN VPS

This section details the work done on the ontology maintenance side of the project. Most authors, e.g. (Sheth et al, 2003) take the view that building an ontology, especially a complex one, is not a task for the average end user. We therefore need to find a mechanism by which the user can, relatively easily, generate and maintain a simple personal ontology and capture the lexical clues that indicate that a document relates to an ontology category. The EzOntoEdit tool (Eimig et al 2006) mentioned above is oriented to importing a pre-existing ontology from a work group or a domain, then tuning it to one’s personal preference, rather than creating it from scratch.
The problem remains of how to capture the lexical clues. Figure 3 shows a range of modes of capturing clues and the ontology structures to which they relate.

![Building ontology structure manually](image)

**Figure 3. Various Modes of Generating Lexical Clues for an Ontology.**

It has been rightly pointed out that, by adopting an ontology-based approach, we might be simply transferring the user’s workload to manually maintaining his or her ontology, instead of ploughing through a mass of emails.

While totally automatic maintenance of lexical clues seems unrealistic, it might be possible to do regular text mining of samples of recent documents and messages, and for the user to interpret those results. However our own experiences with Leximancer (Tagg et al 2008) suggested that such a system might turn out to be very complex. Semi-automatic approaches have been proposed, e.g. (Wang et al 2006, Fortuna et al 2006) but there are not yet many reviews of their user acceptability over an extended period. In our study, we decided on an intermediate - but fundamentally semi-manual - approach, assisted only by the importing of “blueprints” or “stereotypes” of the clues used by users doing similar tasks, and by the process of “snipping”.

![SnipCat main window](image)

**Figure 4. SnipCat main window.**
“Snipping” in our prototype is derived from the ideas of (Snip!t 2008) and from MS Internet Explorer 8 Accelerators (Microsoft 2008). A user highlights a word, phrase or name from any document he or she is working on (e.g. an email, Word document, spreadsheet, PDF file etc) and on the “Copy” command that text appears automatically in the prototype’s window. In Figure 4 above, the user has just highlighted and copied the word Andrew; that text appears in the top left window of our SnipCat (short for Snip Categorization) tool.

The Categories window allows the user to browse the ontology hierarchy and highlight the appropriate category. The Location window allows the user to specify where the text should appear; the options are From, Subject, Body and Anywhere. Indication Strength allows for “maybe” clues – the default is a probability of 1 (slider at extreme right). In the top right window the existing clues for the highlighted category Campus Related are shown. If the green tick is clicked, a Clue instance with Andrew, Subject and 0.9 will be added. Note that “Discriminator” is not the same as Clue, since Andrew might have other meanings, for example if it is in the From location.

One can have a clue that points to a whole Class instead of a particular Instance; currently this is done by highlighting the standard Instance None. Likewise, Other indicates assignment of the string to the next higher level Class.

SnipCat also allows for “Bulk Snipping”. If that button is pressed, the user is prompted for the path of a tabular file (currently MS Excel but potentially also a database), the Sheet (or Table) and the Name of a Column. A typical use would be to import list of contacts such as email addresses, surnames etc. These strings are all set up as clues to the ontology concept chosen.

SnipCat also allows for additions to the ontology, but only at the Instance level. Changes to classes currently require the user to switch to the EzOntoEdit tool mentioned in section 2 above, or some other OWL editor.

SnipCat was tried by 5 users. All found the method fairly easy to use, although to really judge whether or not people would use it, trials would have to continue over a long period. There was a wide range of complexity in peoples’ ontologies; the smallest had fewer than 10 classes with 20 instances and 30 clues, while the largest had:

- 6 work topic classes with 24 subclasses and 98 instances
- 8 task type classes, no subclasses and 27 instances
- 261 clues, split into 25% in the location From, 30% in Subject, 16% in Body and the rest in Anywhere
- The distribution of clues by indication strength (probability) was 67% for a probability of 1.0, 25% for 0.7 to 0.99, 8% for less than 0.7.

5 RELATED WORK

In recent years a number of major prototypes, several of them European funded, have been built addressing similar issues. The systems listed below all follow, to some extent, an ontology-based approach.

- DELOS (Katifori et al 2008) incorporates a PIMS (Personal Interaction Management System) based on a Personal Ontology. This group has explored alternative ontology visualization methods. Their ontology is significantly fuller in detail than that described in the current paper, and they recognize that it is critical to find suitable views to prevent the task of maintaining the ontology from becoming too complex. However, they do not appear to use a clear separation of ontology concepts and clues.

- NEPOMUK (Sauermann et al 2008) includes the concept of a Semantic Desktop, driven by a PIMO (Personal Information Model), which appears to be a form of Personal Ontology. A number of standard classes are provided, and users can add more using a tool. However automatic categorization, in particular of emails, does not seem to be included.
Perfect Digital Assistant (PDA) (Schraefel et al 2008) has been developed by a consortium involving MIT, University of Southampton and Nokia Research Centre. As in our VPS project, it aims to simulate the work of a personal assistant. It is built on a RDF representation of personal knowledge, which is partly fed by activity log mining, but does not yet appear to support email categorization.

DYONIPOS (Rath et al 2008) also builds up a user model by mining logs from the whole range of a user’s frequently-used applications and operating system actions. The resulting context “ontology” is expressed as a hierarchical structure of Tasks, Event Blocks and Events. Tasks are recognized automatically while the user is working.

The Semantic LS system (Krishnan et al 2008) is being developed by IIT Kanpur, India but with European funding. It is targeted at a P2P cooperation model, but for a private group rather than an open internet. A DKS (Domain Knowledge Model) is maintained as a taxonomy in OWL. It does not yet address email documents.

The following noteworthy systems, which are not ontology-based, should also be mentioned.

TV-ACTA (Bellotti et al 2007) is a development of work that has been going on at Xerox PARC since around 2001. TV (Task Vista) is a “to-do” list, while ACTA is an Activity Centered Task Assistant that works as an Outlook add-in. The system takes account of repeated “workflow-like” activity patterns. Earlier work (Taskmaster) looked at threads of messages related to tasks.

SCOUT (Sow et al 2006) was developed by a group within IBM. This appears to be primarily rule-based; the “context” referred to is primarily physical location. Their evaluation showed good results, but they only analyzed those messages that were generated automatically by software, which were only a small percentage of the total.

SmartMail (Corston-Oliver et al 2004), is a Microsoft Research development that “presents the user with a task-focused summary of a message. The summary consists of a list of action items extracted from the message. The user can add these action items to their ‘to do’ list.” Natural language processing is used.

Chandler (Chandler 2008) is a cross-platform personal information management tool that streamlines the links between emails, calendar appointments and tasks. It was originally designed by Mitch Kapor of Lotus 1-2-3 fame, and supported by the Open Source Applications Foundation. It appears to be one integrated PIM that is generally available.

It seems clear that the major software vendors are seeking a way forward in this area, but may be averse to the risk of prematurely launching a major enhancement to popular systems such as Lotus Domino/Notes and MS Exchange/Outlook. The typical user has yet to see any significant products addressing any of context recognition, task recognition or reducing the fragmentation that arises from having to use different applications for each different file type.

6 CONCLUSIONS AND FUTURE WORK

Most of the prototypes emerging from the projects listed above have been wider-ranging and are further developed than our own. We have limited ourselves to addressing email overload problems, partly because of resource and time limitations, and partly because we feel that email is the single most pressing problem.

Our own prototypes have many shortcomings; in particular, our ontology structure as it exists today is little more than a taxonomy, and is-a and part-of relationships are not always strictly distinguished. Our recognition algorithm is very simplistic and needs testing with more variations over more data. Its performance in inferring tasks from emails is so far only moderate.

Our evaluation process was very limited, due mainly to the strict deadlines imposed by the student authors’ thesis and project timetables. However it is worth commenting that it seems very difficult to obtain reliable evaluation of how tools - like those proposed here – would be received by real users in overload situations. Many related projects have used teams of student interns, but these may not be
representative. However the administrative and ethical challenges of testing, using active employees in a cross section of industries, seem daunting.

Our contributions have been:

- to investigate the idea of clue-enhanced personal ontologies
- to test the principle of using these for pre-processing (as a personal assistant might do) emails and potentially other documents
- to explore some of the issues in inferring tasks from incoming emails
- to test an alternative approach to capturing clues and instances in personal ontologies

We recognize that we have only scratched the surface of what is needed if these ideas were to become part of a widely usable tool with “critical mass”. Some of our current ideas for future development are:

- General improvements to our personal ontology facilities, e.g. incorporating WordNet
- Building a library of blueprints (both basic ontology and lexical clues) for common human roles
- Recognition of dates in text, especially where these are deadlines and therefore indicate a task
- Recognition of the relevance of a message to a specific activity within a work topic
- Use of the context recognition system as a component in a system for annotating observed activities in a grounded theoretic approach to group work (Blackburn et al 2007)
- Integration of non-email sources of tasks into the task database, e.g. “actions” from minutes of meetings that are stored in shared folders
- Investigation into improving the contribution of email senders towards categorization for the message receivers.
- An investigation of habits and preferences in personal information management for a range of users in different work environments.

However we feel that we have made a start on a potentially valuable approach to context recognition that strikes a balance between, on the one hand, a high-technology approach that is hard to implement and maintain, and on the other a more easily manageable – but possibly simplistic - solution.

References


System, CHI 2008 Workshop on Personal Information Management, Florence, Italy


Macbeth, M. (2008) Tasks and Time Management in Outlook, Blog entry, 26 October 2006,


Wakefield, J. (2008) Turn off e-mail and do some work. BBC News.


THE EMERGENCE OF LANGUAGE CONSENSUS - INTENSIFYING LANGUAGE INTERACTION IN INFORMATION SYSTEMS DEVELOPMENT

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0638.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Information Systems Development (ISD), Face-to-Face Communications, interpretivist research, Ontology</td>
</tr>
</tbody>
</table>
THE EMERGENCE OF LANGUAGE CONSENSUS –
INTENSIFYING LANGUAGE INTERACTION IN INFORMATION
SYSTEMS DEVELOPMENT

Corvera Vargas, Marianne, Goethe University, Grüneburgplatz 1, RuW Building, 60323
Frankfurt, Germany, corvera@wiwi.uni-frankfurt.de

Abstract

As language is the venue for the concretisation of concepts and entities in information systems
development (ISD), language unambiguousness is understood as an ideal basis for the success of the
consolidation process in information systems development. In contrast, a fundamental postulation of
language theories is the inherent ambiguity of natural language. Therefore, the analysis of how
language consensus can be reached in an effective manner is of particular interest for collaborative
ISD. We claim that the effectiveness of ISD depends on the ability to manage the question of how
people deal with language in practice and reach consensus in the concrete ISD process. This
requirement leads us to the point of emergence of language: the language interaction setting. Hence,
we analysed the language interaction in an ISD project and identified different language interaction
levels which serve as orientation for specific interventions in an interaction setting, enabling an
intensification of consensus negotiation.

Keywords: ISD, Language Interaction, Emergence of Consensus, Communication.
1 INTRODUCTION

A central challenge in information systems development (ISD) is certainly the question how a coherent and meaningful model can be created out of the consolidation of different requirements and perspectives in the enterprise. As we understand ISD as a language development and formalisation process (Lyytinen 1985), the analysis of the consolidation process on a language level can provide an important contribution to this challenge. We argue that by characterising and analysing the use of language as an instrument for communication and for consolidation of an information system (IS) in the ISD setting, we can obtain helpful advice that enable us to enhance effectiveness in the emergence process of language consensus. Our main questions are: how language consensus in ISD processes is reached and how can we influence this process effectively? These questions lead us to the micro-level of ISD, the point where language consensus emerge, that is, the language interaction setting.

The relevance of the improvement of communication issues in ISD is obvious due to its impact on the ISD profitability. Alarming results of different requirements engineering surveys give an account of the failure arising in a significant percentage of software projects (Keil, Mann and Rai 2000, Standish Group International 2001). Some major reasons for this failures are e.g. the insufficient understanding of the integration in practice (Suchman 1995), information asymmetries, and goal incongruence (Keil et al. 2000). We propose: by improving the communication basis for the different stakeholders involved in the ISD process, we can achieve a better alignment of the IS with the business model of the organisation.

Along with the categorisation by Gregor (2006), the contribution of this article is the development of a theory that helps to explain and predict language interaction in ISD. Therefore, we emphasise on the role of language interaction by matching social interaction and communication theories with language theories. Our results emerge from the triangulation of positivist and interpretive research approaches on a research case (Lee 1991) and brought to light different strategies to deal with communication bottlenecks. At the end, we obtained different language interaction levels, which serve as orientation for specific interventions in an interaction setting, enabling an intensification of consensus negotiation.

We proceed as follows: In the next section, we locate the position of this article within the theoretical background of its field and introduce theoretical assumptions relevant to the contribution of this article. Then, we outline our fundamental theoretical propositions based on social interaction, communication and language theories for the particular ISD setting. After drafting our research approach in section 4, we present the results of the interpretive observation and the observed language interaction levels in section 5. Section 6 gives an overview on the application and embeddedness of language interaction levels in different observation cases and finally, section 7 presents an outlook and limitations of this article.

2 STATE OF THE ART AND PROPOSITIONS

Coming from a functional tradition, the IS research has been challenged by new alternative approaches since the 1980s (Lucas 1975). By the consideration of the contribution of those involved in and affected by the system, the understanding of IS as human activity systems gained significant attention (Robey and Markus 1984, Clarke and Lehaney 2000). A further focus addresses communication-related questions about the processes of social interaction in IS and ISD (Hirschheim, Klein and Lyytinen 1995a) and the existence of a practice fraught with volatility, exceptions, unstructured data and unpredictable requirements (Truex, Baskerville and Travis 2000).

As far as language is concerned, although its relevance seems not fully established, the IS research has introduced different approaches and methods focusing on language processes (Lyytinen 1985, Winograd 1988, Helmut 1997, Holten 1999). The main focus on natural language has been set by
the Language Action Perspective (LAP) (Goldkuhl and Lyytinen 1982, Winograd 1988) which focuses on linguistic communication as the basis for understanding IS as well as on the impact of language action on a system (Winograd 2006). However, the focus on action through language was adopted for the analysis of ISs and less for the improvement of language use in ISD.

Yet, the importance of language for ISD can be discerned as dually constituted:
- Language is an instrument for communication in ISD (Pohl 2007)
- Language is deeply involved in the development and formalisation process representing coherent elements and functions of the future IS (Lyytinen 1985, Hirschheim et al. 1995a)

This is the point where the ISD setting defers from many other interaction and communication settings: language is an instrument and a result at the same time. Therefore, a main assumption of this article is that the ISD setting is a language based setting (Lyytinen 1985, Hirschheim, Klein and Lyytinen 1995b, Holten 2007) and that we need a special focus on this topic.

Regarding ISD as a language development and formalisation process (Lyytinen 1985), a main goal in ISD is the development of a coherent and sound IS model (Thomas and Carroll 1981, McDavid 1996, Holten 2007). The resulting models have to be successfully 1) legitimised on a language and knowledge level (Boland 1979) and 2) consolidated by social action and communication (Hirschheim, Klein and Newman 1991) in a collaborative setting, where multiple stakeholders representing the different fields of the organisation and of the system development are involved (Alvarez and Urlo 2002, Kavakli and Loucopoulos 2003). Translating this assumption to the language perspective implies a somehow performed negotiation about and concretisation of language as a sine qua non during the ISD process. Several contributions have been made discussing the role of fixing the requirements elicitation in natural language (Ryan 1993, Rupp and Sophisten 2002). And even if the requirements are fixed in a formal language, we have to acknowledge the use of natural language because it is the most probable common communication media between the different stakeholders, and the most used communication instrument in practice (Pohl 2007).

The advantages of natural language are its universal use in different knowledge areas, its flexibility regarding abstraction grades and its simple use (Kamsties 2001, Pohl 2007). Nevertheless, there are significant disadvantages of natural language. Its inherent ambiguity (Pohl 2007) – e.g. lexical, syntactical, semantical, referential ambiguity and vagueness – provide space for different interpretations of the same requirements (Berry and Kamsties 2003), leading to an increase in negotiation costs and of the risk of misunderstanding in requirements engineering (Holten 2007).

Holten (2007) applied the language critique of the “Erlangen School” (Kamlah and Lorenzen 1984) in ISs and assumes that the reduction of the immanent ambiguity of language and the construction of language consensus between the ISD stakeholders is reached by the creation of a language community as a group of stakeholders with a consensual understanding of relevant language representations. Consequently, the stakeholders of an IS project ought to reach a shared understanding of and an agreement on the same concepts about the system’s elements and functions.

Consensus as a language matter is also addressed by different researchers and practitioners (Rupp et al. 2002, Holten 2007). Different instruments to support the management of terminology were developed; e.g. the construction of dictionaries or lexica for consolidation of the company’s language (Sommerville and Sawyer 2003, Pohl 2007), the language standardisation instances (e.g. ISO TC37) or the development of terminology management systems or support tools (Holten 2001, Kamsties 2001). Nevertheless, none of them gives advices how consensus emerge at the micro level of the communication process in ISD. To analyse this setting we juxtapose in the following communication, social action and language theories with our understanding of ISD outlined in the next section.
3 THE COMMUNICATION, SOCIAL ACTION, AND LANGUAGE VIEW ON THE ISD SETTING

The analysis of communication is usually discussed in the sense of the classical code model (Shannon and Weaver 1949) and refers to the act of transmitting codified knowledge from sender to receiver. Yet, as already outlined, the successful development of an IS model means more than a transmission of requirements, but its social, action-driven consolidation. Consequently, a more adequate focus on communication is taken by circular communication concepts, assuming that communication success is related to the right interpretation of the utterance by the receiver and including his or her reaction (Sperber and Wilson 1995). Furthermore, Watzlawick (1990) sees communication as a reciprocal sequence of utterances between individuals in an action/reaction act. As communication is “carried out by an ensemble of people acting in coordination with each other” (Clark 1996), we use the term interaction, that refers to reciprocal social action. Thereby every stakeholder in the interaction encompasses his or her own actions with the actions of others. The perspective underlying this analysis is based on the social action theory of Max Weber (Weber 1921/1967): people perform instrumental-rational interaction which defines a social interaction form aiming for the achievement of goals through the use of an artefact. In this setting the goal is the development of an IS, and the artefact is language. As a result, we introduce the concept of language interaction.

The linguistic perspective is endorsed by basic semiotic concepts. Saussure (1974), as the precursor of modern linguistics, introduced the term sign as a two-sided psychological entity. It consists of the relationship between the signified as a mental fact related to objects in the real world (concept) and the signifier as its psychological imprint (sound pattern or term). That means, in language interaction we deal with subjective concepts by using terms as their representations. As the communication takes place over the use of terms, we also need a similar understanding behind, that is, the used terms should relate to the same concepts. Yet, the existence of same concept-term relationships is not self-evident, but as Kamlah and Lorenzen (Kamlah et al. 1984) argue, it can make the communication process more effective. The collaborative challenge in this field is thus the design of the process of language interaction in an effective way (Thomas et al. 1981).

Moreover, we postulate that the relationship between concepts and terms is developed in the particular socialisation process of each individual and actualised in the concrete interaction process. There, meaning is developed and modified through schemes of interpretation (Blumer 1973), e.g. subjective language schemes or assumptions. In Figure 1 we draft the role of language interaction in the ISD setting as used in this research.

![Figure 1: Language interaction in the ISD setting.](image-url)
4 RESEARCH METHOD AND RESEARCH FIELD

Our research is developed against the background of the ontological assumption of the existence of a real world but the epistemological assumption that social action always arise out of a subjective cognition of it (Burrell and Morgan 1979). Moreover, we apply the combination of interpretive and positivist research according to Lee (1991).

According to Figure 2, the first step, the interpretive understanding (1) consists of an interpretive observation of the language interaction in the ISD project. The procedure used by the researcher to develop the interpretive understanding from the subjective understanding (Lee 1991) includes the use of the hermeneutical circle to test the validity of the resulting interpretive understanding according to the research principles proposed by Klein and Myers (1999). The qualitative participant observation consists on the observation of the language interaction in the different face-to-face project meetings. As part-involved researchers, the observation was at all times performed in a two-person team, so that one researcher could concentrate on the observation and the other performed the required tasks. A qualitative participant observation was not only the precondition for the realisation of the research but it is moreover a necessary aspect for the understanding of the language community created by the stakeholders. In the second step, the theory-driven development of the positivist understanding (2), we develop hypotheses about the language interaction levels and its impact on the consensus building by analysing the interpretive results and testing them against relevant literature in the areas of communication, social action and language theories. Based on Schutz (1962), in the second step a testing of the developed assumptions in terms of rationality was performed as our assumptions were compared rationally with other theoretical propositions (Lee 1991), in this case, of language interaction. The third step, the positivist analysis (3), consists of a positivist re-observation of the different communication strategies on the ISD setting to confirm the relevance and the impact of these strategies in a qualitative way.

The observed ISD project aimed at the development of an IS for analysis, storage and retrieval of market-specific and user customised information with an expected user group of >500. The work was divided in two main tasks to be fulfilled: 1) the knowledge reproduction task, where the content of the system should be identified and collected and 2) the development of the IS, including the election of required system elements and functions. Our observation focus was set on 2).

The research began on December 2006 until September 2008. The development of the structure of the IS model was based on the acceptance and adoption of used concepts and terms used by the involved
stakeholders. The meetings took place in different project member configurations. By the drafting of this article, the requirements were being successfully realised.

The researchers had a project-inside role, as they were responsible for requirements engineering. This research position was convenient because although the other project stakeholders were informed about the documentation of interaction for research purposes, the reaction towards the researcher’s role was overlapped with the role as stakeholders in the project. Table 1 shows an overview of the stakeholders involved in the research, the research team is underlined. Additionally, the information of the domain knowledge serves as indicator for the existence of different technical languages between the teams.

<table>
<thead>
<tr>
<th><strong>ISD related Knowledge</strong></th>
<th><strong>Business Knowledge</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Management Team (PMT)</td>
<td>Business and organisational knowledge (2 Stakeholders)</td>
</tr>
<tr>
<td>Requirements Engineering Team (RET)</td>
<td>Academic and practical knowledge in IS (2 S.)</td>
</tr>
<tr>
<td>Developer Team (DT)</td>
<td>Academic and practical knowledge in IS (2 S.)</td>
</tr>
</tbody>
</table>

*Table 1. Involved teams in the analysed ISD project.*

## 5 INTERPRETIVE AND THEORY DRIVEN POSITIVIST UNDERSTANDING: LANGUAGE INTERACTION LEVELS

After the first interpretive observations it could be asserted that the stakeholders – using language as an artefact – changed the manner how to deal with language many times and were able to build a common meaning through their interaction. These changes seemed to occur motivated by the necessity to reach an interaction goal and we could observe which impact such changes had on language interaction success.

To classify the areas where changes occurred, we first use the concept of language interaction areas (LIA) as a superordinate concept. Table 2 offers an overview of the observed areas.

<table>
<thead>
<tr>
<th><strong>LIA</strong></th>
<th><strong>Interpretation</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>LIA-1</td>
<td>Existence of phases where the language interaction between the stakeholders strove for a common understanding of terms and concepts and, on the other side, phases where the interaction was built on mutual complementation referring to concept use.</td>
</tr>
<tr>
<td>LIA-2</td>
<td>The language negotiation was in most of the cases not performed on the language interaction surface, but implicitly present in the background of the discussions. Then again, in few cases the negotiation about language occurred explicitly, in the sense of: “what do you mean with the word XY?”.</td>
</tr>
<tr>
<td>LIA-3</td>
<td>Existence of a different use of natural language referring to the elaboration of utterances. There were on the one hand phases were technical language was on the foreground; on the other hand we could observe a high use of colloquial language and narrative elements.</td>
</tr>
<tr>
<td>LIA-4</td>
<td>Besides the oral language, written language was used during the face-to-face language interaction. It seemed to have different purposes, on the one hand, it was used for documentation and in other cases it allowed the visualisation of meaning.</td>
</tr>
</tbody>
</table>

*Table 2. Language interaction areas.*

Within each LIA, the discerned changes of the manner how to deal with language are now referred to as changes of language interaction levels. In the following paragraphs we present the language interaction levels for each LIA as a result of the theory-driven positivist analysis on the interpretive observations. Afterwards Table 3 provides an overview on these language interaction levels.
LIA-1: Symmetrical vs. complementary interaction level

In prior literature the phenomenon observed in LIA-1 was already analysed (Watzlawick 1990, Dennis and Valacich 1999, Holten 2007). We apply Watzlawick's (1990) concept of symmetry and complementarity. Symmetry means an interaction which aims for the building of the same knowledge, complementarity means that one stakeholder's knowledge complement the knowledge of the other. In the case of language, we assert that as long as both stakeholders have the same language basis, the interaction leads to a complementation of knowledge. Otherwise, the interaction leads to the building of a same language basis.

LIA-2: Explicit vs. implicit language interaction level

Based on the results of LIA-2, there are different approaches referring to the manner how language is negotiated constantly in everyday life. As everyday language is learned empratically in the socialisation process (Bourdieu 1990) it has a self evident role on its use (Kamlah et al. 1984, Rupp et al. 2002, Pohl 2007). We use language more than we talk about it. The process of connecting signs with specific concepts is involved in its use in an existing environment of social interaction, and is therefore coined empratically, that is, by its practical use (Kamlah et al. 1984, Bühler 1990). Although an explicit language negotiation may be not usual, it sometimes seems inevitable, as for example in academic fields (Kamlah et al. 1984).

LIA-3: Colloquial vs. technical language interaction level

The questions raised by LIA-3 refer to the level of abstraction of language use. The use of technical language is usually associated with a differentiated and less context dependent vocabulary (Cummins 1979) whereas the use of colloquial language is fraught with narrative contents, less concrete use of terms and context-rich utterances (Cummins 1979). Although the interaction in the level of colloquial language may be an indicator for a weak sign concretisation in the sense of sign unambiguity, it may provide a clearer notion for people with different language backgrounds. These observations are similar with the results of an ISD research performed by Alvarez et al. (2002), which assert that narratives and messy data may provide a deeper insight into the user’s perspective.

LIA-4: Level of language interaction volatileness vs. codification vs. visualisation

The differentiation between oral and written language use, as observed in LIA-4, is of particular interest in communication theory. One important aspect is the question whether the results of language interaction remain fleeting or are perseverant as in the case of written language (codification). The construction of enduring utterances, e.g. in the form of reports and documentations are understood as language actions on time distance (Clark 1996). This implies a further reflection about the expression chosen, which ensures stability and uniformity of what has been written (Hellmuth 1997). Therefore, the change between oral and written language can be classified as more than a change between communication media but as a language interaction modus. A further observation is the use of written specifications, e.g. in form of conceptual models. The ISD has a long tradition of the use of conceptual models as instruments of the visualisation of concepts (Kottemann and Konsynski 1984, Karimi 1988). Therefore, this issue is considered as important as well.
<table>
<thead>
<tr>
<th>LIA</th>
<th>Language Interaction Levels</th>
<th>Description</th>
</tr>
</thead>
</table>
| LIA-1 | Symmetrical vs. Complementary Language Interaction                                      | Symmetrical: Communication leads to the formation of a common language basis.  
Complementary: Communication built on a common language basis. The result is the complementation of knowledge. |
| LIA-2 | Explicit vs. Implicit Language Interaction                                                    | Explicit: Stakeholders negotiate about language in an explicit way.  
Implicit: Stakeholders negotiate about language in an implicit way. |
| LIA-3 | Colloquial vs. Technical Language Interaction                                                 | Colloquial: The language used is an easy structured, common everyday language.  
Technical: The language used is clearly defined, more standardised and not usual in everyday communication. |
| LIA-4 | Language Interaction Volatility vs. Codification vs. Visualisation                           | Volatility: The language used is fugacious and concatenated.  
Codification: The language used is written and used for documentation.  
Visualisation: The language used has the purpose of provide a visible reference or an overview. |

Table 3. Language interaction levels.

6 APPLYING LANGUAGE INTERACTION LEVELS IN THE CONCRETE ISD SETTING

As a conclusion we present three exemplary cases (EC) where the existence of the developed language interaction levels and its impact on language consensus is analysed, showing the appropriateness of the results already presented. The criteria for the choice of the ECs are on the one side their appropriateness to show some important insights in consensus building and language interaction, which can be the basis for further research. On the other side the ECs show a spectrum of different interaction lengths. EC-1 refers to an observation performed over different meetings, whereas EC-2 was observed over one meeting in a time segment of about 2 hours. EC-3 shows a passage of about 10 minutes and gives an outlook on further research. Afterwards we analyse EC-1 to EC-3 altogether and present further results which can be gained over the analysis of language interaction levels.

Analysis of EC-1:

In prior meetings the stakeholders had elaborated some notions about the structure and functions of the IS. Nevertheless the language interaction in one of the concluding meetings between PMT and RET was still fraught with ambiguity about how these foregoing notions should be articulated and referred to. In the specification process about the structure of the IS, there were many confusing situations in language use. Although the stakeholders had developed notions about the concepts behind entity types, they were not able to refer to them. They used different terms or described them by using colloquial language (LIA-3).

During the subsequent language interaction, the search for a language anchor was so obvious that the point arrived where one of the participants explicitly asked for a time-out to order the different concepts behind the terms (LIA-1). By doing so, he changed the current language interaction from the implicit into explicit language interaction (LIA-2). In the course of the conversation following, a negotiation about which terms should be connected with which concept emerged. In this negotiation, the relevant concepts were solidified and linked to terms; furthermore they visualised the relationship between the different terms (LIA-4). To better concretise the negotiation results, the stakeholders were also engaged in writing the produced terms and the relationships between them down, building thereby a structure (LIA-4).
In the next meeting, where the consolidation of the systems structure had to be accomplished, the PMT presented the negotiated terms and its relationships (LIA-1, LIA-2), indicating the importance of the specified terms for a better understanding within the project.

**Analysis of EC-2:**

The first meeting between the PMT, the RET and the DT was held in order to bring the different perspectives of the stakeholders together and develop a first concept of the IS structure. During the meeting, the discussion on the term *content* seemed to be held in an atmosphere of confusion. Although the stakeholders referred to the same term, they seemed to have different technical perspectives on the question of “how to deal with *content*”. By using the term *content*, the PMT referred to the data display whereas the RET and the DT meant the data structure on an abstract level.

In this first phase of the meeting, the language interaction level of all stakeholders can be classified as complementary interaction. All three stakeholders were trying to build new knowledge without realising that the symmetric interaction could more appropriate (LIA-1). Later on, the negotiation seemed to stay in the level of implicit language interaction, as the concept was embedded into further explanations, but not defined as such (LIA-2). In this phase, the stakeholder interchanged between interaction volatileness and interaction codification (LIA-4), and colloquial and technical language interaction (LIA-3), without visible impact on the consensus building process. After a while, the possibility of everyone talking about different concepts was explicitly mentioned by one of the stakeholders (LIA-2). Then, the other stakeholders agreed on receiving an exemplary content document from the PMT to have a look at the meant *content*, thereby the stakeholders achieved a common understanding.

The change into explicit language negotiation (LIA-2) resulted to better suit the requirements of the language interaction. Even if the change between implicit and explicit language interaction as well as symmetric and complementary interaction seemed to be decisive (LIA-1, LIA-2), we suppose that the changes between interaction volatileness and visualisation (LIA-4) and casual and elaborated language interaction (LIA-3) were not unimportant as they brought the situation to escalate and enabled the change into an explicit language negotiation as a last chance of understanding.

**Analysis of EC-3:**

<table>
<thead>
<tr>
<th>Language Interaction</th>
<th>Language Interaction Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A: Member of RET</strong></td>
<td>Complementary</td>
</tr>
<tr>
<td>A: This is then saved as “Users Profile” (…).</td>
<td>X</td>
</tr>
<tr>
<td>T: I would call it “Personalisation area”</td>
<td>X</td>
</tr>
<tr>
<td>A: Uh… “Personalisation area” is the whole thing.</td>
<td>X</td>
</tr>
<tr>
<td>T: Ohh, ok, (…)</td>
<td>X</td>
</tr>
<tr>
<td>T: I interpreted it as an “Information Object”, and then I assumed that it is associated with the user.</td>
<td>X</td>
</tr>
<tr>
<td>A: No…</td>
<td></td>
</tr>
<tr>
<td>T: Look, for us, an “Information Object” is not an “Information Object” anymore; it is e.g. a “Report”. (…) In addition, that what we call “Report” is now in this field the “Personalisation area”, so, the “User”.</td>
<td>X</td>
</tr>
<tr>
<td>A: No, the “Personalisation area” is not the user; it refers to all possible users.</td>
<td>X</td>
</tr>
<tr>
<td>A: This is everything that we can have as users, all this here (ed.: he draws a sketch of the relationship)</td>
<td>X</td>
</tr>
</tbody>
</table>

*Table 4. Exemplary case using language interaction levels.*
The dialog in EC-3 (Table 4) makes clear that as T has another concept of Users Profile, both stakeholders have to negotiate about their understanding of terms (LIA-1, LIA-2). Thereby they were able to find out that they have a different understanding of Information Object. It shows furthermore, how a same term related to different concepts can be the source of misunderstandings and how the stakeholders use language visualisation to find a common understanding (LIA-4).

EC-3 also shows how to classify language interaction situations along the drafted language interaction levels and it indicates possible further research fields, e.g. by coding the language interaction we could search for patterns or analyse the data using quantitative research as well.

**Analysis of EC-1 to EC-3**

EC-1 and EC-2 show that the estimation about the adequate point when the negotiation about language should occur is not self-evident. Both cases showed that the adequate point to negotiate explicitly about language seemed to be missed and, as a consequence, the language interaction remained ineffective in a first place.

In addition, EC-1 and EC-3 suggest that the explicit language negotiation can involve important solutions for ISD problems as:
- the concretisation of the concept-term relationship,
- the development of a conceptual model consisting of sign structure and its relationships and
- the development of common and accessible definitions of the concepts and terms used, which can serve as anchor for referring to already reached consensus.

Insofar, the change from language volatileness to language codification and visualisation can also be assumed as being important for consensus building (EC-1 to EC-3), as the conceptual model served as a basis for future communication.

Below the line, we could observe that the categorisation of language interaction in language interaction levels can provide a helpful perspective to understand the consensus building process. The awareness of language interaction levels could give stakeholders simple strategies to react in the different language interaction situations and to influence the interaction course significantly.

**7 LIMITATIONS AND OUTLOOK**

The detected language interaction levels show how language is used over the whole ISD process. This may contribute to a better language consensus building in ISD by providing orientation for specific interventions in the interaction and enabling an intensification of language interaction. Thereby we can gain a deeper understanding about the emergence of language consensus.

Although the interpretive approach and the process of finding the different language interaction levels seems to be the more sensitive link in the chain, we need to go over this step as language emerges in the interaction. There is a necessity of observation and participation as the only way to find out which language negotiation strategies are used in the micro level of communication is to be present there. There we have the chance to develop useful hypotheses to be able to intercede in language interaction by changing the schemes of interpretation of the stakeholders. By analysing the language interaction in ISD settings, we provide qualitative evidence and analytical generalisation of how language consensus is reached in practice. Using the developed interaction levels proved to be helpful tools for the analysis of single observations cases in the field of language interaction in the ISD setting.

Further research, besides the aforementioned research fields, could focus on stakeholder interviews, in order to get more detailed information about the user interpretation of the strategies to reach consensus. Another alternative could be action research, trying to change the current interaction levels intentional and observing its impact on the development of coherent entities in ISD.
Acknowledgments

This work was supported by the “Innovation With Services Research Program” of the German Federal Ministry of Education and Research (BMBF 01HQ0608).

References


BRIDGING THE GAPS: PACKAGED SOFTWARE IMPLEMENTATION PROJECTS IN VIETNAM

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0171.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Developing countries, Implementation, Grounded Theory, interpretivist research</td>
</tr>
</tbody>
</table>
BRIDGING THE GAPS: PACKAGED SOFTWARE IMPLEMENTATION PROJECTS IN VIETNAM

Vuong Nguyen, The National Australian University, vuong.nguyen@anu.edu.au
Walter Fernández, The National Australian University, walter.fernandez@anu.edu.au

Abstract

Implementations of advanced information systems in developing countries are often perceived as a way to economical prosperity and modernity. However, even in the best of scenarios implementing major information systems is far from a risk-free activity. The reality of developing countries adds extra complexity to an already demanding endeavor. This study explores socio-technical aspects of an unprecedented technological journey in the Vietnamese banking industry, and thus responds to calls for more empirical research in this field. In this paper we present contextual issues affecting the studied software package implementation project and propose a theoretical framework for a basic social process we called “gap bridging.”

Keywords: Packaged software, Implementation projects, Developing countries, ICT for Development, Grounded theory, Case study, Empirical studies, Theory building.

1 INTRODUCTION

Businesses in developing countries often adopt advanced information technology (IT) as a way of keeping pace with the developed world. This is both a necessary and a risky activity. While IT investments in developing countries can contribute and foster economic growth and development (Avgerou, 2003, Yoo, 2003), IT implementation projects are intrinsically prone to failure (Lyytinen and Hirscheim, 1987) and these failures are more significant in developing countries (Heeks, 2002).

The socio-technical conditions of developing countries demand special research attention. On one hand, there is an increasing interest in the use and adoption of information systems (IS) in these countries (Walsham and Sahay, 2006). On the other hand, increasing the knowledge of conditions and processes related to socio-technical changes in developing countries through the adoption of technology is significant to both IS research and practice (Avgerou, 2008). Yet, the number of empirical studies on IS implementation in developing countries is limited (Walsham et al., 2007).

In response to this, our study investigates the implementation process of large package-based information systems (Iivari, 1990) in Vietnam. This study presents a theoretical account that emerged from the perspective of Vietnamese client organizations engaged in “an organizational effort to diffuse and appropriate information technology within a user community” (Kwon and Zmud, 1987 p.231).
2 LITERATURE REVIEW

IS implementation is an important research topic that has been studied since the 1970s (Lucas, 1975, 1981, Ginzberg, 1981, Cooper and Zmud, 1990, Lapointe and Rivard, 2007). While “the early implementation research lacked a sufficiently rich technological, institutional, and historical context” it succeeded in identifying critical success factors and providing multiple insights into the development and implementation process (Lucas et al., 2007 p.208). After its initial concerns with success and failure, the direction of research on IS implementation moved to packaged solutions, enterprise-wide implementations and rapid implementation (Lucas et al., 2007).

Packaged IS implementations have been studied since the 1980’s (e.g., Markus, 1983, Lucas et al., 1988). Packaged software presents benefits over custom-made software; however, the implementation of packaged software is not trouble-free (Iivari, 1990) and organizations are likely to find hidden costs (Lynch, 1984), multiple pitfalls (Lynch, 1985, Ben, 2005) and implementation risks in the external business context, the organizational context, the IS context and the project context (Scott and Vessey, 2002).

IS implementation projects in developing countries add risks to four contexts identified by Scott and Vessey (2002), presenting an interesting area of research with many possibilities. IS adoption in developing countries has been mainly studied from the perspectives of technology transfer and diffusion, social embeddedness, and transformative discourses (Avgerou, 2008). However, these studies can also focus on the role of IS in development, cross-cultural collaborations, localization, and marginalization (Walsham and Sahay, 2006).

The change agency aspect of technology must also be considered. The IS implementation process clearly involves organizational change (Zmud and Cox, 1979, Markus, 1983, Orlikowski, 1993, Lyytinen and Newman, 2008). Markus and Robey (1988) suggested that theories about organizational changes caused by technological innovation should be evaluated in terms of causal agency, logical structure and level or unit of analysis.

Following Markus and Robey (1988) causal structure dimensions, our research adopts an emergent perspective, with the logical structure of process theory at organizational level. This exploratory, theory-making, study follows an interpretive epistemology to analyze empirical data from case studies of large packaged IS implementations in Vietnam.

3 RESEARCH METHODOLOGY

Our objective is to generate a substantive theory of packaged IS implementation in Vietnam while allowing the special economic, political and developmental aspects of Vietnam to emerge without the constraints of theoretical preconceptions. We adopted classic grounded theory methodology (Glaser and Strauss, 1967) to derive theory from case study data (Fernández, 2004). Grounded theory (GTM) is an inductive, contextual and procedural method suitable for exploratory studies (Orlikowski, 1993, Myers, 1997) that facilitates the generation of substantive theory based in the discovery of patterns from participants’ accounts and data collected in the field (Glaser, 1998). GTM fits well with case studies (Eisenhardt, 1989, Orlikowski, 1993), provides well established guidelines for data collection and analysis and is likely to produce a theory that is relevant, rigorous and readily modifiable (Glaser, 1978, Fernández and Lehmann, 2005). Using case-based data for theory building studies allowed us to deal with complex processes enacted by multiple actors in a real life situation and to use multiple sources of evidence to derive knowledge (Eisenhardt, 1989).
3.1 Data collection

It was critical to the success of our study to have access to IS projects that were important to local organizations and exhibited organization-wide impact. To gain access to suitable sites personal contacts were needed. The first author attended a banking conference in Vietnam; from contacts made at this conference, we gained access to six cases. To facilitate theoretical emergence, we choose to start with the case of CORE, a significant IT project at BetaBank (pseudonyms) because it was “the most complicated of the real cases” (Yin, 2003 p.74) and thus well suited as a foundation case (Fernández, 2004). BetaBank is a top commercial bank in Vietnam and CORE presented an excellent foundation case given its magnitude, socio-technical complexity. CORE was also the first system of its kind to be introduced in Vietnam and it was under implementation from 1994 to 2004.

Following Yin (2003), we collected data from documents, records, interviews, and direct observations of the systems in use and the actors in situ. The first author spent 4 months at BetaBank. The first month allowed for familiarization with the case through document reading, observations and informal discussions with project personnel. This activity was very important to become familiar with the case and also for the participants to become familiar with the researcher. The following three months were used for formal interviews; yet, access to documents and informal “chats” continued.

Interviews were the primary source of information. Data from 15 formal interviews were collected and analysed. We interviewed different stakeholders including the bank’s board members, IT department managers and IT staff, relevant business department heads and members, project manager and project team and business unit users. Each interview lasted up to two hours, with average of one and a half hours and was recorded and transcribed, producing more than 250 pages of text.

Following the GTM approach, the interview process evolved over time as the themes emerged from data analysis. For example, our earlier interviews were more open and unstructured, but the later ones were guided by the emergent themes and concepts developed from data analysis (Glaser, 1978). Many brief follow-up discussions were also used to gather further information, to explain emerging issues or to clarify aspects or issues found during data analysis. The strategy to develop rapport during the first month was very effective in facilitating the engagement of participants.

3.2 Data analysis

Data analysis activities included open coding, sorting, selective coding and theoretical coding (Glaser and Strauss, 1967, Glaser, 1978). Open coding produced 217 codes related to meaningful excerpts from the data (or instances). The codes were further sorted into categories while the constant comparison of incidents facilitated the generation of theoretical memos, thus raising the conceptual level (Glaser, 1978). By comparing incidents in each category, we generated theoretical properties of the category. The emerging 26 themes included concepts related to gaps, communication, learning, conflicts, goals, delays and performance.

We conducted further theoretical sampling and selective coding within the core themes. Memo writing proved to be helpful in recording the ideas generated during the analysis and theorize on interrelations. Memos also help to guide our research discussions and to reach higher levels of conceptualization. From this iterative process the “core variable” (Glaser and Strauss, 1967) emerged, we called it the gap bridging process, a basic social process (Glaser, 1978) described in section 5 of this paper.

4 BETABANK’S CONTEXT

This section describes factors in the external and internal environment of the bank that we found influential in the case. The external environment presented the following influential factors:

- Laws and regulations. The banking services were heavily regulated by the state laws and regulations. The CORE implementation was affected by requirements that were mostly suitable to
the manual operations of old-fashioned commercial banks. The fast pace of integration of Vietnam into regional and world economies demanded changes in laws and regulations. These changes were often sudden and unexpected.

- **Industry development.** In the late 1980s, Vietnam’s banking system was transformed into a two-tier system, separating the state bank and commercial banks. Consequently, in the 1990s the banking industry in Vietnam was less developed than those in Asian region and industrialized countries. The CORE project was then seen as the pioneer in modern banking technology and most Vietnamese commercial banks were observing how BetaBank implemented CORE.

- **Technological base.** In the early 1990s, most of the commercial banks used either their in-house developed program or package programs. These programs could only be used in a local area network for each branch with its sub-branches. Data exchange between a bank’s headquarter (HQ) and its branches were through file transfer by modem connection over telephone lines. The national telecommunication infrastructure was in poor condition with low capacity and interrupted services.

BetaBank is one of the four biggest local commercial banks in Vietnam with over 5,000 employees and a branch network covering over the country. It provides wholesale and retail banking services to government agencies, corporate customers and the general public. The bank had strong banking services to support foreign trade, FDI, ODA and remittance. The organizational context had a direct effect on the CORE project. The most influential factors in this context included the followings.

- **Business strategies.** Before CORE, each branch operated as an independent bank with its own funds, management and operations. In the early 1990s, the bank’s management started a process of centralization of fund management and increased control over the branches’ operations; as we shall explain, this was a key driver for the CORE project.

- **Business processes.** Most of the business processes at the bank were poorly-documented. Policies relating to business processes or operations issued by HQ were usually ambiguous, causing different interpretations. Business processes were not standardized across the bank; for example, a loan to a small enterprise would be provided differently across branches. HQ was unable to ensure the observance of its policies by the branches.

- **Management control.** As a consequence of a long tradition of branch autonomy, the organizational structure of each branch developed independently and was different from that of HQ. Each functional department (either at branch or at HQ level) had considerable authority and strong views regarding how their specific business should be conducted. The HQ management control was weak, and rewards and penalties were ineffective and seldom used.

- **Corporate culture.** Following the autonomous traditions, BetaBank’s culture was characterized by a focus on local interest and unwillingness to share information and knowledge. The function-based organizational structure contributed to the lack of cooperation and cross-organizational team spirit.

- **IT strategies.** The bank did not have clear IT strategies. The bank had a tradition of in-house IT provision of all IT supports to business areas as required by the business operations.

- **IS/IT management policies and procedures.** Before the project, all the big applications at the bank’s network were developed by the IT department at HQ. Each branch had its own IT group to support the branch’s operations and to develop applications for its own needs.

## 5 THE CORE BANKING SYSTEM PROJECT

In the early 1990s, while BetaBank was working with partner banks in the developed world, the bank realized that they needed to improve its operations through the use of IT. As a consequence, the bank’s top management decided to implement a core banking system (the CORE project). This was the first core banking system (CBS) to be implemented in Vietnam.
The adopted CBS was packaged software developed in USA and used by many commercial banks in USA as well as in some Asian countries. The CORE project went through a sequence of partly overlapping phases (see Figure 1) classified by the critical events and characteristics of the activities undertaken by the organization and project implementation team.

![Figure 1: Phases of the CORE project.](image)

In the **emerging** phase, BetaBank decided to use a CBS for online service and central processing of transactions for the whole bank, including HQ and its 25 branches. CORE was perceived by top management as a mechanism both to modernize operations and to bring the branches under control. The IT department initiated the project and assisted the bank’s management in making important decisions such as choosing the client-server technology and buying a modern packaged system rather than developing in-house. In early 1996, a vendor from Malaysia was chosen to provide and implement a software package solution developed in USA. The functionality of this system was thought to be aligned with the business objective of centralization of the fund management undertaken by the bank.

At the **exploration** phase, a project taskforce was setup to study the base system provided by the vendor and also the bank’s business processes and operations. The taskforce performed a gap analysis and prepared the user requirement document (URD) for customizing the system. The group encountered several barriers, including: their own lack of knowledge and experience in the technology and modern banking practices; lack of cooperation and contributions from other units at the bank; and, lack of business and IT strategies to guide the development of the URD. Hence, their URD was too general and failed to capture the bank’s operations in sufficient details. Based on this deficient document, the bank and the vendor signed a contract to customize and implement the CBS across BetaBank. A flexible contract condition covered unanticipated changes required by state regulations.

The **evolving** phase was the longest and most difficult for the bank and the vendor. The formal implementation team was set up, consisting of IT staff and business staff from HQ and some branches. The interactions between technical changes and business changes made the implementation more complicated. The technical side required an iterative and evolving process of system customization/development, system testing and gathering further user requirements. The business side required changes to the business processes, organizational structure, policies and procedures were proposed, discussed, and implemented. Also, the implementation teams from client and vendor, this phase required the involvements of many other groups, including the bank’s top management, functional departments, IT department, and users at branches. This phase experienced poor planning and a significant lack of formal and suitable methods. The changes were characterized as an emergent process. Conflicts were frequent in most tasks and among different stakeholders involved. Solving conflicts was seen as a time consuming process that caused delays in the project. Fortunately, the learning was most effective for the implementation team and other stakeholders. This phase lasted over three years and ended with the launch of the new system at a single branch at the bank’s HQ.

The forth phase, **rolling**, aimed at the installation and use of the systems in all of the bank’s branches. This was a hectic process with insufficient time to perform complex tasks such as user training, data conversion and launching of the system. This phase started in early 2000 and lasted over 1 year with 24 branches converted from the in-house developed legacy systems to the new CBS. While the
changes to business operations were implemented, the training was ineffective and users were unable
to exploit the benefits from the new system causing delays in benefit realisation, as explained next.

The last phase, achieving, took place from mid 2001 to the end of 2003. During the first year after the
rollout of the system, the bank-wide use of the new system demanded errors corrections, fine-tuning of
the system, and changes to business processes. The bank started to achieve benefits from the new
system, experiencing growth in customers and services offered. BetaBank was able to introduce many
new financial services thanks to the technological foundation of the new CBS.

From the data analysis the initial framework emerged as a pattern that was relevant and core to the
activities of the CORE project. The participants main concerns were able to be explained by a basic
social process (BSP) that we call Gap-bridging (for an extensive discussion on BSPs, see Glaser,
1978). The following section presents this emerging BSP as a theoretical framework.

6 THE EMERGENT THEORETICAL FRAMEWORK

From data analysis following grounded theories methodology, we discovered the emergent theoretical
framework of package information system implementation in the context of Vietnam – the gap-
bridging process (Figure 2.). A discussion on the nature and role of each sub-process is presented next.

6.1 Gap discovering

Gap discovery is a social process that detects discrepancies between the expected reality and the
reality found in action. We observed many instances where the implementation teams discovered
differences between the system purchased and the bank’s needs. One common cause for discrepancies
related to contextual differences between USA and Vietnam, as one team member said:

The development of the software was based on the thinking and logics of the foreign country [USA];
therefore what we needed to do was to customize totally. The technology of had to be changed totally.

The implementation of the CBS started with the gap discovering process, yet, the gap was initially
perceived in technical terms. Another team member defined gap analysis as follows:

Gap analysis is the comparison between the new system and the old [legacy system] to identify what is
the same and what is different.

Yet, our data suggests that gap discovering refers to gaps in the intersection of context and system,
with social and technical aspects. Thus, the gap discovering process serves to initiate the negotiation
process that will enable the resolution of the gap and the progress of the project. This process detects
mismatches between the characteristics of the system and the local context, including environmental,
organizational and user contexts (Figure 3).
Figure 3: properties of the context-system gaps (CORE data analysis).

Environment-system gaps refer to the mismatches between the system characteristics and the environment of the organization in which it is to be used. We found three elements (or gaps) at this level, including legal regulations, business customs and customer interfaces.

At the organization-system level, the gaps are mismatches between the proposed system and the existing structures, politics and culture. Gaps regarding the organizational structure are often easier to detect, they refer to elements such as communication channels, control mechanisms, decision rules and chain of commands (Markus and Robey, 1983). However, gaps related to organizational politics and culture tend to remain latent (undiscovered) for an extended period. Organizational politics refers to the distribution and use of power to influence decisions (Barbara and Sonny, 1985). Power relations among individuals and groups in BetaBank were important to the project and also difficult to discover and solve. The third element, organizational culture, “a complex set of values, beliefs, assumptions, and symbols that define the way in which a firm conducts its business” (Barney, 1986) acted as a retardant of information sharing, acceptance of alien technologies and collaborative behaviours.

Finally, user-system gaps refers to discrepancies regarding system user interfaces and operational procedures that required a certain level of user’s competence, skills, knowledge and certain cognitive characteristics. For example, assumptions about adequacy of training effort were proven incorrect because computing skills and literacy among the bank employees were lower than the vendor’s assumptions. This resulted in frequent user-induced errors, a more protracted implementation cycle, reduced user acceptance, and a more difficult user support task.

The gap discovering process is a cognitive process that converts unknown gaps into known gaps at all three levels. In the CORE project, both BetaBank and the vendor intended to carry out the gap discovering process during the Exploring phase in order to develop a complete specification of user requirements. However, this intension failed due to natural limitations in human cognition. Clearly, to imagine the unknown under the influence of the known and with imperfect information is a toll-order, as a team leader attested:

[The development of URD, because of the thinking of the old ways and not grasping of all the complexity of the project, was very limited. We found that it is really funny when we read that URD now.]

As seen in this indicator, the gaps were obvious, even funny, when observed with the benefit of the experience. Yet, the actual gap discovering was an on-going, incremental, process that went along with the gap closing process throughout the implementation resulting in a more suitable system. As one team member put it:

[The [system] was customized over the years into something totally different from the initial URDs. … When we actually worked on the program, the development of URD never stop. For over 10 years, it is still being developed.]

The delays in the process of customization were caused by several factor including the following: (a) the organization-system gap was very complex and contained tacit knowledge, embedded in the organization’s contexts (at structural political and cultural levels) and in the CBS; (b) the project
implementation teams lacked bank-wide support to investigate the organization’s contexts; and (c) the project team consisted of young, inexperienced, people with limited knowledge of the bank’s existing operations and management. Prior to the CORE project, the implementation team had no exposure to modern CBS or to an IT implementation project; they discovered their way while doing. One team leader stated:

“We did this project for the first time and had no training or education on how to carry out such a big project in a professional way. Mostly, we learned by doing.”

It could be argued that a more comprehensive identification of context-system gaps through the initial implementation team’s investigation could have contributed to forming more clear and realistic project goals and to a more efficient gap closing process. Yet, the observed gap discovering process was an important process leading to gap closing throughout the implementation, and had an important role in influencing the goal forming process.

6.2 Goal forming

The goal forming process consisted of three interrelated sub-processes: the identifying (needs and constraints), the balancing and the evolving processes, which are next explained in more detail. The discovery of gaps often caused the implementation team and BetaBank’s management to struggle with the choice of keeping the bank’s existing business processes or followings the functions and processes in the vendor’s base system. Consequently, one of the emerging themes was need-identifying. This process required the implementation team to clarify expectations and desires of both the bank’s management and the system users in order to achieve more precision in identifying clarify the goals. However, the requirements of these two groups were often tacit and divergent, demanding of the implementation team a great deal interactions to discover hidden details.

Another emerging process was the constraints identifying process, whereby the implementation team realized that not all identified needs could be satisfied with the available time, resources or because of external factors. Only when the identified needs were balanced with the identified constraints, feasible goals could be formed. However, balancing needs and constraints was an imprecise science requiring significant skills and knowledge. During the early stages of the project the goals were vague and not uniformly understood among the project stakeholders. As a result, the project team and management often underestimated the constraints, resulting in unrealistic expectations or unrealisable goals. As the project progressed, the project goals became clearer and more realistic; the efficacy of the goal forming process increased as the team’s capability improved overtime.

Implementation goals are important elements, they guide the change process by specifying what, when and how the system or the organization should be changed. The data showed that evolving goals acted as a guide to the gap closing process, which is discussed next.

6.3 Gap closing

Gap closing is defined as the process of creating organizational fit (Iivari, 1992) between the system and the organization by eliminating the discovered context-system gaps or by reducing the gaps to an acceptable level. This process consists of two sub-processes: changing and conflict solving.

Based on our data, we agree with Orlikowski and Hofman (1997) in that “when organizations are using information technologies to attempt unprecedented and complex changes … the depth and complexity of the interactions among theses activities is only fully understood as the changes are implemented.” The observed change in the CORE project provided evidence for three types of changes: emergent, anticipated and opportunistic.

At a high level of abstraction, the change process consists of technological change and business change, with each following its own process and interacting with the other. The change to the packaged software was initially planned to be done by sequential steps, but the actual method used
was iterative process of requirements, customizations, testing and further requirements. The cycle continued until the users accepted most of the functions in the system.

The lengthy customization process was influenced by the following factors. First, the users, including the end users and functional departments, were not involved in the earlier phases, but they had most important voice in testing and accepting the system. Only at this phase, the users fed their requirements to the customization along their testing of each system function by test cases. Second, the requirements of the laws and regulations were found to be also a significant source of changes to the customization. Considerable delays were caused by ambiguous legal requirements to banking services and accounting procedures discovered during the customization stage and during tests.

The business changes included organizational structure, business processes, management control and IT management. In this domain, the implementation team regularly interacted with different functional departments and top management. The prominent issues here were related to knowledge gaps, management bureaucracy and lack of responsibility and accountability. Several attributes influenced this process, including the following: (a) differences in understanding between groups about certain implementation issues, (b) the difference between the realities embedded in the organization and the software system and the perceptions of the people, (c) the management bureaucracy resulting in a lengthy decision making process that involved many functional departments and different management levels, (d) multiplicity of implicit procedures for different types of issues made the process more complicated, and (e) unwillingness to accept responsibility causing delays in decisions.

Change was often a source of conflicts among the groups involved in the implementation. Change resistance existed at all levels of the organization and appeared in all phases of the implementation process. Conflict existed between the change promoting (i.e., the implementation team) and change resistant groups (e.g., management, functional departments, users and vendor’s team). Changes also brought uncertainties and at times altered the distribution of power and benefits among the groups involved. In pursuing their different self-interests, these groups created conflicts. Lastly, gaps in knowledge caused misunderstandings and opposing views regarding implementation issues.

In the CORE project, conflicts happened in almost all issues emerging from the implementation process. The project implementation team was central in the conflict resolution process, balancing the different goals and interests of top management, users and vendor. Yet, satisfying the parties was difficult, as the following quotations illustrate:

..not all disagreements on the issues was solved satisfactorily. … [W]hen there were requirements from BetaBank, they [vendor] asked why we needed in that way. Both two sides were tense with each other.

… the technology and IT department …tries to keep the functions in the base system as much as possible. The users tried to keep as what they were doing. There were battles between the two. … Even in the board of management, there were fights on these issues.

With unclear responsibilities and accountabilities and a tendency to delay decision, the main conflict resolution method was a compromising process by which the involved parties went through lengthy discussions and negotiations to arrive at short-term solutions. As one informant put it:

[o]ur opinion was that one thing was necessary, we needed to persuade the other side about that, but you could not force them understand in our way. It was also true that they could not force us to understand in their way. It was human aspect.

The continuous nature of the change process coupled with the constant resolution of conflicts fostered the necessary learning in the project for all parties involved. The numerous issues emerging during the implementation required constant discussion and negotiations among the groups, closing the discovered gaps contributed to building critical capabilities in the implementation team.
6.4 Capability building

The implementation team was central to the implementation process. Their capability consisted of knowledge, experience, and skills that contributed to the deployment of CORE. Team capability at any point in the project is the results from a dynamic process involving staff turnover/recruitment and learning process.

Knowledge, experience and skills were limited at the beginning of the project, affecting both the client and vendor implementation teams. The client’s team members were either newly graduates recruited for this project or staff from functional departments with limited working experience in BetaBank. This caused certain level of frustration, as one of the users stated:

[the client team’s] operational knowledge and experience was an issue. When they received our complaints and requests, they could not understand why they should be as we requested.

In addition, the team knowledge of modern banking operations and management and of the capabilities of the CORE technology was also lacking. Fore example, a team member stated that:

All the people in the project team, when joined the project, did not know about the modern banking operations embedded in the new program. We had to investigate the system to understand the technology as well as those embedded business processes.

On the other hand, the vendor’s team was knowledgeable in the technology but had no knowledge, experience and skills in implementing such project in Vietnam. This problem would have affected any vendor as the project was the first implementation of a CBS in Vietnam. Thus, both implementation teams had significantly reduced capabilities at the beginning of the CORE project, the only way out for them was by learning in action.

However, learning during the implementation process was challenging for the individuals and groups involved in the project. They faced work overload, uncertainties, language difference, and staff turnover. To overcome the challenge, the learning process emphasized on learning focus, learning from others and learning by doing, as this quotation indicates:

We did not have education or training [on IS implementation]; therefore we learnt by doing and self adjusting. ... We exchanged ideas, for example, something we did but we heard that people said it was not reasonable and should be done like this or that. We had to think thoroughly what people said was reasonable or not. We adjusted the way of working for my self and for the whole group in the direction that we believed it would be better.

Learning happened during the change process in both technological domain and business domain and also during conflict resolution of the many issues that the client and the vendor had to confront.

Another key element to build the capability of the implementation team was its recruitment process. As the project developed through the phases, the implementation team changed in size and expertise composition. While recruiting more staff contributed to capability building, staff turnovers decreased the team’s capability, as experience and knowledge in the project were lost.

The implementation team was a key actor in the gap closing process. By proposing and implementing changes and solving the conflicts throughout the project, the team not only enabled the delivery of a complex project in difficult circumstances but also became knowledgeable and effective.

7 CONCLUSION

The paper discussed the process of packaged IS implementation in the context of Vietnam and confirmed the importance of the context to the implementation process, indicating areas for attention. This case study contributed to knowledge in the area of IS implementation in developing countries. By using grounded theory with case-based data, the study developed a theoretical framework called “gap bridging” with four sub-processes – gap discovering, goal forming, gap closing and capability
building. We used the framework to explain the roles and characteristics of these sub-processes as well as how they were interrelated. It also provides evidence of issues confronted while implementing and using application packages as means of modernisation of business practices.

However, this is only an emerging framework developed during the first phase of a major study, thus the framework needs more empirical work to further refine and develop its properties and produce a substantive theory. To achieve theoretical saturation, we are currently working with data from two subsequent cases. From the integration of these data with the relevant extant literature, we will propose a substantive theory on package IS implementation in developing countries.

The research in progress reported in this paper highlights an important set of processes that need to be considered when software packages developed and used in the developed world are implemented in developing countries. The topology of context-system gaps and gap discovering process developed in this study may be useful for client organizations in selecting and implementing such application packages. Moreover, lessons learned from the proposed gap-bridging process can be utilized to overcome prominent issues in countries such as Vietnam, and thus facilitating technological adoption in the developing world.

References


When standards is not enough to secure interoperability and competitiveness for European exporters

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0362.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Standards, Regulation / Deregulation, Implementation, Cross-national study</td>
</tr>
</tbody>
</table>
WHEN STANDARDS ARE NOT ENOUGH TO SECURE INTEROPERABILITY AND COMPETITIVENESS FOR EUROPEAN EXPORTERS

Stefan Henningsson, Center for Applied ICT, Copenhagen Business School, Howitzvej 60, 2000 Fredriksberg, Denmark, sh.caict@cbs.dk

Niels Bjørn-Andersen, Center for Applied ICT, Copenhagen Business School, Howitzvej 60, 2000 Fredriksberg, Denmark, nba@cbs.dk

Abstract

This paper addresses Information Systems (IS) standards from a managerial perspective. Standards can be seen as behavior regulations where there is a lack of hierarchical and market steering mechanisms. The use of government standards to harmonize IS is complicated, since in comparison to hardware and software products, IS are specific to the context in which they are implemented. Only hardware and software can be pre-fabricated. Based on a comparative case study of implementations of regulation from the European Commission seeking to harmonize European e-Customs, this paper develops a taxonomy of deviation sources for IS standards implementation divergence – reasons why implementations of IS standards differ. In addition to the previously known deviation sources: a) work process, b) data meaning, and c) data model, we found two new potential deviation sources in mode of transfer, and timing of implementation. Furthermore, all these sources seems to be of the nature that even the slightest deviation in any of the areas leads to the standard failing to meet its objective of increased efficiency with detrimental effects to the competitiveness of European exporters.

Keywords: Standards, Regulation / Deregulation, Implementation, Cross national study
1 INTRODUCTION

The European food industry face a number of major challenges in the immediate future. Consumers and governmental agencies demand increasing control, traceability and security from producer to the dining table (Tan et al., 2006). At the same time, a recent report from the European Commission (Wijnands et al., 2007) reveals that the European food industry is poorly equipped to meet the increasing worldwide competition. It will be severely hit if/when the EU realizes its plans to diminish subvention of the industry. The industry face the seemingly impossible demand of increasing control, traceability, and security while at the same time reduce the administrative burden of these functions for actors in the European food industry and its value networks in order to ensure competitiveness.

Adoption of standards is a major driver for industry productivity (Shapiro & Varian, 1996). Following industry-wide standards decreases the total resources spent on organizational transactions and interactions (Schilling & Steensma, 2001). Network effects of standards use enable increased efficiency as more and more actors adopt a standard. In a number of national and pan-European initiatives (e.g. FoodTrace, Traceback), Information Systems (IS) in various forms have been or currently are being introduced as the answer to the seemingly impossible demand for increased control and decreased cost of administration. In this paper we report on the efforts of harmonizing all European export procedures, using the same standard for e-Customs (here defined as “electronic transfer of information needed for customs processes”) in all European countries.

In the absence of hierarchical or market structures to enforce behaviour, management through standards is a possible way to achieve coordination among independent actors (Brunsson & Jacobsson, 2000). Standardizing through the creation of structures is also one of the most effective managerial toolboxes as suggested in the classic framework for exertion of power (Markus & Bjørn-Andersen, 1987) in the sense that standards are often perceived as less obtrusive than other regulations. But standardizing IS pose a particular set of challenges on the standardizing body due to its socio-technical nature (Hanseth et al., 2006). Introducing the European e-Customs standard as defined by Regulation (EC) No 648/2005 has the outspoken managerial objective to harmonize European customs processes by an IS standard, in order to provide a common interface to companies active in several European countries. However, as we shall demonstrate with our analysis, standardising IS is not as easy as one might expect in spite of the obvious advantages if all actors standardized.

The dairy company “Dairy Corp” has business in several European countries and export their products from two European countries. Both countries’ customs organizations, Custom A and Custom B, have adopted the European e-Customs standard launched by the European Commission. Yet, to report the customs data electronically Dairy Corp has to implement and maintain two separate interfaces to the two e-Customs implementations. Even though Dairy Corp is one of the largest dairy companies in Europe and must be regarded as one of the more frequent exporters in Europe, the cost of implementing and maintaining the new system extensions has been estimated to exceed the benefits in efficiency by e-Customs. When complying with the standard does not mean that the objective of harmonization is met it is a relevant question to ask why. In this paper we address the question why although Customs A and Customs B complies with the same e-Customs standard, the outcome in terms of IS is so different that they have to be treated as completely different systems. We refer to these reasons as the deviation sources of divergence in IS standards implementation. Our objective in this paper is to develop a taxonomy of sources for standards implementation deviation, which is a contribution of descriptive theory according to the theory classification by Gregor (2006).

The relevance of seeking to explain the reasons for standards compliance not fulfilling the managerial objective should not be limited to the context of customs. Research on IS and standards is limited despite the increasing importance in managing and using IS (Lyytinen & King, 2006). In addition, the existing research has mostly investigated content of new anticipatory standards rather than processes and factors that explains why and how standards emerge, are used and with which consequences
Among the explanatory works, standards development is almost exclusively the focus leaving the subject of standards implementation almost unexplored (Egyedi, 2007; Egyedi & Dahanayake, 2003; Söderström, 2004). The exceptions include a few writings on institutional factors of standardizing organizations that eventually leads to implementation problems (e.g. Egyedi, 2007; Damsgaard & Lyttinen, 2001; King et al., 1994). Not at least the past stream of EDI research has highlighted the need for synchronization between standards on IS and organizational level (e.g. Brousseau, 1994), but not specifically addressed the different levels as sources of deviations.

In the next section we will introduce the field of IS standards implementation and the current understanding of why implementations of standards may deviate. Thereafter we present the case of standardization of the European e-Customs, which introduces two yet undocumented sources of standard implementation divergence. Based on existing literature and our empirical investigations, we suggest a taxonomy for sources of divergence in standards implementation. Given the taxonomy we discuss and make tentative suggestions on how these problems can be met. Finally we provide conclusions for practice and academia.

2 MANAGEMENT THROUGH IS STANDARDS

In settings where hierarchical structures that enables enforcing behaviour lacks standards can be used as a managerial technique to regulate behaviour. The regulating mechanism is driven by a number of mutually independent actors identify themselves as a group and adopt a defined way of carrying out specific tasks (Brunsson & Jacobsson, 2000).

2.1 Management through IS standards

IS management has the ultimate objective of ensuring that the IS contribute to the organization reaching its goals (Alter, 2003). Drawing on general IS management research (e.g. Clemons & Row, 1991; Gottschalk, 2000; Kalling, 2003; Mata et al., 1995; Pyburn, 1983) and research on IS and IT governance (e.g. Van Grembergen, 2005; Willcocks et al., 2006; Brown & Grant, 2005), IS integration management may be defined as consisting of two main tasks: a) identify the different basic structural options that exist for IS management, and b) make and implement decisions upon an understanding of how the alternatives relate to organizational objectives. This is parallel to the two tasks of IS standardization: a) defining and developing IS standards, and b) adopting and implementing IS standards (c.f. Markus et al., 2006).

Defining and developing an IS standard is part of the options identification process. Few standards are the result of purely rational economic decisions (Backhouse et al., 2006). It is more than a little naïve to think that power and politics does not play an important role in the creation of standards (Cargill, 1989). However, this paper is not on how the objectives with an IS standards are developed, it is about how the objectives are realized in implementations. Standards implementation is thus part of the second managerial task above. The objective of standardizing is to harmonize and coordinate. Exactly how the harmonization shall be achieved is stipulated in the standards specification, but this might leave too much room for different interpretations leading to different implementations.

2.2 IS standards

IS standards are fundamental specifications that shapes the configuration of IS (Backhouse et al., 2006), and how the IS are used and managed in their organizational implementations (Hanseth & Braa, 2001). IS standards contain inscribed actions and processes that affects the way an organization is operating (Backhouse et al., 2006). The relationship between IS and IT standards can be compared to the discourse on differences between IS and IT in general (c.f. Iivari, 2003; Myers, 2003). IT standards are technical standards that can be defined as an agreed-upon specification for how to communicate and perform actions (Nickerson & zur Muehlen, 2006). Although various IT standards
can be important components of IS standards, just as IT frequently are important parts of IS. IS standards exists on an infological level (c.f. Iivari, 2003) that is based on the conception of information transfer from one human to another (Checkland & Holwell, 1988). An IS standard can thus be defined as a standardization of how information about something should be transferred from one individual to another individual who is in need of that piece of information.

A common distinction between standards is between de jure and de facto standards. Regulations from the EC can be seen as de jure standard, since the regulation works as what in the discipline of law is called “soft laws”, i.e. the use is mandatory but the means of enforcing use are limited. EC regulations can also be described as open standards, in contrast to proprietary standards, since the EC regulations are publicly available to everyone who wants to adopt them. The non-electronic standard that preceded the e-Customs initiative – the Single Administrative Document (SAD) – was adopted by several non-EU countries that had extensive trade with the EU.

2.3 Deviation in standards implementation

Stories of intentional and unintentional standards deviation are frequent in the business press. Microsoft’s use of Sun’s Java standard (Blundon, 1997) is one classical example of an intentional standards extension. Organizations may also choose to only adopt parts of an existing standard - intentional partial standards adoption. The reasons for intentional deviation from the standard might be both strategic and technical (Egyedi & Dahanayake, 2003). The focus of this paper, however, is unintentional standards deviation which is a less explored phenomenon (Egyedi & Dahanayake, 2003).

While IT software and hardware can be duplicated and sold as complying to a particular standard, IS are specific to the organizational or inter-organizational context in which they are implemented (Iivari, 2003). IS can be defined on three different levels: technical, infological, and organizational (Iivari, 2003). Consequently, deviation sources can be found on each of these three levels.

Egyedi and Dahanayake (2003) addressed deviation sources on a technical level. They found that these deviations stemmed from pure errors, ambiguities, and specification to parallel options, and functional deviation. Inter-organizational standards are based on a data model, describing which data should be transferred through the system are intended to eliminate such differences. If ambiguities, inconsistencies, or options exist in the data model, deviation in the implemented IS is possible.

The infological level of IS exists in a language context (Iivari, 2003). This means that the objective of an information system is to supply its user with information to support its activities. The interpretation and meaning of transferred data is another potential source of IS deviation (Gustafsson et al., 1982). In the case of e-Customs, deviations can occur because data fields such as “means of transport”, “description of goods”, and “exporter” are interpreted differently in each e-Customs implementation. An exporter might be the company that produced the goods and who is now sending it to the customer, but it might also be the logistics service provider that actually transports the goods across the border. In some cases it is even the receiving part who actually acts as exporter. The same problems exist for almost all data that is shipped: shall weight of goods be specified with or without wrapping? How does one write an appropriate description? Is the receiving part represented by its name or organizational number, and if so the organizational number in which country? Data transferred in the IS should have similar meanings to its users. Data meaning is another potential source of implementation deviation.

The organizational level of IS refers to a process view of how data is transformed through the systems to its users. IS can be seen as work systems that as one constituent part contains a work process (Alter, 2003). To complete transfer the user has to manipulate the system by a logic that is partly built into the system and partly defined by the implementation context. The process view addresses when to send which data and to whom. It is possible that national customs offices wants some part of the data model to be submitted to VAT-controlling authorities, some data to health authorities, and some data to the authority responsible for controls of dangerous goods. When and where data is transmitted is defined by the IS’ work process, which is a third source of standards implementation deviation.
Although the three levels of deviations described above are important sources for variations in implementation of standard, they are to the best of our knowledge not explicitly described in the literature as sources of deviation; they are only deductively derived as potential sources of errors. In the absence of previous works the deduction of a tentative framework based on closely related research is a possible way forward.

3 RESEARCH APPROACH

This paper presents a comparative case study, following the structured case study approach (Carroll & Swatman, 2000), of implementation of Regulation (EC) No 648/2005 in two European countries.

3.1 Research approach

Our purpose was to develop a taxonomy of sources for standards implementation deviation, which is a contribution of descriptive theory according to the classification by Gregor (2006). The taxonomy also contains an explanatory element since by using it, it is possible to explain why standards implementations differ from each other. But the study is explanatory in a general sense and not in the notation of explanatory case study as used by for example Yin (1994). As descriptive theory we suggest drawing on Gregor (2006) the usefulness of this type of theory should be evaluated by its completeness, distinctiveness, and simplicity. Completeness means that important categories or elements should not be omitted from the classification system, that is, the theory should be able to capture all important resources. Distinctiveness means that boundaries between categories and characteristics that define each category are clear. The empirical phenomena encountered should be possible to categorize according to these criteria without too much difficulty (Gregor, 2006). Simplicity refers to that by making a model or framework too elaborated or comprehensive makes it hard to work with and in the end makes it useless for its purpose.

The words: IS, standard, and deviation can be combined in a way that deviation refers to a discrepancy between what is described in the standard specification and how the IS performs, but it can also mean that two IS are performing differently although they both conforms to the same specification. Given the managerial perspective and the research question regarding the use of standards to regulate behaviour, it is the latter deviation that is of interest in this paper. The cases are thus the two implementations of Regulation (EC) No 648/2005 and the focus is the difference between these implementations which is addresses by a cross case comparison.

3.2 Data collection and data analysis

The essence of the structured case study approach is that it forms an iterative research cycle upon a formal theoretical framework. The framework creates a structure that permits collection of relevant field data and in later stages, the case results enables traceability of conclusions and theoretical generalizations (c.f. Yin, 1994). The framework in this study was based on the three levels of implemented IS (work process-, data model- and data meaning level) and the deduced sources of standards implementation deviation. Empirical data was collected along the three levels of IS in both standards specification and implemented IS.

Existing documentation about the European e-Customs initiative and the implemented systems are extensive and was an important source of empirical data. The standards specification Regulation (EC) No 648/2005 was a natural starting point. Other important documents produced by the EC include Regulations, Decisions, and Communications regarding customs code, as well as e-Customs and the Multi Annual Strategic Plan that sets the strategy for developing European customs. To grasp the implemented systems, user manuals, system documentation, specification of applied data model, and specifications of UN/EDIFACT and XML-schemas for data transfer were investigated. In addition, specifications of technical standards by ISO, DG/TAXAUD, and UN/CEFACT were investigated.
The second empirical level was participations in meetings, workshops, and interviews of IT personnel at Customs A and Customs B. In total 12 meetings and workshops were held on e-Customs with representatives from Customs A, Customs B, the EC and Dairy Corp. These were complemented by 10 semi-structured interviews to understand the specific details of the customs processes. Visits at Dairy Corp’s production facilities and export office complemented with the organizational context.

4 HARMONIZING EUROPEAN E-CUSTOMS

A number of contributing factors have forced the EC to reconsider the customs processes in the member countries. This section presents the story of how the EC seeks to increase control in customs processes and simultaneously lower the administrative burden for exporting companies in the EU by introducing a standard for e-Customs.

The competitiveness of the European food industry is weak from an international perspective. Productivity growth is less in the EU than in the largest competitive economies, the world market share is decreasing, and rate of innovation in the EU is minor compared to innovations in the US and Canada (Wijnands et al., 2007). There are several reasons for the deprived. One contributing factor is the administrative burden caused by comprehensive export processes. This non-productive cost is estimated to be as high as two percent of the total turnover for the average exporting company. The current complexity of export processes is manifested by the sometimes more than 20 km long lines of freight trucks waiting at the Finnish-Russian border or the more than 40 paper documents in four copies that has to accompany a container shipped by vessel (Tan et al., 2006). So far, only 4% of all trade documents exist only in electronic form. (Tan et al., 2006). This is a huge waste of paper (causing CO\text{2} emissions), a huge administrative burden, and in a number of countries, the passing of paper documents from one person to another is a potential for bribing/kick-back.

4.1 Harmonizing regulation and process standardization

Until now national customs authorities has been able to individually decide how they asses risks, control goods and carries out administrative tasks. A pan-European company thus have had to deal with national customs individually, complying to different legislation and processes. With the modernized customs code (TAXAUD, 2004) and the e-Customs standardization (Regulation (EC) No 648/2005) the EC seeks to harmonize all European processes and at the same time increase efficiency by the move to electronic processes.

With the focus on divergence in the standards implementations the objective of harmonization is in focus. Harmonization would imply substantial benefits as companies would be able to merge export departments and reduce the numbers of connections reporting electronically to e-Customs solutions.

4.2 Customs A’s e-Export system

Customs A introduced its first IT-supported export system in 1988. This pioneering system was solely used internally in the local tax and customs office by employees at the tax offices to register export statements in an IT system. From 1990, it became possible for companies to report their export statements into the system via an electronic connection. The possibility to report export statements via Internet using a web-interface or FTP was introduced in 2000, but adoption has been sparse.

In 2002 Customs A decided to reengineer the e-Export system in collaboration with an industrial advisory board. The objective was to achieve at least 95 % electronic export declarations. Two years after the implementation, practically all export declarations were reported electronically. As the EU has refined and elaborated its demands on national customs authorities, these have been successively implemented in similar but not identical e-Export system across Europe. Current alternations concerns complying with Regulation (EC) No 648/2005, which should be implemented by June 2009.
Figure 1 outlines the main steps associated with export declaration and control when exporting from Country A to outside the EU. Before the goods are transported across the border, the exporter has to notify Customs A with what is called a pre-advice message (step 1 in Figure 1). When this message has to be sent is dependent on the status of the exporter. If the exporter is certified by the customs as an approved exporter and use electronic transfer of declaration data, the data can be submitted one hour in advance. Based on the pre-advice, Customs A runs a risk analysis (step 2) and then either a) decides that additional control is required (step 3), possibly with a physical inspection at the loading place or b) releases to goods for export (step 4). Assuming that nothing is discovered in the additional inspection, the goods are finally released for export (step 4) and shipped across the border (step 5).

### Figure 1  **Generic export process in Country A**

Custom A’s e-Export system is a direct computerization of a previous manual and paper-based system. Messages sent have their paper equivalent which are still possible to use for companies who for some reason does not want to submit export data electronically. The paper-equivalent is the SAD. The SAD, specified by the EC in Regulation No 1875/2006 and No 648/2005, is a standardized form for customs data that is accepted in the EU, EFTA-countries and many other countries, including for example Russia. The data model is a field by field translation of the SAD-fields that Custom A was using prior to the computerization. Customs A e-Export system can be reached using standards like UN/EDIFACT messages or XML messages. However, it should be noted that using XML does not mean that the contents is standardized. There are no international XML standards for electronic customs declaration messages. Regarding UN/EDIFACT directory D96.B is used. The messages sent are based on the following UN/EDIFACT documents:
- UN/EDIFACT Syntax Rules
- UN/EDIFACT Syntax Implementation Guidelines
- UN Trade Data Interchange Directory Issue D96.B
- UN/EDIFACT Message Design Guidelines document
- UN/EDIFACT Procedure documents
- EWOS Technical Guide for EDI Message Profiles

Regarding data meaning, Customs A’s e-Export system benefits from being a direct derivation from the SAD, since the need for specifying data meaning was highlighted already when the SAD was introduced. A substantial amount of work was already done by UN/CEFACT, ISO, and the EC/TAXAUD to specify the meaning of data. Also for such fairly interpretable fields such as “product description” there are appropriate guidelines on how to provide and interpret data.

### 4.4 Customs B’s e-Export system

Deviation in the work process means that the basic process logic of how to make an export declaration could deviate. The EU regulation and directives does not contain a generic process model, specifying how an export declaration shall be made in each of the EU countries. However, customs processes are based on a long tradition and have many similarities on a high level: they are based on the same need of controlling what goods are leaving the country and for what destination. In the two countries
investigated, the processes of how to do an export declaration, meaning which steps are required to complete, are basically similar by tradition – not by formal standardization.

Customs B also used the SAD as template for creating the *data model* in their e-Export system. However, besides the mandatory fields the SAD presents a number of optional fields that national customs can demand if they want. In reality, this has resulted in the use of different XML schemas for the data transfer related to export declaration. In our analysis, it is very unlikely that any two European countries would have implemented exactly the same XML schema without directly collaborating on the matter. Consequently, although both e-customs solutions for the two countries are within the general EU regulations, the data transferred differs from country to country.

Regarding data meaning a substantial amount of work is already done by UN/CEFACT, ISO, and the EC/TAXAUD to specify the meaning of data. Already in the work of the SAD many of these issues were brought into daylight and had to be solved. Also for such fairly interpretable fields such as “product description” there are appropriate guidelines on how to provide and interpret data.

4.5 Consequences for pan-European business

The implications for European exporters are grave. For instance, when we look at country A, Dairy Corp has set up a system that links their enterprise system (SAP) to Customs A’s e-Export system. Dairy Corp uploads UN/EDIFACT messages via FTP to the EDI server of Customs A. But Dairy Corp is an approved economic operator in that country, and it is delivering their pre-advice notifications in batch mode every five minutes.

Customs B does not require Dairy Corp to report export declarations electronically, and Dairy Corp has chosen not to. The reason is that the development and future maintenance cost of extending the existing system with a module for handling also electronic export declaration in the second country would far overshadow the efficiency gains. Dairy Corp investigated the cost of developing such a module and were given the offer of about 100 000 Euro. Bearing in mind that it has to be considered normal that factual development cost is most often up to twice of calculated costs and that maintenance costs throughout the lifetime of a system is normally much higher than the development cost, implementing electronic export declarations to Customs B would be a fairly costly development. Therefore, Dairy Corp has chosen not to implement electronic transactions with Customs B.

Accordingly, the consequence for Dairy Corp could be summarized as follows:

- Two different export offices – one for each country
- Two different system modules reporting to Customs A and Customs B respectively
- Additional development costs of >100 000 Euro
- Additional maintenance costs, normally substantially larger than development costs for IT systems

5 ANALYSIS: STANDARDS DEVIATION

The objective of the above described regulations and directives were partly to harmonize export declaration processes in the EU. It should be said that the European e-Customs initiative is far from finalized, but directives regarding export declaration is already being implemented in the two investigate countries. Therefore differences in how these regulations were interpreted, adapted to the national context and implemented can already be seen. Our objective was to describe sources of standard divergence, and as viewpoint to identify divergence we used Dairy Corp that has to deal with both Customs A and Customs B. To assess the usefulness of our initial framework, the question is whether it can explain the differences between the two e-Export systems? Or, in other words, if harmonization in the three stipulated dimensions would imply Dairy Corp being able to deal with the two customs organizations using the same IS? The answer is no. As will be explain in the following the framework has to be extended with at least two empirically derived sources of deviation.
5.1 Extending the framework: empirically identified sources

As far as the authors of this paper has been able to identify, a fourth and a fifth source of deviation is largely ignored in the current regulations and directives. Even if the same data, with the same meaning is transferred there can be deviations in the way it is transferred. We refer to this as the *mode of transfer*. Comparing e-Export A and B we found that the mode of transfer differed in encryption required for submission and in requirement of digital signature. Currently neither Customs A nor B employs encryption, but considering the data submitted this is a must for future developments. In the country of Customs B, legislation demands identification of the person who submits data, therefore Customs B employs digital signature since long. Customs A does not. As will be further elaborated later, this source of deviation may be very difficult to deal with due to the roots in national legislation.

Finally, *timing* is an almost ignored aspect of implementation deviation. On a very general level the EC has set deadline for the phases of e-Customs, but on a practical level national customs are free to implement regulation and updates during a time span. As updates, modification and alternations are frequent during systems development processes and as the systems development process of European e-Customs is likely to be ongoing for a long time, timing of implementation creates at the same time different systems in all European countries.

5.2 A taxonomy of deviation sources in standards implementations

Recapitulating both theoretically and empirically derived sources of deviation we found five distinct sources of deviation: a) declaration process, b) data model, c) data meaning, d) mode of transfer, and e) timing in implementation. The five sources are summarized in Table 1 as *A taxonomy of deviation sources in standards implementations*.

<table>
<thead>
<tr>
<th>Deviation source</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Work process</td>
<td>Logic by which the IS is manipulated to fulfil its use</td>
</tr>
<tr>
<td>B Data model</td>
<td>Data processed by the IS</td>
</tr>
<tr>
<td>C Data meaning</td>
<td>The meaning of processed data for IS’ users</td>
</tr>
<tr>
<td>D Mode of transfer</td>
<td>How data is transferred through the IS</td>
</tr>
<tr>
<td>E Time of implementation</td>
<td>When the IS and changes to the IS are implemented</td>
</tr>
</tbody>
</table>

Table 1 A taxonomy of deviation sources in standards implementations

It is noteworthy that all five sources of deviations are so crucial that only the slightest deviation severely damages the harmonization idea. A new step in the work processes, slightly different data models, incompatible meaning of data fields, different encryption techniques, or different implementation dates would from the perspective of Dairy Corp force different versions of e-Customs connections. Although development cost for each deviation might be limited seen in isolation, the complexity of having to maintain 27 e-Customs solutions for the pan-European trader is a substantial overhead cost, which is reducing the competitiveness of EU companies. And it is problematic, since it can be avoided, if government agencies across the EU collaborated to a higher degree on standards. In the next we will discuss possible ways of dealing with deviations in the implementation of standards.

6 DISCUSSION: DEALING WITH DEVIATIONS

Trough the current harmonization-efforts European export declaration processes has becoming more similar. But, as shown above, close is not close enough if the ambition is to substantially lower cost for pan-European business. We will discuss three different ways forward that can lead to the foreseen benefits being met.
The first redesign solution is to seek to achieve what can be called “true harmonization”. True harmonization means that a company can use the same electronic connection, being based on EDIFACT, XML or whatever, to report to all European customs authorities. To achieve this, the level of specification on how to implement the customs side of the export declaration system must be substantially more elaborated. It must cover at least all five sources of divergence as described earlier in this document. Today, the functionality of the export declaration is standardized, but not how it should be implemented - therefore the difference in implementation. We suggest that a substantially more detailed blueprint, which is sufficiently detailed to eliminate all potential sources of divergence in order to achieve true harmonization must be developed, agreed upon and implemented.

A second redesign solution is to keep the current peer-to-peer structure, but create a “complete European export declaration protocol” that contains all data demanded by national customs organizations. That means that a company like Dairy Corp, when exporting through Customs A would submit all data that is needed by Customs A, but also all data that is needed by other European customs. Customs A would be able to receive such a message and extract the information relevant to them. The benefit for Dairy Corp is that they could send the same message to all European customs. The downside would be that they had to send irrelevant data at each submission, but since companies such as Dairy Corp do already have all necessary data in their internal systems, the cost of providing irrelevant data is marginal. There are several upsides for national customs in this suggestion, and the downside is very limited. Being able to distil relevant data from a complete European export declaration would not be a challenging task.

The third redesign suggestion is to allow companies to deal with only one national custom authority. For example, Dairy Corp could submit export declarations to Customs A regardless from which country they export. The problem of implementation differences is thus reallocated from business to customs. However, since it is the authorities who decide the rules of the game it is unlikely that they will take on additional responsibility although from a macro-economic perspective it would mean that fewer resources are being spent on setting up infrastructure for European wide e-Customs. The main reason why this would be less resource demanding is that implementing connections between the 27 member countries is far from the effort needed to interconnect the almost one million exporting companies with as much as 27 national customs each. This third option is in EC writings referred to as centralized clearance. This is a complex enterprise if not limiting the scope to just passing on information, but also include actually making the customs clearance in one single country and then divide fees, reimbursements etc according to predefined schemas. But once established, this would significantly improve the competitiveness of EU companies.

7 CONCLUSIONS

In our investigation we found that differences in the implementation of e-customs systems potentially could originate from five distinct sources. In addition to the previously known deviation stemming from a) work processes, b) data model, and c) data meaning, we found two new potential deviation sources in d) mode of transfer, and e) timing of implementation. All these sources seems to be of the nature that even the slightest deviation in any of the areas leads to the standard failing to meet its objective of increased efficiency.

The idea behind the above mentioned regulations of total electronic data transmission is that it will lead to increasing the security, higher transparency and less fraud while not adding to increasing the administrative burden on pan-European actors. So far the results have been that:

- Security, traceability and control have increased.
- The cost of reaching the new level of security, traceability and control without electronic submission of data seem to be higher in most cases.
- The additional burden on pan-European companies has not yet been lowered but rather increased as a result of the new export declaration processes.
We believe that to a large extent this is due to the current deviations in the implementations of export declaration processes, which are in a clear conflict with the idea of creating one common, inner European market place. With the current situation, if a company is active on one member country and would like to expand business with another production and export facility, everything else being equal, the cost is higher if placing that facility in another country than the country that the company is already exporting from due to the need of creating new export processes and developing a new electronic channel to the national customs.

We believe that this research highlights the importance of taking the step from research on IT standards to research on IS standards. As heavily emphasized in the IS literature, IT has no organizational value in its own right, but only as an enabler for information sharing. What we show in this research is that IT standards compliance is a prerequisite but not sufficient requirement for leveraging the benefits that is expected by standardization. More research is needed to understand the potential and problems with regulating behaviour by IS standardization.

More research is also needed on the complex issue of harmonizing the way national authorities work and interface to global business. Business is steadily getting more and more global while national agencies seem to be lagging behind in adaption. Failure to adapt might have far reaching effects as competition is extending to include competition between value chains, industries and regions.

Acknowledgements

This research is part of the integrated project ITAIDE (Nr.027829), funded by the 6th Framework IST Programme of the European Commission (see www.itaide.org). The ideas and opinions expressed by the authors do not necessarily reflect the views/insights/interests of all ITAIDE partners.

References


DO USERS GO WITH THE NEW WORKFLOW? FROM USER PARTICIPATION TO QUALITY OF WORK DURING WFM DEPLOYMENT

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0617.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>User participation, User Satisfaction, Workflow, Post-implementation</td>
</tr>
</tbody>
</table>
DO USERS GO WITH THE NEW WORKFLOW? FROM USER PARTICIPATION TO QUALITY OF WORK DURING WFM DEPLOYMENT

De Waal, Benny, Research Centre for Innovation and Business, University of Applied Sciences Utrecht, Padualaan 101, 3508 AA Utrecht, The Netherlands, benny.dewaal@hu.nl

Batenburg, Ronald, Institute of Information and Computing Science, Utrecht University, Padualaan 14, Utrecht, The Netherlands, r.s.batenburg@cs.uu.nl

Abstract

What determines the involvement of users to ‘go with the new workflow’ in the case of workflow management (WFM) deployment, and how is this related to their perceived quality of work? This key question is addressed in this paper. Customized for our empirical case context – a large Dutch social insurance organization that recently deployed a WFM system – we developed a conceptual model based on the models and concepts of DeLone & McLean, Hartwick & Barki, and Kappelman & McLean. Our model subsequently relates user participation, user satisfaction and quality of work, shortly after WFM deployment. The expected relations are tested by survey data collected from 143 end-users. Results of correlation and regression analysis show that (1) the degree of influence is a key determinant for the perceived quality of the system, (2) perceived system quality is a main driver for user satisfaction, and (3) user satisfaction is a main determinant of the users’ perceived quality of work.

Keywords: user participation, user satisfaction, workflow management systems, quality of work, public service organizations.

1 INTRODUCTION

Workflow management (WFM) and WFM systems have been around for a long time (cf. Kobielus 1997, Jablonski 1995, Hollingsworth 1994), but its adoption and deployment remains a relevant topic. An editorial article of Information & Management suggests that WFM has enjoyed regular attention over the years 1998 to 2005 with mostly three articles published per year (Palvia et al., 2007). Most recently, WFM has been extended to the broader concept of Business Process Management (cf. Smith and Fingar 2004, Weske et al. 2004).

The general premise of applying workflow management, is that the coordination of work can become easier, a higher quality of services is delivered, the work is executed more efficiently, and the process becomes more flexible (Reijers et al. 2003). Despite the successes (Fischer 2007), there are also many questions about the implementation of workflow management. Most of the problems appear to be of an organizational nature, rather than technical. Moore (2002) found that excessive activity automation and poor design of work assignment strategies are critical in workflow projects. In an influential study on the effects of workflow systems on organizations, Küng (2000) concludes that workflow design should actually be extended towards job design and organizational design. Most critically, employees outside the IT department should play an active role from the beginning of a workflow project. This essentially implies that user participation and user involvement in WFM deployment will improve its success.

Not much empirical research has been conducted on the success of WFM systems applying survey research among users. There seems to be a white spot with regard to the organizational benefits from WFM deployment in terms of the effects of WFM on the quality of work of end-users. This is
remarkable given the extensive lines of information systems (IS) and information technology (IT) research on the perceived usefulness and perceived ease of use of IS/IT (cf. Markus and Keil 1994, Venkatesh et al. 2003, Konradt et al. 2006, Sabherwal et al. 2006), and the line of social science research on job changes and quality of work in relation to technology and ICT (Zuboff 1984, Benders 1993, Mullarkey et al. 1997, Andries et al. 2002). Apparently, these two lines of research have developed quite separately from each other.

This paper aims to contribute to the different research problems and issues addressed above. By approaching the deployment and effects of WFM in organizations from both an IS/IT and social science perspective, we go beyond the studies that analyse WFM either from a technical, organizational or a job perspective. The lack of empirical research drives our objective to develop and test explicit hypothesis about the relationship between WFM deployment and the users’ quality of work. We firstly do this by building upon the research on upgrading and downgrading of skills and jobs in relation to IS/IT (cf. Grint and Woolgar 1997, De Witte and Steijn 2000, Bresnahan et al. 2002). Secondly, we develop testable hypothesis on the intermediating role of user participation for WFM success and quality of work. All hypotheses are part of an integrative conceptual model that provides an answer to the central question how WFM deployment effects the perceived quality of work of users. A key element of this paper is the data by which our model and hypothesis are tested. Contrary to the ‘common’ structure of most papers, we start by describing the case context of an empirical study on a large Dutch social insurance organization that recently deployed a new WFM system. The the conceptual model and its underlying expectations are tested by data collected from a survey among 143 end-users of the WFM system. After a discussion of the results the paper concludes with some recommendations for further research.

2 THE CONTEXT: WFM DEPLOYMENT AT A LARGE DUTCH SOCIAL INSURANCE ORGANIZATION

The empirical research for this paper is based on a case study and survey within a large Dutch organization for the execution of employee (social) insurances. Since its foundation in 2002, as a result of a merger of five public organizations, the formal mission of this organization is “to stimulate people to work, and if work is impossible to provide a temporary income”. The organization is an independent policy agency that resorts under the Dutch Ministry of Social Affairs. It is a large administrative organization that has a national task to execute several social security laws. The organization controls all administrative and financial processes to execute these laws, and is dedicated to control misuse and fraud. A second task is to realize the reintegration of (partly) unemployed and disabled employees. For this task, it collaborates with private health companies. The organization has 17,292 employees amounting to 14,563 Full Time Equivalents (FTEs) in 2007. The research took place within the so-called Social Medical Function (SMF) domain. The focus of this domain is to examine the medical- and work capabilities and to judge the claims of the clients within the scope of the national laws. Within the SMF-domain the number FTEs in 2007 is 6,359.

In 2005, a project was started to support the front office to integrate existing applications. As part of that project a WMF system was developed. This WFM system was primarily aimed to streamline the core administrative processes of the social security laws within the SMF organization and improve its quality in terms of compliance with rules and procedures that are mostly of a legal kind. The WFM system consisted of a standard software package with custom-build components to support the transition of information and tasks from one ‘resource’ to another. A resource can be a person (i.e. an employee) or a system (i.e. an application). Besides achieving more efficiency and higher quality of services, the goal of the WFM project was to reach uniformity in case handling. To support this, the concept of teamwork was introduced. Teams were created by joining five different types of employees namely team-support employees, process support employees, work/job consultants, insurance medical doctors, and sometimes medical assistants. To ensure that the composition of the team fitted with the system, the size of a team was flexible. It was intended that a proper composition of the team will increase the (user) performance of the new WFM system.
The project organization consisted of a steering group supported by a sounding board and a project quality group. The steering group was responsible for the end result and consisted of senior users, senior developers, a representative of the user organization, a project advisor and the project manager. The sounding board was responsible for the definition of the requirements of the user organization and the acceptation of business products. This group consisted of managers and staff employees of the user organization, one end user, a quality advisor and development manager. It was agreed that members of the sounding board informed the employees they represent. The project quality group was responsible to monitor the quality aspects of the project and the deliveries. Members were quality advisors and domain experts. The overall responsibility of the project manager was to assure that the project delivers the right products within the appointments that have been made.

In 2007 the WFM system was nationally deployed along all establishments of the Dutch social insurance organization. During the implementation phase at all locations, new releases of the WFM system were developed. To set priorities for the releases a user group was installed with members of the different function groups of the user organization. Priorities were discussed with members of the design team. Users could send issues to the user group and received information about its progress. In this phase the users were actually confronted with the new system for the first time. From the start it was planned that the use of the new WFM system was mandatory. During the introduction users received professional training lasting one to two days. After that, training on the job was supported by a trainer and a mentor. In the beginning it occurred that the new system sometimes failed. Besides that, during that period there was a huge workload because of new legislation. Also, there was significant time pressure to reach productivity norms. In some cases, users were allowed to use the old systems if the new system failed – or if it was easier to achieve the production norms.

This case context should be explicitly taken into account, as the survey that is reported in this paper took place 6 to 12 months after the deployment of WFM system. The specific conditions as just described were therefore incorporated in the development of the conceptual model from which expectations about the interrelation between this new WFM system and users’ quality of work and their role as user participation is derived.

3 A CONCEPTUAL MODEL OF USER PARTICIPATION, USER SATISFACTION AND QUALITY OF WORK

3.1 The D&M model

There is an expanding field of research on the measurement and determinants IS/IT in organizations. Several theories and models from several perspectives have been developed in this domain. One stream of research departs from models on adoption and use of IS, with theories such as Diffusion of Innovations (DoI), Theory of Reasoned Action (TRA), the Technology Acceptance Model (TAM), Social Exchange Theory, Theory of Planned Behaviour (TPB) and Task Technology Fit (cf. Van de Wijngaert et al. 2008, Dwivedi 2008). The major stream of research however, is dedicated at exploring the direct determinants and measurements of success (or failure) of IS. Important studies in this area are Delone and Mclean (1992, 2003), Seddon et al. (1999), Smithson and Hirschheim (1998), Grover et al. (1996) and Sauer (1993). For the study reported in this paper, the departure point is the leading, and probably most-cited, model in the research of IS success, namely the updated DeLone and McLean Model of IS Success (see Figure 1).

The original D&M Model was first formulated in 1992 (DeLone and McLean 1992) and updated in 2003 (DeLone and Mclean 2003). In the updated D&M Model the quality of a system depends on three dimensions namely ‘information quality’ (e.g. completeness, ease of understanding, relevance), ‘system quality’ (e.g. usability, availability, reliability) and ‘service quality’ (e.g. empathy, responsiveness). Each dimension will affect ‘use’ or ‘intention to use’, and ‘user satisfaction’. ‘Use’ is a behaviour, whereas ‘Intention to use’ is an attitude. ‘Use’ and ‘user satisfaction’ are interrelated. If there is a positive experience with ‘use’, this will lead to greater ‘user satisfaction’. Similarly, increased ‘user satisfaction’ will lead to increased ‘intention to use’ and thus ‘use’. As a result of this ‘use’ and ‘user satisfaction’, certain ‘net benefits’ will occur. The ‘net benefits’ can be individual or
organizational. The ‘net benefits’ will in turn influence, positive or negative, ‘intention to use’ and ‘user satisfaction’.

Concerning the case context of the study as described in the previous section, the D&M model was adapted in two ways. First there is focus on system quality only, excluding information quality and service quality as other drivers for intention to use and user satisfaction. The reason for this is that interest for the research reported here lies in the evaluation of the deployment of a specific application (a WFM system), and a definition of the perceived quality of work of users as the net benefits. As was mentioned in the previous section, during the introduction users were for the first time confronted with the WFM system. System quality is therefore the most significant reason for (potential) changes in use, user satisfaction and their effects. Although use was mandatory, during the deployment there were several reasons to escape from this rule, e.g. system failure and production norms. The second adaptation was therefore to adjust the D&M model by selecting intention to use, not use, that is both related to user satisfaction and also to net benefits. Because this was a study of the deployment of a new WFM system that was not fully in use at the moment of data collection, its actual use is a less valid measurement since this depends on many other practical factors such as technical implementation problems and other delays. Therefore, the original D&M model that assumes that intention to use leads to actual use, with feedback effects from actual use to user satisfaction and intention to use was not used.

3.2  Extending the D&M model with Hartwick and Barki

It would be prudent to refer to the very early notion of Conway (1968), or more prominently Orlikowski (1992) that an important issue to the use and success of IS/IT is in the interaction between designers and users. In the study reported here the process between designer and user was explicitly included, and the D&M Model was extended by bringing in the concepts of user participation and user involvement. Recalling the case context, this is relevant as during the process of deployment of the WFM system users could bring in issues for new releases through user participation by delegation.
recently were confronted with the deployment of a “new, business oriented, IS application in the near future” (p. 446) in their organization, is shown in Figure 2.

They define user participation as the observable behaviour of system users in the information system development process, i.e. their participation in information system development and implementation activities. Barki and Hartwick (1994) identified three dimensions of user participation: overall responsibility, user-IS relationship and hands-on activity. Overall responsibility refers to user activities and assignments reflecting overall leadership or accountability for system development project. User-IS relationship refers to development activities reflecting user-IS communication and influence. Hands-on activity refers to specific physical design and implementation tasks performed by users. Next, user involvement is defined as a psychological state of system users, i.e. as the importance and personal relevance of a system to use (Barki and Hartwick 1989). The authors differentiate this psychological state from another psychological state, namely attitude. This is generally conceptualized as an affective or evaluative judgment of some artefact and can be measured to locate one’s position on a bipolar affective or evaluative dimension, e.g. bad/good. Finally, they incorporated in their model Fishbein and Ajzen’s Theory of Reasoned Action (1975). From their study (Hartwick and Barki 1994), it is claimed that user participation influences user involvement and that the effect of user participation on intentions and use is mediated by the psychological constructs of involvement, attitude, and subjective norm. User participation and involvement are particularly important for the voluntary users of a system.

For this study the Hartwick and Barki (1994) model is used in particular the three user participation concepts and their relationship with IS success, i.e. intention to use, user satisfaction, and quality of work (as perceived by the end-users). The TRA concepts and actual use concept from their model is not used. As described previously, external factors in this case organization strongly hindered the actual use of the system during the WFM deployment. This makes the TRA constructs (attitude concerning use and subjective norm concerning use) as well as the actual use concept less relevant as concept to be included in our conceptual model.

3.3 Extending the D&M model with Kappelman and McLean

While Hartwick and Barki (1994) relate user participation to the Delone and McLean concept of intention to use, Kappelman and McLean (1991) shows that there is also a relation between user participation and user satisfaction. As stated before, the case context in the study reported here concerns a period where participation and involvement is critical. Therefore it is also important to link user participation to the concept of user satisfaction of the D&M IS Success model. The model that Kappelman and McLean (1991) tested is called the Behavioural-Attitudinal Model and is depicted in Figure 3. The study was conducted at 52 branches of a regional interstate bank holding company, during the installation and conversion of the information processing sub-system at each of the branches. Although the operationalization of user participation was not the same as Barki and Hartwick, the underlying items were based on the same literature (Olson and Ives 1980, 1982).

As in the case reported in this paper users in the Kappelman and McLean (1991) study participated only in the system installation and conversion phase, or what they label as ‘later-phase user participation’. This concept corresponds with the ‘hand-on activity’ dimension of Barki and Hartwick (1994). User involvement in the model of Kappelman and McLean also overlaps with the model of Barki and Hartwick (1994). Finally we adopted the concept of user satisfaction. Kappelman and McLean (date) convincingly showed that user participation does induce user involvement, i.e. that involvement actually intervenes in the participation-satisfaction relationship. Further, they argued that involvement is more important in understanding user satisfaction than user participation.

Figure 3. The Kappelman and McLean Behavioral Attitudinal Model.

As in the case reported in this paper users in the Kappelman and McLean (1991) study participated only in the system installation and conversion phase, or what they label as ‘later-phase user participation’. This concept corresponds with the ‘hand-on activity’ dimension of Barki and Hartwick (1994). User involvement in the model of Kappelman and McLean also overlaps with the model of Barki and Hartwick (1994). Finally we adopted the concept of user satisfaction. Kappelman and McLean (date) convincingly showed that user participation does induce user involvement, i.e. that involvement actually intervenes in the participation-satisfaction relationship. Further, they argued that involvement is more important in understanding user satisfaction than user participation.
3.4 A conceptual model of user participation, user satisfaction and quality of work

The results of the adaptation and extension of the D&M IS Success model are combined and jointly depicted in Figure 4 below. In addition to the previous sub-sections, a number of other decisions were taken in designing the final conceptual model to suit the case study context.

To measure user participation, the concept of ‘overall responsibility’ was replaced by the concept of ‘degree of influence’. As will be described in the next section, the actual measurement was how satisfied users were with the degree of influence offered by the sounding board and user group. Further, user satisfaction is measured with the User Information Satisfaction (UIS) instrument as used by Shaw et al. (2002). With this instrument the perceived service satisfaction and information satisfaction is measured, because this is relevant in view of the D&M IS Success model. The net benefits in this study are defined as the quality of work of the WFM system users. As mentioned in the introduction, WFM systems are designed to coordinate processes and can potentially make work more efficient by improved information and steering of tasks and roles. WFM systems potentially effect the autonomy and workload of users, as well the information they receive. Therefore, it was decided to specifically focus on autonomy, workload and information as the main indicators of the quality of work of the WFM users. Further elaboration of measurements are presented in the next section.

The conceptual model in Figure 4 below depicts the expected relationships. From left to right, between: (1) the three user participation concepts on the one hand, and the three concepts system quality, attitude toward the system and user involvement on the other, (2) subsequently between these three concepts and intention to use and user satisfaction, and (3) between intention to use, user satisfaction and net benefits of the WFM deployment being the three dimensions of quality of work. By colours, Figure 4 also depicts the theoretical origins of the concepts and their interrelationships.

The main hypotheses from our conceptual model that will be tested are:

1. user participation is positively related with system quality, attitude toward the system and user involvement;
2. system quality, attitude toward the system and user involvement are positively related with intention to use and user satisfaction;
3. intention to use and user satisfaction are positively related to net benefits.

4 THE SURVEY AMONG THE WFM SYSTEM USERS

The survey research took place at six offices of the Social Medical Function (SMF) domain and was conducted in May and June 2008. The focus of this domain is to examine the medical- and work
capabilities and to judge the claims of the clients within the scope of national laws. Within the SMF-domain over 7,000 UWV employees are active. Within these locations, the sampling was stratified according to the four main job categories that would jointly become team members as main users of the system. The job categories included team support employees, process support employees, work/job consultants and insurance medical doctors.

The (web-)questionnaire was sent out through internet. As not all team support employees and process support employees had access to the internet some completed the questionnaire from another computer or a hardcopy. A total 143 employees returned a complete questionnaire. The response rate was 30%, not biased according to job category as the main stratification criteria. The sample consisted of 13% team support employees, 21% process support employees, 42% job consultants and 24% medical insurance doctors – which is largely representative for the total organization. With regard to some other background characteristics, the sample can be described as:

- 60% male, 40% female;
- average age is 48.5 years;
- 27.3 percent of the respondents has an university degree;
- respondents have 19 years average work experience at the organization;
- 35 percent stated that they have more than average computing skills, compared to their colleagues.

4.1 Measurements

Below the measurements of the elements of the conceptual model (Figure 4) is described from right to left. The results of construct validity and reliability testing is presented.

Perceived Net benefits (i.e. Quality of Work) – As explained in section 3.4, the focus is on the three main characteristics from the total quality of work concept that are most relevant in relation to the potential effects of the WFM system under study, i.e. Workload, Information of Work, and Autonomy. The characteristics were measured with scales developed by Schouteten and Benders (2004) and Dhondt et al. (2002). Principle Component Analysis (PCA) shows that the 11 items for workload all contribute to a 1-factor solution with an eigenvalue of 5.58. Similar results were achieved for the 6 items for information of work (eigenvalue 2.87), and the 9 items for autonomy (eigenvalue 4.04). The overall accounted variance of the three 1-factor solutions was 50.8%, 47.8% and 44.9% respectively. In addition, Cronbach’s alphas were calculated for the three item sets. Their scale reliability was confirmed as the Cronbach alphas were 0.90, 0.77 and 0.84 for the Workload, Information of Work and Autonomy constructs.

User Satisfaction – To measure user satisfaction the Information Satisfaction and Service Satisfaction scales developed by Shaw et al. (2002) was used. Factor analysis was performed to analyze the construct validity of the group of 9 items. PCA resulted into a 1-factor solution with an eigenvalue of 4.50, accounting for 50.0% of the overall variance. Reliability of this User Satisfaction scale was confirmed by a Cronbach’s alpha of 0.87.

Intention to Use – The ‘intention to use’ concept can be defined by one relatively clear item, i.e. “If the system was not mandatory, I would still use it”, as suggested by Seddon and Kiew (1996) as one of the first scholars with regard to this issue. This single item holds 4 answer scale categories, as most of the other items used for the model.

System Quality – For the purpose of this study, specifically the aspects Usability and Usefulness of DeLone and McLean’s concept of system quality are measured. Usability is measured by 6 items from a (Dutch) questionnaire developed by Tijdens and Steijn (2002) and the classic complexity scale of Thompson (1991). Usefulness is measured by 6 items from the classic compatibility scale of Moore and Benbasat (1991) and the work-with-computer-scale from Tijdens and Steijn (2002). Factor analysis was performed to analyze the construct validity of the total set of 12 items. The one-factor solution from PCA holds an eigenvalue of 6.43, accounting for 53.6% of the overall variance. Reliability analysis confirmed the scalability of the 12 items, as Cronbach’s alpha is 0.91.
Attitude towards the System – Attitude toward the system is measured by four items from a scale developed by Hartwick and Barki (1994). Different from the previous items described, they developed as 7-point items by which users could judge the IS/IT application by contrasting labels as “good/bad”, “terrible/terrific”, “useful/useless” and “worthless/valuable”. PCA showed that the four items clearly load on one factor, with an eigenvalue of 3.21 and an overall variance of 80.2%. As can be expected, the reliability of the scale was also confirmed by a Cronbach’s alpha of 0.92.

User Involvement – To measure User Involvement, five items from Hartwick and Barki (1994) were used. They defined the User Involvement concept by items that express the judgment of the IS/IT application in terms of “important/not important”, “relevant/irrelevant”, “fundamental/trivial”, “essential/non essential” and “significant/insignificant” on a 7-point scale. For this dataset, PCA on the items showed that there is one strong latent factor (eigenvalue was 3.36, overall variance 67.3%). Reliability analysis supported the scalability of the item set (Cronbach’s alpha = 0.88).

User-IS relationship – The User-IS relationship scale, as part of the User Participation concept, is based on the work of Hartwick and Barki (1994) as well. They developed 7 items to measure 7 different types of participation that could be either scored as “no” or “yes”. As these items refer to specific actions that, in many cases, were actually not applied in IS/IT project (for all kinds of reasons), it was decided to sum the “yes” scores of these items instead of applying PCA and reliability analysis which is more appropriate for ‘subjective’ items as opinions and Likert-type statements.

Hands-on Activity – Hands-on Activity concerns a scale that is comparable with the User-IS relationship scale, also developed by Hartwick and Barki (1994). Likewise, the sum of the 8 items they defined was used, addressing 8 different types of hands-on activities to let users participate in the IS/IT development.

Degree of influence – Finally, the Degree of influence as developed by Lynch and Gregor (2004) was applied. Using their four (4-point) items, PCA on the dataset confirmed that these items can be aggregated into one latent factor, with an eigenvalue of 2.67 and overall variance of the 1-factor solution of 66.6%. Reliability analysis confirmed the scalability of these four items: Cronbach’s alpha is 0.83.

5 RESULTS

In Figure 5 below the results of three OLS regression analyses that correspond with the three decomposed parts of the conceptual model is presented. These are labelled as ‘Model I’, ‘Model II’ and ‘Model III’. The one-way-directed arrows in the figure represent the significant (standardized) regression (beta) coefficients (p<.01), while the two-way-directed arrows represent the significant correlations (p<.01) between the independent variables within the subsequent regression models. For all three OLS regression models, the potential problem of multicollinearity was investigated by computing VIF-factors for each predictor in the regression model. Although in some cases correlations between independent variables were relatively high, VIF factors in none of the three models exceeded 5 – a commonly applied rule of thumb.

The results from regression Model I show that primarily User Satisfaction holds a significant relationship with each of the three Quality of Work dimensions. Although Intention to Use is strongly correlated with User Satisfaction, this holds no significant relation with the users’ Quality of Work. This might be partly due to the problems with the actual use of the WFM system at the social insurance organization as we described in section two. On the other hand, this result also supports that satisfaction through ‘actual use’ is an important determinant. It is confirmed that satisfied users significantly experience higher Autonomy, Work-related Information and less Workload. The explained variance (adjusted R$^2$) of the three regression models is 7.4% (F=6.613, df=141, p=.002), 16.4% (F=14.802, df=141, p=.000) and 15.0% (F=13.424, df=141, p=.000) respectively.

With regard to model II, the three assumed determinants of User Satisfaction and Intention to Use are strongly inter-correlated. In joint regression analysis however, only System Quality holds a significant relation with Intention to Use and User Satisfaction. Apparently, the usability and usefulness of the WFM system dominates the actual satisfaction and intention to use of the users at the social insurance,
while their attitude and involvement only supports this, i.e. not significantly driving it. Given the context as described earlier – a new WFM system fully in the phase of adaptation and deployment, combined with limited use of the WFM system – this result is as expected. The explained variance of both regression models (adjusted R²) is relatively high: 47.8% (F=11.976, df=142, p=.000) for Intention to Use, and 45.0% (F=37.843, df=142, p=.000) for User Satisfaction as a dependent variable.

Model III finally, shows that Degree of Influence is the prominent determinant of User Involvement, Attitude towards the System and System Quality. Again, the three predictors of Model III do correlate and clearly coincide, but ‘in competition’ the perceived Degree of Influence is the clearest driver, probably as the participation items of Hands-on Activity and User IS relationship were lacking in practice as well. The explained variance of the three regression model varies: 23.5% (F=15.580, df=142, p=.000) for System Quality as a dependent variable, 23.4% (F=15.445, df=142, p=.000) for Attitude towards the System, and 10% (F=12.660, df=142, p=.000) for User Involvement.

![Conceptual Model](image.png)

Figure 5. The conceptual model tested: reliability of the item scales (stats within the boxes), significant correlation coefficients (two-way arrows between the boxes), and significant standardized regression coefficients (one-way arrows between the boxes) for three sets of dependent variables.

### 6 CONCLUSION AND FURTHER RESEARCH

This paper presents and effort to build upon two research traditions that both deal with the question of how the success of IS/IT deployment can be understood in terms of individuals benefits, more specifically the users’ quality of work. In developing a conceptual model, different models and concepts from the IS-tradition were brought together (i.e. the model IS success model of DeLone & McLean, extended by model elements from Hartwick & Barki, and Kappelman & McLean). In addition, concepts from the social science tradition on quality of work and its relevant dimensions in relation to IS use, more specifically WFM systems were used.
Based on the context of a large Dutch social insurance organization that recently deployed a new WFM system for all their employees, a conceptual model was customized. In this model it is claimed that the quality of work of users is directly related to their intention to use the WFM system and their user satisfaction. Next, it was hypothesised that these two concepts in turn are related to the quality of the (WFM) system, the users’ attitude and their user involvement. Finally, the three types of user participation are included into the model as drivers of the perceived system quality, system attitude and user involvement.

Data collected from a survey among 143 users within the case study organization enabled the construction of valid and reliable measurements and constructs as defined in the conceptual model. Correlation and regression analysis showed that for each of the three sub-models significant relations were found, although not all of the hypothesized relationships were confirmed. It was shown that (1) the degree of influence is a key determinant for the perceived quality of the system, (2) perceived system quality is a main driver for user satisfaction, and (3) user satisfaction is a main determinant of the users’ perceived quality of work. From these results it could be argued that the concept of (perceived) system quality appear to play a central role. On the one hand, system quality is the main determinant of user satisfaction which subsequently holds the strongest relation with the users’ quality of work. On the other hand, system quality is as strongly related to user participation as user involvement. One way to further validate our findings is to estimate the complete model using Structural Equation Modeling (SEM). An additional method to validate our scale construction is to apply factor analyses over all items.

If these (quantitative) results are interpreted against the (qualitative) case context as described earlier in this paper, some further conclusions can be drawn. While the value of the WFM system within the Dutch social insurance organization is recognized, its importance for all processes and tasks has not been proven yet. The disruptions during the introduction of the WFM system caused employees to kept working with the old system to keep up with their targets. This resulted into a lack of use and experience with the new system. Although the benefits and added value of the new system still remain unproven, it can be seen from the survey results that user satisfaction and system quality do matter. If users increasingly value the quality of the new WFM system, their satisfaction and perceived quality of work will increase – and most likely their productivity.

With regard to participation and involvement, it was determined that users at the Dutch social insurance organization were not much directly activated during the deployment of the WFM system. Instead, their participation was mainly organized through delegation, and in addition users were frequently informed during the project. From the analysis it was found that if users experience this degree of influence, this holds a positive relation with the experienced quality of the system, user involvement and attitude towards the system. As this subsequently increases the users’ satisfaction and quality of work, this influence can thus be considered as an important ‘background driver’ for the benefits of the new WFM system. These results support the basic statement that employees should play an active role in WFM projects (Küng 2000). High impact IS/IT application as WFM can indeed upgrade jobs, but only if this is mediated with user participation.

The measurements and analysis of user participation in the context of this case of WFM at a large Dutch social insurance organization also provides new roads and ideas for further research. In particular, finding the ‘right’ type of user participation for the ‘right’ type of user/system combination can be a crucial new research question. It was found that ‘indirect’ ways of participation hold positive relations with user involvement and perceived system quality, while direct participation activities showed no relation with these factors. This is a particularly interesting result for larger organizations. Obviously, with larger number of users the costs of user participation will be higher too, challenging new and ‘efficient’ ways to let users participate during deployment projects. Further research within other organizations is needed to validate if the size of the organization is indeed an important condition for selecting the ‘optimal’ type of user participation (cf. Markus and Mao 2004, Nordheim and Nielsen 2008). Finding the right mix of user participation makes it possible that users go with the (new) workflow, and subsequently experience an improvement of their quality of work as well.
References


**COMPONENT-BASED PROCESS MODELLING IN HEALTH CARE**

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0131.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Process Management, Workflow, Design research, E-health</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
COMPONENT-BASED PROCESS MODELLING IN HEALTH CARE

Baacke, Lars, University of St. Gallen, Mueller-Friedberg-Strasse 8, CH-9000 St. Gallen, Switzerland, lars.baacke@unisg.ch
Mettler, Tobias, University of St. Gallen, Mueller-Friedberg-Strasse 8, CH-9000 St. Gallen, Switzerland, tobias.mettler@unisg.ch
Rohner, Peter, University of St. Gallen, Mueller-Friedberg-Strasse 8, CH-9000 St. Gallen, Switzerland, peter.rohner@unisg.ch

Abstract

Structural changes and increasing market dynamics in the health care sector intensify the hospitals’ need for cost-savings and process optimization. A first step is the documentation of processes in order to clarify the actual needs. As in health care processes are rather complex and often different players with divergent demands are involved, a disciplined approach to effectively and efficiently model processes is required. For this purpose, in this contribution a component-based modelling approach is presented and applied.

Keywords: Business Process Management, Design Research, Health care, Workflow.
1 INTRODUCTION

The health care sector is confronted with unsatisfactory performance in costs (i.e. health expenditure) and quality (i.e. effectiveness of service delivery and medical malpractices) for years (Porter & Olmsted Teisberg, 2004). Affected organisations counteract these deficiencies by various changes on strategic, procedural, technological, and cultural level. Horizontal mergers, outsourcing and shared services, public-private-partnerships, managed competition, integrated and preventive care, new medical procedures and information systems, clinical information interfaces, and digital imaging are only few examples for the dynamic environment and related challenges of health care organisations (Denis et al., 2000; Ginter et al., 2002; Ham, 2003). However, transformation projects often fail or at least do not meet raised expectations due to unexpected influences and obstacles (Berg, 2001; Lorenzi & Riley, 2004). These are in most cases the result from the high complexity of change projects. Complexity is related to the number of elements of a whole system as well as their interrelations (Schneberger & McLean, 2003). Looking at an enterprise or hospital from an intra-organisational perspective, this understanding covers organisational and procedural structures as well as the information processing landscape and the target systems of employees (Mettler et al., 2008). From an inter-organisational perspective, involved stakeholders (e.g. patients, health care service providers, suppliers, administrations, competitors, insurance providers), regulatory conditions, requirements of the market (e.g. collaborative service production), etc. have to be considered (Denis et al., 2000). These circumstances point out the complexity of health care organisations (Anderson & McDaniel, 2000) and respective changes, and clarify the need for transparency, in order to identify potential areas of modernisation as well as to derive concrete actions for a successful transformation.

Prerequisite for transparency is the comprehensive understanding of relevant structures and influences basing on respective documentations. However, looking at current change projects (regardless whether technically, organisationally or strategically motivated), required information is costly gathered primarily within those projects by internally responsible persons or expensive external consultants. Resulting documentations are closely aligned to the goals of the respective project and their quality, to large extents, depends on knowledge and experiences of the modeller. Thus, they are often not reusable for other purposes and integrable to existing documentations. A model-based representation of knowledge provide the foundation for documenting organisational architectures in a standardised, comprehensible, and interoperable way which is the basis for a cross-organisational identification of deficiencies and the design of a targeted state, for communication with involved stakeholders as well as for derivation of the transformation process itself.

Although models are well known and accepted for abstracting real life systems in medical research, they are not yet broadly diffused for documenting business-related knowledge in health care organisations. In contrast to medical or biological models which are widely accepted (e.g. Rector et al., 1997; National Library of Medicine, 2005), business-related models are applied to homogeneous groups of organisations, single organisations or even specific units. While re-use is limited, effort for creating new business-related models is unequally high. In addition, business process models typically involve several people with different perspectives and comprehension (Vassilacopoulos & Paraskevopoulou, 1997). However, there are much more influencing factors that constrict broad diffusion of process modelling, such as the confusing variety of modelling notations for similar purposes, the incompatibility of different tools and the integrability of respective models, the dependency on modellers’ knowledge and experiences, the diverging levels of detail and abstraction, the inconsistent use and poor definitions of organisational terminologies, or the lack of maintenance and up-to-dateness which are limiting re-usability (Chen et al., 2008; Curtis et al., 1992; Christov et al., 2007).

This contribution addresses this heterogeneity and uncertainties by proposing a component-based modelling approach. The composition of predefined Generic Activities (GAs) establishes a common terminology and understanding of what a process is. It aligns the levels of detail and abstraction, and
finally enables compatibility of independently created model parts as well as their combination to complete processes. For that purpose, a proven concept for development of GAs is being adopted from public sector research, and its applicability in the health care domain is discussed.

After the description of the research method in the subsequent section, the conceptual foundations and the procedural model for identifying the demanded model components are presented in section 3. Then, the identified model components and practical aspects of applying GAs for modelling processes of the health care domain are discussed in section 4 by exemplifying a concrete model instance. The main findings and topics for further research are summarised in the last section.

2 RESEARCH METHOD

As the documentation of organisational knowledge is a rather pertinent and practical issue, engaged research is needed in order to provide rigorous solutions for this relevant problem. A theoretical basis that serves both relevancy and rigour of research and requires engaged research is that of design science research (Hevner et al., 2004; March & Smith, 1995; Venable, 2006). While natural sciences try to explain and predict behavioural aspects of the reality by developing and verifying theories, design-oriented research aims to build and evaluate innovative artefacts, in order to extend existing capability limitations. Artefacts represent the actual results of a design process. They can be characterised as “constructs, models, methods, and implementations” (March & Smith, 1995). The model components presented in this contribution can be categorised as constructs since they are used as essential building blocks within process models.

In order to ensure the quality of a new artefact, the design-oriented approaches inherently consist of two iterative steps: build (i.e. construction of the artefact in a transparent and traceable way) and evaluate (i.e. activities to prove innovativeness and ability to solve the addressed problem). In order to demonstrate the functional capability, effectiveness and efficiency of the developed problem solution, analytical (e.g. architecture analysis), descriptive (e.g. scenarios), experimental (e.g. simulation), or observational (e.g. case study) evaluation methods can be applied (Hevner et al., 2004). As it was our aim to test the GAs (which are considered the artefacts of this research) in a ‘real world’ setting and it is difficult to ‘simulate’ a health care environment, we chose the case study method to evaluate our findings (Kitchenham et al., 1995; Smith, 1990). In doing so, a project with five hospitals was started aiming at the depiction of the logistic processes using the identified GAs. An extract of the findings is presented in the subsequent sections. For illustration purpose, the Business Process Modeling Notation (BPMN) is used. However, other modelling notations such as Event-Driven Process Chains, Object Process Methodology, or flow charts can be applied since the proposed findings are notation independent to a large extent.

3 IDENTIFICATION OF GENERIC ACTIVITIES

3.1 Configurative and Compositional Mechanisms of Reference Modelling

For purposes of reducing heterogeneity of process descriptions by means of standardisation and improving syntactic, semantic as well as pragmatic model quality (Lindland et al., 1994), the concept of reference modelling provides generic mechanisms for reuse and adaptation based on predefined models or model components (Thomas, 2005). Mechanisms of adaptation can be categorized into configurative (Gottschalk et al., 2007) and compositional concepts (vom Brocke, 2007). Examples are the configuration of business processes within standard software applications (Dreiling et al., 2005) or the composition of independent software services to complete business workflows (Szyperski et al., 2002). The configurative top-down mechanism requires extensive knowledge about feasible variants potentially occurring within the process and the respective conditions for those variants which have to
be formalised to configuration parameters. Although the adaptation effort is considered low, the construction effort – especially in complex domains – is comparatively high (vom Brocke, 2007). As allowed variants have to be known in the development phase of the reference model, flexibility in creating new (innovative) process solutions during adaptation phase is as limited as the applicability of the predefined models for differing situations which are not supported by the reference model and its configuration parameters. Thus, the configurative mechanism is suitable for highly standardised and well structured processes.

In order to model processes in a more complex, distributed environment, compositional mechanisms of adaptation are considered more suitable (Reichert & Dadam, 1998; Baacke et al., 2007). Especially in information processing domains, process structures are very heterogeneous but basic activities (i.e. model components) recur regularly. Bottom-up composition of modular activities promises moderate effort for adaptation. Although composition supports restricting mechanisms (e.g. by using combination rules) to ensure model consistency and plausibility, the potential for process innovation and flexibility to depict both the individual as-is and the potential to-be situation are high. As a consequence, composition of standardised GAs is considered an appropriate adaptation mechanism for distributed modelling of cooperative processes.

3.2 Applicability of Public Sector Modelling in the Health Care Domain

As the component-based approach used for this contribution originates from the public sector (PICTURE Consortium, 2006), the question arose if it is applicable for business process modelling in the health care domain as well. Although processes in health care organisations are considered rather complex and correct outcome is crucial, there are many similarities in the service production of public and health care services. Example characteristics are

- Division of labour through specialisation and function-oriented organisation
- Cooperative service production involving several specialists and resulting in a number of organisational interfaces
- High autonomy of cooperating units
- Decision-intensive processes (and variable results) cause structural complexity
- Occurrence of similar activities in different processes
- Little experience in process modelling and heterogeneity of existing documentations.

At least for workflows of information exchange and processing as well as for administrative processes the appropriateness of compositional (compared to configurative) process modelling can be derived. Due to the specificity of medical treatment and extensive analogies to administrational processes, this contribution excludes processes of diagnostics and therapy and, instead, focuses on modelling more administrative processes of enterprise management, patient management, care administration, communication, supporting services and business support (Helfert et al., 2005) which predominantly focus on information management as well as procurement and provision of medical commodities for actual treatments. Such surrounding processes (internal and external) in general aim at providing a required quantity of information or products (in the required conditions and in time) for appropriate cost to the right customer (Plowman, 1964).
## 3.3 Decomposition of Modelling Constructs

As already mentioned, the component-oriented modelling approach is based on standardised GAs. Depending on the used notation and the understanding of the modeller, tasks are not only represented by activities or functions but also by organisational units, documents, events, etc. In order to standardise process models and to ensure compatibility to other process models and other model types, a precise differentiation of *Generic Activity* and respective *Processed Object* (PO) is demanded (Baacke et al., 2007). This differentiation is visualised in Figure 1.

![Figure 1. Differentiation of process tasks into GAs and POs.](image)

Although for this contribution medical treatment processes and their specific activities have not been analysed, the conceptual differentiation between GAs and POs supports consistent access and processing of information also for medical purposes (cp. Figure 2).

![Figure 2. Surrounding vs. medical treatment processes.](image)

Depending on the notation used for modelling, further entity types can be connected, e.g. responsible organisational units or supporting information systems (Baacke et al., 2008).
3.4 Procedure for Identification of Generic Activities

The procedure for identification of GAs conforms to the iterative build and evaluate phases of design research (cp. Section 2). The main steps as well as expected results are specified in Table 1.

<table>
<thead>
<tr>
<th>Design activities</th>
<th>Phases of design research</th>
<th>Expected results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Identification of GAs</td>
<td>Build</td>
<td>Initial GA catalogue</td>
</tr>
<tr>
<td>(2) Initial evaluation</td>
<td>Evaluate</td>
<td>User feedback</td>
</tr>
<tr>
<td>(3) Revision</td>
<td>Build</td>
<td>Example process models</td>
</tr>
<tr>
<td>(4) Final evaluation</td>
<td>Evaluate</td>
<td>Example process models</td>
</tr>
</tbody>
</table>

*Table 1. Design activities and expected results of the methodology.*

Due to the similar characteristics of processes in public administrations and support processes in health care organisations, not all GAs had to be newly identified. Instead, already developed GAs (Baacke et al., 2008) have been reused to some extent. However, the procedure for identification has been applied without modifications (cp. Figure 3).

*Figure 3. Procedure for identification of GAs.*

Analysing existing process models, an activity, function, task, or corresponding element was selected in the first step in order to identify a new GA candidate. In doing so, the core activity had to be extracted (generalisation). As not all selectable components necessarily need to be activities (cp. discussion of heterogeneity of models in Section 3.1), they had to be (re-) phrased as a verb (e.g. “receive” instead of “reception”). This activity then became a potential GA candidate. In the third step, the catalogue of existing GA candidates – which due to the reused GAs was not empty – has been checked whether the new candidate or a synonym was already contained. If a synonym was contained, the new GA candidate has been added as an alternative. Subsequently, the procedure iterated until no new activities were identified respectively the GA catalogue was considered to be complete. As the
GA catalogue was not empty at the beginning, the case of synonym occurrence has frequently been experienced.

3.5 Resulting List of Generic Activities

Result of these design activities is the catalogue of GA candidates (including synonyms) which had to be evaluated. Evaluation was especially important as existing GAs from the public sector domain had been adopted without further reflections. Evaluation should identify those GAs and synonyms which are considered inappropriate or irrelevant for health care processes. In total, the catalogue consists of 19 GAs which are listed and categorised in Table 2. Thereby, 17 GAs from the public sector could have been reused with minor modifications (esp. regarding synonyms). Only two GAs had to be newly added to the catalogue which confirms the theoretically derived similarities of service delivery processes in public administrations and non-medical processes in health care organisation (cp. Section 3.2).

<table>
<thead>
<tr>
<th>Categories</th>
<th>Generic Activities (GAs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transfer/Communicate</td>
<td>Send/Hand over/Deliver, Receive/Accept, Demand/Follow up</td>
</tr>
<tr>
<td>Analysis</td>
<td>Retrieve/Enquire, Check/Verify, Examine/Analyse</td>
</tr>
<tr>
<td>Production</td>
<td>Produce/Create, Change/Update/Complete</td>
</tr>
<tr>
<td>Transformation</td>
<td>Capture/Enter/Fill in, Print, Scan, Copy</td>
</tr>
<tr>
<td>Administration</td>
<td>Archive, Dispose/Delete, Start/Open</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>Pay, Sign, Prepare/Setting up, Store/Relocate</td>
</tr>
</tbody>
</table>

Table 2. Categories and related GAs.

Such categorisation does not only support users in quickly finding and selecting appropriate GAs but also enables implementation of mechanisms to ensure consistency and plausibility of the resulting models by double-checking syntactic and semantic rules (e.g. whether the GA “send” is connected with a corresponding “receive” or not). In addition, each GA is further specified defining its meaning, category, evaluation status, etc.

Although it was not in the focus of this investigation, typical POs have also been identified using the same, slightly modified procedure. Instead of extracting activities and re-phasing as a verb, components have been analysed with regards to contained objects such as documents, information or materials (e.g. medical products). Respective objects have been extracted and added to a PO catalogue. Some examples of the identified POs are drugs, internal requisitions and external purchase orders, invoices, and delivery notes.

4 CASE STUDY

In order to exemplify and evaluate the theoretical results presented above, and to practically prove the applicability of the concept in the health care domain, a case is presented in this section. Thereby, special emphasis is given to the illustration of the expressive power of the identified components. However, to validate our proposal not only the semantic, syntactic and pragmatic aspects are important (Lindland et al., 1994) but also the domain appropriateness since the model components are intended to capture the knowledge of the problem domain for the purpose of communication and understanding among the stakeholders (Siau & Rossi, 1998; Mylopoulos, 1992). Thus, Krogstie et al. (Krogstie et al., 1995) propose to subsume this as explicit evaluation criterion called ‘knowledge quality’.
4.1 What is the Case?

Hospital pharmacy is an interesting arena to bring many functions and relationships together for study. On the one hand, the hospital pharmacy acts as an interface between the many external stakeholders involved in supply of drugs and the particular hospital. The process documentation therefore has to comply with many, sometimes divergent, needs. On the other hand, it serves as a focal point for many internal departments (medical and non-medical). For example, the pharmacy is not only responsible for drugs supply but also offers consultancy services to clinics, prepares medical devices for the laboratory, provides additional information for the patient administration and financial department etc. Having this gatekeeper role, the hospital pharmacy is supposed to be an important opinion leader when logistic processes have to be re-engineered.

For the case study expert interviews represent the essential sources of information, and are the primary means for evaluating the utility of our approach. Semi-structured interviews were conducted with three hospital pharmacists and two purchasing managers who have at least 10 years practical knowledge in medical as well as in administrative needs. Main characteristics of the involved hospitals are provided in Table 3.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Units of Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hospital A</td>
</tr>
<tr>
<td># Beds</td>
<td>400</td>
</tr>
<tr>
<td># Inpatients/year</td>
<td>9,500</td>
</tr>
<tr>
<td># Outpatients/year</td>
<td>46,000</td>
</tr>
<tr>
<td>Automation level</td>
<td>low</td>
</tr>
</tbody>
</table>

Table 3. Characteristics of the surveyed hospitals.

It is important to note that we did not involve explicit control or manipulation of any variables of the subject under study. However, to gain that deeper understanding we acknowledge our own subjectivity as part of research.

4.2 Application of the Components in the Context of a Hospital Pharmacy

Hospital pharmacies have a permanent obligation for service delivery, not only under usual circumstances, but also in case of a crisis. Therefore only a marginal fault tolerance in the procurement processes is admissible. Sophisticated mechanisms to manage the entire supply chain are needed. However, in health care this is more aspiration than reality (Burns, 2002). To understand the current drawbacks it is crucial to have an overview of the present situation. A key process is the general order routine for purchasing regular drugs (cp. Figure 4). We used this process to evaluate our proposal against the ‘knowledge quality’ criterion (cp. introduction of Section 4).
As described in Section 3.3, essential components for modelling processes are GAs and POs. When using BPMN as standard modelling language, the GAs can be used to determine the activity (italic text in rounded-corner rectangle) to be done within a task. The POs (plain text in rounded-corner rectangle) define the object an activity is operating on. However, other BPMN standard elements such as organisational responsibilities (swim lanes), events (circles) and gateways (diamond shapes) can be used as well. For instance, the presented process model starts with an event called ‘demand for drugs’. This event triggers a task within the ward containing the GA ‘search’ and PO ‘drugs’ which on its part causes a decision. The interface to the medical treatment process is represented by a black circle labelled ‘start medical treatment’.

### 4.3 Lessons Learned

Although not accustomed to think ‘in boxes and arrows’, the hospital pharmacists were able to understand and interpret the process descriptions. Furthermore, with a brief training they were able to model even complex processes by their own (simplicity of modelling). After a short period of learning how to model with predefined GAs, the time to depict a process or sub-process was significantly reduced even for inexperienced partners. In addition, comparability of processes has been improved: Same processes were modelled for different hospital pharmacies. Although there were considerable differences in the realisation of the processes, the flow of activities was still comparable since a standardised terminology was used. The terminological standardisation based on predefined GAs not only improves comparability of process models but provides an important foundation:

- to diffuse modelling in health care organisations by addressing also inexperienced users,
- to harmonise decentralised modelling activities,
- to enhance the efficiency of modelling and
- to enable extensive, model-spanning analyses.
Although the case study with hospital pharmacists could not evaluate all of these potentials, as a first step the applicability of the GA-based modelling concept by inexperienced users has been proven successfully. However, the suggested standardisation of modelling terminology still leaves room for misinterpretation. In order to further support the modeller, a system of rules should be incorporated which limit reasonable GA combinations or even suggests frequently used GAs (e.g. based on GA patterns). Such mechanisms will improve model quality by extending consistency and plausibility checks.

5 CONCLUSION AND FURTHER RESEARCH

In this contribution, an approach of component-based modelling developed for public sector organisations was introduced and adopted in the health care domain. Summarising, this adoption is considered quite successful by hospital pharmacists as modelling with GAs considerably improved standardisation and comprehensibility of process models. After brief introductions even inexperienced case study partners were able to conceptualise their process knowledge re-using predefined GAs. In addition, the link-up from GAs to POs supports consistency and model-spanning analyses.

However, some limitations became visible. Although only few GAs had to be newly introduced, the list of GAs is not yet representative as it bases on a limited number of processes used for identification. Thus, further evaluation needs to be conducted in order to complete the specification. Furthermore, the occurrence of specific GA combinations should be analysed in order to identify typical process patterns. Such rules and patterns will significantly improve effectiveness, efficiency and quality of process modelling. In addition, more POs should be identified and categorised in order to further standardise modelling. As competition is not that intensive and externalisation of internal knowledge is not considered that crucial, it is furthermore planned to record and provide process description (best practices) in a common repository.

In order to address divergent processes perceptions of involved people (e.g. medical specialists or administrative staff) and to harmonise their model-based conceptualisation, the composition of GAs should provide alternative terminologies (e.g. synonyms), a set of individual rules to restrict combinations of GAs for plausibility and consistency purposes as well as transformations into user-oriented model views. Those semantic features extend classical syntactical checks significantly. They can be implemented by means of ontologies (Uschold & Jasper, 2001). If GAs and POs are depicted in an ontology they can be re-used (instantiation and composition) by mapping the ontology concepts and properties to the meta model elements of a suitable notation, such as UML Class Diagram or BPMN (Saeki & Kaiya, 2006). The ontology-based standardisation of model content not only improves transparency and model quality but especially provides the foundation for other powerful mechanisms such as semantic analyses, e.g. regarding concrete weakness patterns (Baacke et al., 2008). Weakness patterns are combinations of model components that represent a deficiency, such as a media break or an inefficient communication channel. On that basis, analyses can be automated to some extent. As automated analyses are not limited to single processes, they can easily be applied to the whole process landscape. Hence, further research will aim at the development of standardised analysis patterns specialised on issues of the health care domain.

In order to realise and prove these potentials, the component-oriented and ontology-based modelling approach is currently being implemented in a software tool which will be used for extensive evaluations and further developments (PICTURE Consortium, 2006). As the case study presented in this paper just prove applicability in the context of hospital pharmacies, further – more extensive – evaluations have to address syntactic and semantic interoperability, comparability, processability and quality of resulting models as well as efficiency of process modelling. Last, it has to be investigated if this approach can be adopted to other industries, such as finance or insurance, as well.
References


AN AGENT-BASED APPROACH TO IMPROVING RESOURCE ALLOCATION IN THE DUTCH YOUTH HEALTH CARE SECTOR

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0621.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Agent-based model, Simulation, Supply chain management, Health information systems / medical record systems / care delivery /</td>
</tr>
</tbody>
</table>
AN AGENT-BASED APPROACH TO IMPROVING RESOURCE ALLOCATION IN THE DUTCH YOUTH HEALTH CARE SECTOR

Giesen, Erik, INITI8, Van Nelleweg 1, Rotterdam, NL, giesen@initi8.nl
Ketter, Wolfgang, Department of Decision and Information Sciences, Rotterdam School of Management, Erasmus University, Rotterdam, NL, wketter@rsm.nl
Zuidwijk, Rob, Department of Decision and Information Sciences, Rotterdam School of Management, Erasmus University, Rotterdam, NL, rzuidwijk@rsm.nl

Abstract

We show how agent-based simulation is used for analyzing different queuing strategies in the youth health care sector. The simulation model represents an authentic business case and is parameterized with actual market data. We discuss the differences between four queuing strategies which are based on push/pull allocation and centralized/decentralized queuing strategies. The model incorporates, among others, a withdrawal and return mechanism, a non-stationary Poisson arrival process, and a preference algorithm to include a care provider’s case preference. The investigated system accommodates extensive waiting lines which are currently solely judged on their length. We have identified that performance measurement in youth health care should not be focused on queue lengths alone, but should include a case difficulty parameter as well. The simulation results, together with contextual data obtained from stakeholder interviews, indicate that a push strategy with a centralized queue suits the sector best. Most related research in health care focuses on queuing theory which fails to address the complexity of the case. Our simulation approach incorporates additional complexities of the case at hand which turn out to be relevant for the queuing strategy decision. We validate the model and strategies by comparison with real market data and field expert discussions.

Keywords: Agent-based simulation, resource allocation, youth health care, queuing strategies, scenario analysis
1 INTRODUCTION

The Dutch youth health care sector aims to provide care to children on demand. Children can enter the system at the institution for youth health care at their own initiative and obtain an indication for professional help. A formal indication includes a diagnosis and entitles the child to receive care at a care provider of its own preference. The system is publicly funded and enforced by the law on youth health care since 2003. The national system is subdivided in multiple regional systems which cover provinces or urbanized regions. These local systems implement their own structures, including local offices of youth health institutions, care provider agreements, and financial systems (Figure 1).

Over the last few years, the Dutch youth care sector has faced long waiting lists and long waiting times, a problem that has received a lot of media attention. The government has provided funding on several occasions, which did not result in a permanent solution. Hiring interim managers to reorganize internal operations at for instance a local office of the institution for youth health care turned out to be unsuccessful as well. The approach taken in this research starts with the observation that the real problem is not solely the waiting list length or waiting time as such. We claim that it is rather a problem of the current strategy of allocating cases to care providers in general. In our opinion, methods that solely address the symptom of long waiting lists will be ineffective in the long run. We shall consider solution directions that not only focuses on the handling of contemporary waiting lists, but that may require structural changes in the system.

We elaborate on such structural changes by presenting an overview of multiple allocation strategies, based on a combination of push/pull strategies and centralized/decentralized queuing strategies. The push and pull strategies define the party which ultimately makes the actual allocation decision. This could either be the institution for youth health care that pushes specific cases towards care providers, or the care provider that pulls the specific demand for care towards its resources. Centralized and decentralized queuing strategies define the moment at which the actual allocation will take place. A centralized strategy postpones allocation of children until the moment they can instantly get care; in other words, at the moment a care position has become available at a care provider and can be dedicated to the specific case. This strategy results in waiting lines at the institution for youth health care. A decentralized strategy allocates cases immediately as much as possible when they arrive in the system; this will result in waiting lines both at the institution for youth health care and at the care providers. Allocation decision problems, as presented by the youth health care case, suit very well a multi-agent simulation approach, as the allocation decisions depend on communication between the
different parties in the system. Furthermore, institutions and persons have their own objectives and the coordination thereof needs to be addressed explicitly. The actual clientflow through the system is the result of a negotiation process between several parties in the supply chain. Indeed, a case allocation procedure requires input from other parties in the sector on which the final decision can be based. Since a multi-agent simulation is built of individual agents that pursue a specific personal goal, it is fairly easy to configure and study alternative allocation strategies in a simulation environment.

Our objective is to provide insights into the allocation problem and to advice on possible improvements by discussing and evaluating a variety of allocation strategies. The analyses are supported by an agent-based simulation model. The model simulates the allocation process in the youth care sector as well as the directly coupled processes of case indication and the actual treatment. The parameters of the model are loaded with stochastic distributions which are based on actual data taken from one province in the youth care sector. While it is difficult to play around with the actual system itself, the simulation model provides a rich but risk free environment for testing varying resource allocation strategies. A simulated environment has the advantage that multiple decision algorithms and coordination structures can be used for experiments. This approach enables analyses on long term effects of certain strategies while maintaining focus on inter-agent interaction. Such studies would be hard, if not impossible, to perform in the real world system, because one cannot make structural changes in the real world system just to study a certain long term effect. Our approach contributes to the research in information systems and agent-based simulation, since it proves the usability of an agent-based approach in a real world environment by not only matching the current decision making process but also by studying varying alternatives. The model is loaded with an extensive amount of stochastic distributions based on actual market data and successfully matches the performance of the real world system. This research contributes to the research in resource allocation in health care by providing a currently unused approach to counter queuing related issues. Simulation of the resource allocation process helps to understand and test long term effects of varying allocation strategies and coordination decisions. We contribute to research in information systems by improving the human decision-making process. Our study on the different strategies on the youth health care system decreases the information overload which increases the rate of fair child allocations. This will improve socially responsible welfare decision-making.

2 LITERATURE REVIEW

Waiting lines in the health care sector have received little attention in the scientific literature, although hospitals and general health care face similar problems with waiting lines. A common approach taken by governments to tackle these problems is the injection of capital which is used to increase capacity. This provides a short term solution, as available capacity and queue lengths reach a new equilibrium after a short while (Hurst & Siciliani 2003; Postl 2006). Saulnier, Short & Gruenwaldt (2004) identified five popular approaches to decrease waiting times: monitoring of procedures, using priority scoring tools, setting waiting time targets, using an external advisory body, and registering online. However, Rachlis (2005) argues that such methods do not work by themselves; better coordination and flow control should increase performance at the public sector. Waiting lines in health care feature withdrawals when clients have to wait for an extended amount of time. Several studies have shown that the amount of time that a client is willing to wait for care is related to the urgency of the problem (Goodacre & Webster 2003, Goldman et al. 2005). More urgent problems are difficult to treat elsewhere, while they genuinely require attention. These cases will therefore accept longer waiting times. The converse holds for less urgent problems. Most literature on waiting line management in health care are based on the mathematical approach described by queuing theory (Torgerson & Mcnntosh 2006). These studies focus mainly on utilization of resources and calculations of the minimum required amount of treatment positions while maintaining a high service rate (Gorunescu, McClean & Millard 2002a & 2002b; McManus et al. 2004). These studies show how queuing theory struggles with phenomena like seasonal effects and withdrawals. These issues are also identified by
Brown et al. (2003) who argue that traditional queuing theory has a series of shortcomings like, among other things, the absence of customer withdrawal behaviour, time-dependent parameters or customer heterogeneity. These three characteristics are all notably present in a health care system. Mandelbaum & Shimkin (2000) tried to construct a model for withdrawal. They acknowledged that a lot of work needs to be done to achieve practical usability of queuing theory.

Only a limited amount of literature studied the waiting line problem by means of event-based simulation (Ridge et al. 1998, Bagust, Place & Posnett 1999). While these studies include stochastic processes and basic withdrawal schemes, they still solely focus on utilization issues and capacity planning. To the knowledge of the authors, agent-based approaches have not yet been used in this context. Event-based simulation is able to include an extensive level of complexity in the model like withdrawal & return behaviour, seasonal arrivals and client heterogeneity. An agent-based model in addition is also suited to quickly implement and experiment with alternative strategies and coordination flows. These experiments are difficult, if not impossible, to implement with queuing theory. Research on waiting lines and times by means of simulation is rarely used in the health care sector in general. While other sectors with comparable cross organisational networking structures like container ports were studied with agent-based approaches (Moonen et al 2005), an agent based approach in health care is even unique in its kind. This study will therefore provide a first example of the usability of agent-based methods for improving performance in the health care sector. It will contribute to scientific research in studying an agent based simulation approach on a genuine and new business case in practice with a significant amount of market data in the youth health care sector. Unfortunately, research is not able to provide a clear definition of what an agent is. While most researchers agree on certain aspects of an agent, it seems not easy to integrate all views in one grand perspective (Lang et al 2008). The four and most appreciated aspects are autonomy, social ability, reactivity, and pro-activeness, this set has become known as the weak notion of agency (Wooldridge & Jennings 1995). The authors state that an agent is a hardware, or software based computer system, or both, displaying the properties of autonomy, social adeptness, reactivity, and pro-activity.

3 THE SIMULATION MODEL

The agent-based model extends the Distributed Simulation Object Library (DSOL) simulation environment (Jacobs et al 2002) and was run on a conventional desktop computer with settings as shown in Table 1. This environment provides the discrete event simulation engine, the theoretical distributions and a toolkit to design a simulation model as a set of loosely coupled objects. Because of the latter, the suite is ideal for implementing an agent-based model since it enables the design of agent who are physically disconnected from each other. While the simulation environment provides all the elementary building blocks of a simulation, the model itself only concerns the agent specific parts. Other research studied agent-based simulation architectures and the required components for building an agent-based model (Ketter et al 2009, Collins et al 2002).

<table>
<thead>
<tr>
<th>System</th>
<th></th>
<th>Simulation model</th>
</tr>
</thead>
<tbody>
<tr>
<td>System</td>
<td>Dell Vostro 200</td>
<td>Java version JDK 1.6.0.11</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel Pentium E2140</td>
<td>Java VM arguments -jar -Xms512m -Xmx1024m -server</td>
</tr>
<tr>
<td>RAM</td>
<td>2GB DDR2 (PC2-5300)</td>
<td>Format Access database</td>
</tr>
<tr>
<td>OS</td>
<td>Windows Vista Ultimate 32bit</td>
<td>Data per replication ± 70-77MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CPU time per replication ± 10 minutes</td>
</tr>
</tbody>
</table>

*Table 1 Technical information of system used for analysis.*

1 The total amount of data used for the study was about 20 GB from 256 replication runs.
2 The model runs on one processor core, therefore the model can be ran on both processor cores simultaneously without drastic influence on the average processing time. The given value is based on the average processing time while running the model twice in parallel on the same machine.
There are several types of information identified in the model. First, there are parameters which define fixed values like agent names, the theoretical distributions and the geographical home location of an agent. These parameters are mined from real world health care data. Second, there are dynamic data stores which hold process information upon which an agent can make decisions. This group of information can be divided in two groups; the transactional data store and the decision data store. The transactional data store holds records of the overall process of an agent. For instance, the agents of the institution for youth health care maintain an internal database of care lines which holds all relevant information of a case. The care provider agent hold a similar database, however holding only information which is relevant and accessible for these agent. The transactional data store in such is factual information which represents historical record keeping. The decision data stores, on the other hand, hold short term information which is relevant in executing the allocation strategy on a certain moment in time. The appraisal of this data decays over time since its trustworthiness decreases as time passes by. For instance, the institution for youth health care who decides on the best location for allocating a case based on known data and estimates of dated data.

The core model is constructed around a strict path which all cases in the system will follow. This flow is used in all four strategies. The white box indicates that the used allocation strategy might differ, while the remaining parts stay the same. Each child or case will be created at the top-left. The model includes mechanisms for the allocation process, the waiting phase and the treatment itself. Further it includes withdrawals during the waiting phase and returns after treatment or withdrawal. A core concept in this model is the use of a case selection process for treatment at the care providers. An overview of this process is shown in Figure 2. When a care provider selects the next case for treatment, he will judge the cases in his queue on certain characteristics for fitting with the available treatment location. This is modelled as an fit with the age of the child which results is a potential list of cases for treatment. While the order of these cases should be on ‘first come, first serve’ basis, this rule is violated by the care providers when possible because they prefer easier cases. More easy cases means more throughput which increases profit for the care providers. The political influenced decision algorithm, which is shown in Figure 2, will order the potential cases by preference of the care provider. The political influenced decisions are responsible for the major part of the differences in performance between the allocation strategies. Without this algorithm all strategies would perform roughly similar because the selection procedure would maintain a strict ‘first come, first serve’ order.

![Figure 2](image-url)  
*Figure 2* Overview of the case selection process for treatment at the care providers. The dashed boxes represent factors which set restrictions on the set of potential cases The single lined boxes represent the set of cases which can potentially be selected while de last (double lined) box represents the final case selected for treatment.

Equation 1 is responsible for the political influenced ordering of the cases which has been developed by use of a balanced scorecard technique (Kaplan & Norton, 2005) based on interviews with field experts and the evaluation of real world data. The matter at hand is intangible, complex and hard to express in words; often these issues live in one’s mind rather than being written down in some kind of procedures, therefore the interviews were of an unstructured kind in order to manoeuvre the discussion into the most fruitful directions during the interview. Among the experts are: a youth health care consultant with a high level of experience in the sector, a case manager at the institution of youth health care with operational experience, a financial director at a care provider with operational experience and some strategic experience, a director at a care provider with strategic experience and some operational experience. The results of these interviews are translated into a set of constraints...
which the equation should obey. These constraints are parameterized into the resulting equation which consists out of two parts; the influence of a case’s waiting time (left) and the influence of the expected treatment time (right). Waiting time is estimated to influence the decision mainly in the first weeks; a case which is waiting just a few days cannot precede a case which is already waiting a long time because the care provider will not be able to justify such a decision to the institution for youth healthcare. This effect is estimated to be decelerating decreasing with the turning point at the threshold value\(^3\). This simulates the effect of a rapid increase in preference in the first weeks of waiting while it flattens out afterwards. The second part represents the expected treatment time by a care provider. Cases with below average treatment times are preferred in any case, however the higher the treatment time, the lesser the preference which is expected to be exponentially increasing. The threshold value is related to the average treatment time in order to set the estimated steepness of the preference line. The graph in Figure 3 shows an example of four potential cases after making a subset for age characteristics of the open treatment position. The cases differ on waiting time and expected treatment time. It can be seen that the waiting time order (when solely looking at the waiting value in decreasing order; B-C-E-G) differs from the preference order (looking at the acceptance value in decreasing order; C-E-B-G). When solely looking at the waiting time, child B should be selected, however the preference function will prefer child C.

\[
\text{acceptance factor} = \frac{\text{threshold\_waiting}}{\text{waiting\_time} + 1} + \frac{\text{treatment\_time}^2}{2 \cdot \text{threshold\_treatment}^2}
\]

\(1\)

![Figure 3](image.jpg)  

**Figure 3** Indifference curves of the political influenced decision algorithm for specific acceptance factors including a example subset of children ready for allocation.

This core model, including the political influenced decision algorithm, is the same for all four strategies. The difference will be in the process which performs the actual allocation. While the current allocation strategy in the youth care sector is based on pushing cases towards the care providers at the moment they receive an indication for care needs, it is considered that a push can be postponed until the moment that a care position becomes available. This centralizes the queue at the institution of youth healthcare which will enable the supply chain to more effectively match cases to care positions. Such an system would reduce the ability of care providers to make preference based selections since the institution of youth healthcare has a much better view and control on final case selection which increases fairness in the system. Care providers are likely better in estimating the actual treatment plan for a child than the institution of youth healthcare since they are the party

\(^3\) Threshold\_waiting is set to 32 days, threshold\_treatment is set to the average treatment time of the care provider in question. These settings are based on interviews with field experts and evaluation of real world data.
providing the actual care to the children. Therefore it would increase the efficiency of the system when the care provider would also decide which child should be treated next. This strategy changes the push system into a pull system in which the care providers select cases from a pool of children requiring treatment. Note that the moment of selection still enables the possibility of maintaining centralized or decentralized queues, therefore we will address the following four types of allocation strategies.

1. **Pushing cases to decentralized queues [Decentralized Pushing].** When a case is indicated at the institution for youth health care, it is immediately pushed to one of the care providers. The push is performed by the institution for youth health care. This strategy is currently implemented in the youth care sector. In this case the care providers maintain and control their own queues while the institution of youth health care only tracks the development of a case. The political influenced decision algorithm can significantly influence the order of treatment because this strategy positions all cases at the care providers. The allocation itself is performed by the institution of youth health care which makes the decision based on a estimated ‘shortest queue’ algorithm. Each institution agent maintains a list of known queue lengths to support this decision. The value of each item in the list however decays when time passes by. The algorithm estimates the current queue length at a care provider by judging the available information while it gets maximal three requests for an updated queue length per case. The agent therefore makes the final decision on partially uncertain information.

2. **Pushing cases from a centralized queue [Centralized Pushing].** When a case is indicated at the institution for youth health care, it is held in a centralized queue until a fitting treatment position becomes available. The institution for youth health care maintains and controls the central queue while the care providers have no queue at all. The care provider announces that treatment positions become available and the institution of youth health care pushes the next case for treatment. The political influenced decision algorithm at the care providers is blocked in this strategy since there are no cases to choose from except for the case pushed by the institution for youth health care.

3. **Pulling cases from a centralized queue [Centralized Pulling].** When a case is indicated at the institution for youth health care, it is held in a centralized queue until it is pulled by a care provider which offers an open treatment position. The institution for youth health care publishes the waiting list on a bulletin board for evaluation by the care providers. The care providers do not maintain their own queues. The institution for youth health care monitors selection behaviour by maintaining ‘first come, first serve’ order on comparable cases. Care providers anticipate on this rule by applying the political influenced decision algorithm to choose between multiple fitting, but incomparable cases on the bulletin board to select the following case.

4. **Pulling cases to decentralized queues [Decentralized Pulling].** When a case is indicated at the institution for youth health care, it is published on the bulletin board until it is selected by a care provider which promises to offer a treatment position in the future when it becomes available. Both the institution for youth health care and the care providers maintain queues in this strategy. A care provider prefers easy cases but must also select the comparable preceding cases if it wants the preferred case. Postponing a preferred selection however increases the chance that another care provider selects the case while it also increases the chance of ending up with an empty queue. An empty queue means that a care provider cannot find a fitting case for a treatment position which is considered worse than treating a unprofitable case. The care provider will evaluate the trade-off between the inclusion of less easy cases in the allocation and the chance on running his own queues empty. This selection behaviour by care providers is simulated by the function from Equation 2. The equation shows the algorithm used to judge cases on the bulletin board. The equation consists out of three parts; influence of current queue length (left), influence of estimated treatment time (middle) and the influence of the interaction between these two factors (right). The equation is based on interviews with field experts\(^4\) which prescribed a set of constraints; When a care provider faces a low queue length for a given treatment type, he is likely more easily willing

\(^4\)Threshold\_queue, threshold\_treatment and threshold\_steep are based on interviews with field experts in order to create the desired effect of decision making for this allocation strategy.
to accept new cases in order to ensure full utilization of resources (i.e. bedrooms). The longer his queue, the less willing he is to accept new cases which will decrease exponentially as the queue increases. The expected treatment time of a case will also influence the decision because the care provider will try to avoid the most difficult and therefore unprofitable cases. Ultimately these two factors interact with each other which weaknesses the willingness to accept cases even further if at least one of the two dimensions scores rather high.

\[
\text{acceptance factor} = \frac{\text{queue length}^2}{\text{threshold queue}} + \frac{\text{treatment time}^2}{(2 \cdot \text{threshold treatment})^2} + \sqrt{\frac{\text{queue length} \cdot \text{treatment time}}{\text{threshold steep}}}
\]  

(2)

## 4 MODEL ANALYSES

The model is parameterized, verified and validated with real world data during three months of full-time analyses. The model is initiated with 7 youth care institutions and 8 care providers which have a capacity and a geographical location in the studied region. These will be unchanged during the entire simulation. The children are generated during the time of the simulation by a non-stationary Poisson distribution to include a seasonal influenced arrival effect. The age of the arriving children is included in the simulation to include the non-uniform division of age over the generated children. The age of a child is important for the age related aspect of the treatment positions. Further a distinction is made on the crises level of a case. Some cases are remarked crises which means that these cases are allocated and get treatment at once. These cases bypass the allocation strategy but do influence the usage of capacity in the model. Note that a ‘crisis’ denotes a case of extreme urgency and that its level of difficulty can be of any kind. Each generated child will be indicated with a varying amount of care needs. These needs can be indicated simultaneously at the first indication or re-indicated after a withdrawal or a successful treatment. This also involves the analyses for withdrawal chances during the waiting phase and return chances when withdrawing or ending care. A return further involves a return interval since the child will not return immediately but after a varying amount of time. Since a child cannot be geographically allocated at random due to practical distance limitation, the analyses include a set of chances for maximal allocation distance from the view of the child’s home location. A case is provided with a identifier which represents an attribute for a case’s difficulty. The difficulty of a case cannot be analysed but is considered to be uniform distributed for the purpose of this research. The treatment time distributions differ among the care providers, the treatment time is therefore coupled to the difficulty identifier and the care provider. Table 2 lists the parameters and types of distributions as they are used in the model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Valid</th>
<th>Type of distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacities</td>
<td>+</td>
<td>Absolute value</td>
</tr>
<tr>
<td>Arrival distribution</td>
<td>++</td>
<td>Non-stationary Poisson distribution</td>
</tr>
<tr>
<td>Age distribution</td>
<td>0</td>
<td>Empirical distribution</td>
</tr>
<tr>
<td>Crises distribution</td>
<td>++</td>
<td>Absolute chance value (%)</td>
</tr>
<tr>
<td>Parallel tracks</td>
<td>++</td>
<td>Absolute chance value (%)</td>
</tr>
<tr>
<td>Repeating tracks</td>
<td>0</td>
<td>Absolute chance value (%)</td>
</tr>
<tr>
<td>Difficulty</td>
<td>++</td>
<td>Uniform distribution</td>
</tr>
<tr>
<td>Geographical arrival distribution</td>
<td>++</td>
<td>Uniform distribution within the borders of the region</td>
</tr>
<tr>
<td>Geographical range limitations</td>
<td>++</td>
<td>Absolute chance value (%)</td>
</tr>
<tr>
<td>Care duration</td>
<td>0</td>
<td>Empirical distribution</td>
</tr>
<tr>
<td>Withdrawal ratios</td>
<td>++</td>
<td>Calculated chance value (%)</td>
</tr>
<tr>
<td>Return rate</td>
<td>++</td>
<td>Calculated chance value (%)</td>
</tr>
</tbody>
</table>

Table 2 Model parameters with validation and type of distribution remarks. The validity results range from bad to good, respectively: --, -, 0, +, ++.
The model is initiated as a non-terminating system since decisions and performance measures depend on long lasting developments. The model is pre-filled at start in a fully utilized state at the care providers while there are no waiting lists. This procedure will decrease the required warm-up length of the model. Warm-up time has been determined by the method of Welch (1981, 1983) which resulted in a warm-up time of 4 years simulation time. The replication length has been set to 20 years simulation time in total.

The verification of the model is split in two groups; first, the introduction of state-transition control and the implementation of numerous checks during the simulation which ensure a correct flow of cases through the system. Second, in-depth source review by others who didn’t participate in the design of the model verified the correct coding of the model. These two types of verification resulted in the acceptance of the model as technically correct.

The validation process is split in three groups; first, the results for the input parameters of the model are compared to the expected values which are based on the real world data which was used for the input analysis. The comparison of chance variables is performed by matching the average results with the expected value. The theoretical and empirical distributions are visually compared with the results. The results are split into multiple independent sections which enables the comparison with a specific certainty. This procedure is required to avoid biased judgement due to the stochastic nature of the parameters in question. The results of this comparison phase are listed in Table 2; the column ‘valid’ represents the result of the validation process of a particular parameter in the simulation model. Second, two of the most important but less understood parameters of the model are analysed for sensitivity. The determination of a case’s fit by comparing the age of the child to the preferred age of the treatment position depends on a certain range which is considered acceptable. The results showed low sensitivity on this setting. The other value for sensitivity analyses is the political influenced decision algorithm. The results showed moderate sensitivity on these values. The third validation group is the user validation by field experts (1 youth health care consultant with a high level of experience in the sector, 1 case manager at the institution of youth health care with operational experience, 1 financial director at a care provider with operational experience and some strategic experience, 1 director at a care provider with strategic experience and some operational experience). The results show that the model mimics expected behaviour accurately. The field experts recognized much of the real world system in the model’s output. For example, the arrival distribution including seasonal effects and the construction of treatment trajectories including crises, parallel and repeating cases were found realistic representations of reality. The model has been found valid on almost all aspects and is therefore assumed sufficiently accurate for this research.

5 RESULTS AND DISCUSSION

The strategies are judged on a set of 5 Key Performance Indicators (KPIs); 1) Ratio of (difficult) cases that are taken into care, 2) Ratio of (difficult) case withdrawals from the waiting lines, 3) Amount of (difficult) cases on the waiting lines, 4) Aggregated waiting times for the (difficult) cases and 5) Utilization of care positions at the care providers. The KPIs are developed by discussion with field experts, publications from the youth care sector and evaluation of real world data which indicate the importance of these measures. Interviews with field experts indicated that a major shortcoming of the current system is the neglecting of difficult cases. The law on youth care indicates that children in need of care should be able to receive it. However Andriessen and Besseling (2008) argue that there are many cases which receive help via the institution for youth health care are not genuine cases requiring professional help. The authors indicate that these cases shouldn’t enter the system because either the indication of a problem is falsely recognized or the problem is of such a low level that they are able to help themselves. The field experts join this conclusion. The KPIs therefore include a division in average performance and performance for the more difficult cases which are often neglected by the system.
The analyses of the KPIs are based on 10 replications per strategy to avoid biased judgment due to the stochastic nature of the model’s parameters. The replications are joined together in order to construct 95% confidence intervals for each KPI in each allocation strategy. Difficulty has been separated in 5 groups sorted from minimal to maximal difficulty. The two groups of the highest difficulty represent the performance on difficult cases. By means of Pareto analysis we illustrate the trade-off between efficiency and fairness in the system as summarized in Figure 4 which shows the average performance and performance on difficult cases in one overview. The spheres represent the performance scores of an allocation strategy. The horizontal axis illustrates the performance on the service level related KPIs while the vertical axis illustrates waiting line related KPIs. A lower value (towards the lower-left corner) represents better performance. The line in-between the spheres illustrate the continuum on which combined performance will be found depending on the relative importance of the performance types.

The graph shows that the average performance and performance for difficult cases differs significantly among the strategies. None of the strategies score similar performance. It also shows that the choice between a focus on average performance or difficult case performance greatly impacts the order of performance between the strategies as indicated by the rather long lines (except for centralized pulling). It further illustrates a trade-off between waiting line and service level aspects. It appears that decentralized pushing is performing best on waiting line aspects, whether it is on average or difficult case performance. It however loses on the service level aspect from both centralized pushing and pulling. The graph therefore helps to understand the effect of preferring one type of KPI above the other. In this analyses however, the importance of the KPIs is considered equal.

![Figure 4 Trade-off between waiting line and service level aspects.](image)

The relative importance of difficult case performance over average performance is difficult to quantify. There the following assumptions about its ratio explain the effect of a certain choice.

I. Average performance is important, difficult case performance not. This is the set of small spheres, which results in an order of strategies (best strategy first): 1,3,4,2
II. Average performance is equally important as difficult case performance. This is the spot in the middle of the lines. Order of strategies (ampersand means equal): 1,2&3,4.

III. Average performance is not important, difficult case performance is the important measure. This is the set of big spheres. Order of strategies: 2,3,1,4.

IV. Average performance should be respected, however difficult case performance is of a higher importance to avoid the neglecting of these difficult cases. This is the spot on the line which is found between the middle of the line and the bigger spheres. Order: 2,1&3,4.

These assumptions are not randomly chosen but rest on the way the system is judged by different people. Assumption I is related to the traditional way of judging the performance in the youth care system as it is still used currently. Assumption III, on the other hand, is related to the perception of field experts and researchers who claim that the difficult cases should be the primary focus of performance measurement since they are in the greatest need for care. Assumption II is the simplest method to share the importance of both these views and will likely be the result when discussions in the youth care sector don’t bring forward a preference for one or the other. Assumption IV is our point of view since it emphasizes the importance of the difficult case performance while it still respects the average performance. It is considered that a slightly worsened average performance isn’t bad as long as the difficult cases can profit significantly from it. The balance should be towards the point that these difficult cases cannot profit anymore without significantly decreasing the average performance. It is acknowledged that such an strategy would require a change in the financial system because the current system would underpay the care providers who are paid on an average case basis.

6 CONCLUSIONS AND FUTURE WORK

We have presented an approach for analyzing a number of resource allocation strategies in the youth health care sector while including an extensive set of constraints and behaviours from the real world system. The model successfully simulated many of the complex relations between the involved parties in the system. We demonstrated the ability of the model to incorporate different allocation strategies while maintaining an overall structure which deals with the common tasks outside the allocation procedure. We discussed the differences between the scenarios and their impact on system performance. The introduction of a case’s urgency into performance measurement leads us to the advice of the push from a centralized queue strategy for future resource allocation in the youth health care sector. The postponement of the actual allocation in this strategy ensures a higher level of fairness in treatment provision by the care providers because they cannot avoid the difficult cases anymore.

Our approach shows the usefulness of agent-based modelling in complex environments like the youth health care sector where much of the problem is related to coordination and communication between different parties. The research therefore contributes to research in information systems and agent-based simulation by not only showing its usability in such a setting but also showing the ability to study alternative strategies which couldn’t be studied otherwise with this level of complexity. We also contribute to research in information systems by advising a different allocation strategy. This strategy will increase social welfare by increasing the system’s fairness for provision of treatment.

We intend to study further alternative strategies. Within the four strategies there is plenty of room to introduce alternative concepts such as auctioning and case exchange between care providers. We further intend to introduce an implementation of a human preference model into the decision making process as well as human computer interaction. The agent-based architecture of the simulation model is ideally suited to implement such alternatives since it only involves changes in the respective agents. Currently, our evaluation process for the simulation results are based on a balanced scorecard technique. We are planning to define a multi-attribute utility function as a comparison and extension to the current approach. Such a function would allow the floating transition of the trade-offs between different desired properties of patients’ allocation.
References

RFID-ENABLED PROCESS CAPABILITIES AND ITS IMPACTS ON HEALTHCARE PROCESS PERFORMANCE: A MULTI-LEVEL ANALYSIS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0675.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Health information systems / medical record systems / care delivery /, Business Process Improvement, Business value of IT, IT innovation</td>
</tr>
</tbody>
</table>
RFID-ENABLED PROCESS CAPABILITIES AND ITS IMPACTS ON HEALTHCARE PROCESS PERFORMANCE: A MULTI-LEVEL ANALYSIS

Abstract

In recent years, hospitals have begun to invest in RFID systems to control costs, reduce errors, and improve quality of care. Despite the obvious benefits of RFID in healthcare settings, potential obstacles to effective deployment also exist. The purpose of this study is to systematically understand how hospitals can apply RFID to transform work practices and address cost, safety, and quality of care issues, most notably in inventory management. We leverage an interdisciplinary framework to explore adoption and use of RFID at multiple levels of analysis and adopt a multi-method approach to explore the research questions guiding this study. Our study is expected to contribute to a growing body of research related to the adoption and use of IT in healthcare settings and the enabling role of IT for innovating work practices and improving process performance.

Keywords: RFID, Healthcare IS, Multi-level Analysis, IT-enabled Process Capabilities, IT Adoption
1 INTRODUCTION

The adoption and use of innovative technological solutions for solving business problems have been arguably slow within the healthcare industry (Herzlinger, 2006). However, in recent years, burdened with rising healthcare costs and increased competition, hospitals have begun to invest in advanced IT systems such as RFID-based systems to control costs, reduce errors, and improve quality of care (Datamonitor Report, 2008). In fact, as referenced in Lee and Shin (2007), the global market for RFID systems and tags in hospital settings is expected to increase from $90 million in 2006 to $2.1 billion by 2016 (Harrop and Das, 2006). Nevertheless, despite the obvious benefits of deploying RFID technology, there are some apparent downsides, too. High costs, difficulty integrating with legacy systems and interoperability with life saving medical equipment, incompatible communication standards across major entities in a healthcare network, and entrenched work practices are all potential obstacles to effective adoption and use of RFID in healthcare settings (Smith, 2005).

The focus of this research project, therefore, is to systematically understand how to apply RFID as a transformative technology to address cost, safety, and quality problems in this ever important industry. To narrow our scope, this research focuses on the use of RFID within a bounded set of activities and work processes in the healthcare industry. Namely, we focus on inventory management processes that support the delivery of two types of healthcare services. These services differ on key dimensions related to the variety and value of products used during procedures, the range of procedures performed within the service areas, and the urgency with which services are often performed. We draw from prior research to deepen our understanding of how IT can be leveraged to improve inventory management and purchasing processes (Cohen and Lee, 1998) as well as a broader array of supply chain activities (Rai et al., 2006, Scott and Westbrook, 1991). Although the findings from prior research are relevant and useful, the healthcare context presents some unique challenges for adopting and using IT to improve work practices and the delivery of key services.

First, products and equipment used in medical procedures can be extremely expensive; a basic stent used in a cardiac catheterization can easily exceed three thousand dollars. Second, doctors face high uncertainty in terms of the types and amount of product they will need, due to the diversity in patient characteristics (one size does not fit all when it comes to a stint) and fluctuating volumes. Third, the urgency of medical procedures—where the difference between life and death can be at stake—means that inventory checked out from the holding bins may not be accounted for properly, which makes inventory tracking difficult. Fourth, most hospitals utilize a consignment process for managing inventory, wherein vendors maintain a minimum level of inventory, and hospitals only pay for over-and-above the minimum inventory level. While this lowers the risk for hospitals, it creates issues related to product expiration and tracking due to lack of accountability of the representatives. These contextual characteristics that are specific to healthcare settings lead to our first research question:

1) How do contextual characteristics of healthcare services impact the adoption and use of RFID technologies for inventory management?

In addition to understanding the contextual characteristics that impact RFID adoption and use, it is also important to develop a more sophisticated understanding of the mechanisms through which RFID generates value in specific healthcare settings. Though past studies have moved from a firm-level to a process-level focus for understanding how IT impacts performance (Rai et al., 2006), both types of studies typically remain at a singular unit of analysis. Recent IS research argues that studies adopting a singular unit of analysis have inherent shortcomings, and encourages IS researchers to explore the linkages
between levels of analysis and the processes by which system adoption and use creates value (Burton Jones and Gallivan, 2007). Though the primary focus of this research is on the adoption and use of RFID for improving inventory management processes in healthcare services, we are also interested in the downstream consequences to the work practices that are embedded in the delivery of the specific healthcare services. Thus, inventory management processes fulfill and interface with the delivery of healthcare services, such as the two we have targeted for this study. Therefore, to understand the mechanisms through which RFID generates value (or the obstacles to generating it), it becomes critical to evaluate how the work practices of inventory management change, as well as the corresponding impacts to the work practices of the providers that deliver key healthcare services. Only then can RFID be effectively leveraged and the performance impacts accurately measured. Therefore, the importance of a multi-level analysis in this setting leads to our second research question:

2) How do changes in inventory management work practices (enabled by RFID) correspond to changes in work practices that are embedded in the delivery of key healthcare services?

In the following section, we provide the theoretical background and research model that guide this study, followed by a discussion of the methodology and the plan for completing the project.

2 THEORETICAL BACKGROUND

This article utilizes the IT-enabled process capabilities perspectives (Sambamurthy et al., 2003; Barua et al., 2004; Pavlou et al. 2006, Rai et al., 2006), to study the impact of RFID in healthcare inventory management. According to this perspective, resources refer to the firm's assets such as investments in IT infrastructure, whereas capabilities refer to processes and routines that help a firm to apply and reconfigure resources. The extent to which a firm is able to leverage, reapply and reconfigure its internal resources forms the fundamental basis of its competitive advantage.

Consistent with the IT-enabled process capabilities perspective (Rai et al., 2006), we characterize the investments in RFID technologies as resources, and the corresponding capabilities developed, as RFID-enabled process capabilities. Several technologies have been used for inventory management in the healthcare industry. For instance, bar code technologies combined with inventory management systems have been used in several cases for handling inventory (HealthCareIT News, 2004). However, the characteristics of an RFID-based system (discussed in RFID-enabled process capabilities section) provide unique opportunities for a firm to develop important process capabilities that were previously unavailable to the focal firm (Curtin et al., 2007; Dutta et al., 2007). To develop our arguments further, we draw from extant research on healthcare and supply chain management literature, to build our research model.

3 RESEARCH MODEL

The research model proposed in this article is shown in figure 1. The model postulates that healthcare process performance consists of three dimensions namely economic, structural and clinical performance. Healthcare characteristics, which include inventory uncertainty, patient safety and consignment characteristics are postulated to have a direct impact on healthcare process performance. By eliminating several inventory bottlenecks, RFID-enabled process capabilities are postulated to moderate the relationship between characteristics of healthcare service and process performance. Each of the constructs in figure 1 are described below.

3.1 Healthcare Process Performance
Healthcare process performance refers to the performance benefits realized by the focal hospital for inventory management. It refers to the overall value derived by the focal hospital through its implementation of inventory management strategies and is defined as consisting of three dimensions namely economic performance, structural performance and clinical performance (Solovy and Chaiken, 2003). These three dimensions refer respectively to the financial, work-related processes, and care-related processes in the healthcare context.

Economic performance refers to the return on investment (ROI) from the implementation of inventory management strategies (Solovy and Chaiken, 2003). IT investments are shown to contribute positively to the production of services (Menon et al., 2000) and to financial performance (Menachami et al., 2006) in the healthcare industry. However, several researchers including Bauer (2004) and Vogel (2003) argue that conventional analysis of ROI, as increase in revenues, has led to a misconception among healthcare practitioners that IT investments in healthcare does not yield positive ROI. They argue instead that in healthcare setting ROI should be computed as reduction in costs, and should be based on future cost of health personnel.

Structural performance refers to the process improvements in the functioning of various activities, through the implementation of the inventory management strategies (Solovy and Chaiken, 2003). It includes improvements such as workflow efficiencies in inventory management process, streamlining of inventory management activities, and reduction in processing time, enhanced inventory turns, less inventory shrinkage, increased visibility and provision of secure information.

Clinical performance refers to the process improvements in the various patient-care processes, through the implementation of the inventory management strategies (Solovy and Chaiken, 2003). It includes
improvements in processes such as reductions in inventory errors associated with a patient, reductions in inventory use due to mistakes, and streamlining of inventory ordering for patients.

These performance dimensions refer to the multi-level impacts in healthcare. While economic and structural performance refer to the inventory process performance (therefore at the process-level), clinical performance relates more to the delivery of care, which we consider to be the higher-order impacts. Therefore, to understand the mechanisms through which RFID generates value (or the obstacles to generating it), it becomes critical to evaluate how the work practices of inventory management change, as well as the corresponding impacts to the work practices of the providers that deliver key healthcare services within this context.

3.2 Characteristics of Healthcare Service

Healthcare context presents some unique challenges for realizing inventory management performance benefits. These contextual considerations are documented under the umbrella term healthcare characteristics. We refer specifically to three healthcare characteristics that have an impact on inventory management practices, and subsequently to healthcare process performance, namely: Inventory uncertainty, Task urgency and Consignment accountability.

3.2.1 Inventory Uncertainty

Hospitals typically manage their inventory in a centralized repository, and is transferred to individual departments based on the patient needs and the procedure performed (Ducllos, 1993). Most often, doctors have specific preferences for the type of inventory (such as stents, catheters) to be used in a procedure, than being dictated by hospitals or vendor consultants, and is argued to be a “big source of the inefficiency in healthcare” (McCarthy, 2006, p. 291). Also, predicting the exact product to be used in a procedure is difficult. This uncertainty in inventory requirements means that the number of inventory picked from the central location to the procedure room is usually over-and-above the required set of inventory. Inventory wastage in hospitals happens because such additional inventory is not replaced, accidentally misplaced, or is trashed. Such inventory wastage could contribute to lowering of economic performance in inventory management.

3.2.2 Task Urgency

Task urgency refers to the stress related to the task and the corresponding time pressure faced by the doctors, nurses and hospital staff while performing a particular procedure. Prior research on the work conditions in hospitals, has shown that work-related stress experienced by hospital staff and nurses are consistently high (Caplan, 1994; Stordeur et al., 2001). The stress level in operating rooms and catheter labs are even more pronounced, given the nature of the task performed, such as surgical and emergency procedures. As doctors and nurses face task urgency, the focus is more on the procedure performed, rather than on the handling or management of inventory used in the procedure. This implies that inventory tracking takes lower precedence compared to other important tasks, and could lead to decreased economic performance because of inventory wastage. In other cases, wrong inventory might be chosen in a hurry by the hospital staff and nurses. As the possibility of medical errors due to incorrect inventory exists (McNutt et al., 1997), task urgency could also lead to a decrease in clinical and structural performance.

3.2.3 Consignment Accountability

Consignment system refers to the practice by which vendors own the inventory, until such time it is actually used by the hospital (Ballard, 1991). This practice is a variant of the vendor managed inventory
(VMI) system (Williams, 2000), wherein the vendor assumes inventory risks while stocking, and takes responsibility for replenishing it, and the hospitals pay only after the inventory is used.

Although hospitals stand to benefit from such a system due to decreased inventory risks and costs, consignment system also has important limitations. First, in the consignment system, inventory check is performed by vendor consultants on a scheduled basis (batch activity), instead of on-demand basis (transaction-based) (DiGiacomo, 1991). For hospitals that see fluctuations in demand in inventory (such as accidents increase in certain times of the year), having scheduled inventory replenishment by the vendor could be problematic, due to inventory unavailability. Second, inventory tracking is typically done manually by a vendor consultant. This could lead to errors in inventory tracking, and can subsequently lead to ordering of unnecessary inventory. Both these aspects affect economic performance, as inventory costs increase due to potential last minute ordering of inventory. Clinical performance is affected because inventory expiry could cause problems in providing effective patient care. In addition, due to additional labor involved in constant inventory tracking, structural performance could also be affected.

3.3 RFID-Enabled Process Capabilities

Despite the commonly held belief that digital capabilities are fundamental to an organizations survival and growth, scholars still struggle to specify the underlying mechanisms that link IT investments to improved performance in organizational settings (Bharadwaj 2000). Recent work in this area continues to call on researchers to investigate the causal pathways that connect IT investment and firm performance (Rai et al., 2006; Sambamurthy et al., 2003, Ho et al., 2002). For instance, one area of particular interest in IS research today is the role and potential impact of IT in healthcare organizations (Kohli and Grover, 2008; Fichman et al., 2008).

As noted, the healthcare industry provides a unique context for organizational scholars. For instance, the management and distribution of hospital inventory is often discussed under the broad heading of inventory management (Nicholson et al., 2003). Inventory management is an often researched topic among operations research scholars (Scott and Westbrook, 1991; Turner, 1993) who focus largely on functional problems (Rai et al., 2006). However, inventory management takes on a new light in hospital settings, as one distinct feature relates to the use of periodic review par levels (also termed order-up-to level) for managing inventory (Nicholson et al., 2003). It is not the use of periodic review or par levels that creates the distinction in hospital settings. Rather, these par levels are seldom based on actual usage patterns, and instead reflect the desires of doctors and nurses which are often based on subjective criteria (Prashant, 1991).

In contrast to their storied past, hospitals and other healthcare establishments are expected to increase their investments in advanced technologies—such as RFID—to improve operational efficiencies in areas such as inventory management, and to enhance patient care (Datamonitor Report, 2008). However, to generate value from these IT investments, scholars increasingly argue that organizations must leverage the functionalities of these digital resources to create higher order capabilities, such as innovating business processes, shaping new strategy, and extending the enterprise network (Sambamurthy et al., 2003). This logic extends from the resource-based view of the firm (Barney, 1991), and suggests that digital assets alone are not sufficient for increasing firm performance (Powell and Dent-Micallef, 1997; Zara and Covin, 1993).

RFID enabled process capabilities can improve inventory management in specialty hospital settings by reducing par levels through increased product use visibility. Specialty hospital settings relate to practice areas where high risk procedures are conducted and high value (greater than $2000) products are likely used. By using RFID in these settings, the collection of information across activities—those that are coordinated within a specialty process (such as a cardiac catheterization or interventional radiology...
procedure)—is automated. As a result of the limited human intervention necessary for information collection, materials managers can get a near real time digital representation of material flow within practice areas. This newfound visibility can also be combined with and enable new work practices to develop higher order capabilities for improving cost management, enhancing patient safety, and enabling regulatory compliance in hospital settings.

4 RESEARCH METHOD

We will use a multi-case study method to explore the possible impact of RFID technologies on inventory management practices in healthcare settings, as well as the potential corresponding changes to work practices that are embedded in two distinct service areas: Cardiac Catheterization (CCL) and Interventional Radiology (IR). The case study approach was chosen because it affords the best opportunity to dive deep into the contextual conditions (Yin, 1994) that impact process performance in hospital settings, as well as to explore how RFID is used to improve work practices at diverse contexts. To determine case sites, we are guided by principles of theoretical replication, which lead us to choose sites where the phenomenon under investigation can be explored in its natural setting (Yin, 1989), and so we can test our assumptions that predict divergent results but for predictable reasons (Yin, 2003). Guided by this philosophy, we will conduct a retrospective analysis of four service delivery sites that exist in separate hospitals (two CCL labs and two IR labs). This approach allows us to ensure variance along important contextual characteristics that are assumed to impact process performance over time.

In addition to cross case comparison, we have also identified key roles within each study to explore the impact of RFID technologies at multiple levels of analysis. In doing so, we will interview individuals that have adopted and are now responsible for using RFID to improve inventory management within the healthcare settings (supply coordinators, purchasing managers, and inventory clerks). Additionally, we will interview individuals that participate in the work practices that are directly responsible for the delivery of services rendered in the CCL and IR labs in major US hospitals (doctors, nurses, lab attendants). Miles and Huberman (1994) refer to this as within case sampling, which is an additional way to develop rich insight from case study research (Chu and Robey, 2008). By conducting a multi-level analysis, our interest will not only be on the change that occurs within practice areas, but on the causal linkages that connect changes in inventory management practices with changes in the primary work practices, those that are directly related to the delivery of key healthcare services.

4.1 Data collection

This multi-case study will use three data collection techniques to explore the phenomenon and ensure adequate validity of research findings (Yin, 2003; Miles and Huberman 1994). The primary source of data for this study will be a series of semi-structured interviews. Interviews were first conducted from the perspective of a vendor of RFID solutions in the healthcare industry, and included discussions with C-Level employees such as the Chief Executive Officer, the Vice President of Hospital Services for CCL labs, and the Director of Hospital Services for IR labs. Upon completion of these early exploratory interviews with the vendor, we conducted our initial interviews with members of the first site identified, a CCL lab from a large hospital in the Northeastern part of the United States.

The interview protocol is semi-structured and consists of questions that are both targeted and open-ended to encourage normal conversation related to the peculiarities of the healthcare context and the impact of RFID on work practices and process performance. More specifically, questions ask respondents to assess aspects of the healthcare context that most influence their work, how their perceptions and work practices change at early and late stages of RFID adoption and use, how they use the RFID technology to alter their work practices, and how the overall performance of the unit has changed (in terms of employee
The second source of data will be direct observation, which reflects a data collection method where the researcher acts as a passive observer (Yin 2003). Observational evidence is particularly valuable when acting as a supplement to other data collection strategies (Mason, 2002). Mason (2002) states that the term ‘observation’ “usually refer to methods of generating data which entail the researcher immersing himself in a research ‘setting’ so that they can experience and observe at first hand a range of dimensions in and of that setting” (pg. 84). We leverage this method to observe firsthand the activities that occur in CCL and IR labs, how RFID technology is used to alter inventory management work practices that support these service areas, the procedural linkages that connect inventory management practices with the work practices that enable service delivery, and the contextual characteristics that shape behavior in these work domains.

The third source of data will be archival, which is an important source of data for this study as it provides a means to triangulate findings that emerge from semi-structured interviews and direct observations. In addition to power point slides, meeting notes, and white papers generated from both the vendor and hospital perspectives, we will leverage data generated from use of the RFID system. This data reflects usage patterns and changes to inventory levels of tracked products over time.

### Dimensions Working definitions and key interview questions

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Working definitions and key interview questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthcare Context</td>
<td>We focus on how actors view the unique aspects of the work domain and how they related to characteristics of the healthcare context. Such aspects relate to time pressure, issues with consignment, process importance (life and death scenario), and process complexity.</td>
</tr>
<tr>
<td></td>
<td><strong>Examples questions:</strong> What is unique about the healthcare context that impacts adoption and use of RFID technology? Please describe the contextual situation surrounding hospital X when they purchased the RFID solution. Are there unique aspects of the following (CCL, IR, OR, etc.) hospital units that affect adoption and implementation in different ways?</td>
</tr>
<tr>
<td>RFID-Enabled Process Capabilities</td>
<td>We focus on how individuals perceive that the RFID system has enabled changes in inventory management and healthcare service delivery in specific healthcare settings.</td>
</tr>
<tr>
<td></td>
<td><strong>Example questions:</strong> How has the use of RFID afforded you the opportunity to alter your work practices? How does increased visibility impact decision making related to inventory par levels and acquisition of new and existing assets? How do innovations in inventory management work practices impact patient care?</td>
</tr>
</tbody>
</table>
Process Performance

We focus on how participants view the impact of RFID technology on key performance outcomes within their work domain related to economic, structural, and clinical outcomes.

*Example questions:* How has RFID technology and the subsequent changes in work practices impacted inventory par levels of key products in your work domain? How has RFID technology and the subsequent changes in work practices impacted patient safety in your work domain?

<table>
<thead>
<tr>
<th>Table 1. Protocol for data collection</th>
</tr>
</thead>
</table>

### 4.2 Data analysis

We leverage a hybrid analytical technique to analyze the qualitative data, incorporating both inductive and deductive coding and thematic development procedures to generate rich insight (Fereday and Cochrane, 2006; Chiasson et al., 2008). The hybrid approach is particularly useful in novel settings where extant work is limited; it allows us to begin the analytical process by working from the data and move from specific cases to more general conclusions (Schwandt, 2001). Moreover, by alternating inductive and deductive analysis, we are able to leverage the IT-enabled organizational capabilities framework to interpret emerging themes, which will result in richer and more robust findings.

Two of the authors are conducting the interviews for this investigation. After each interview, both authors meet repeatedly in recap sessions to converse on major themes that emerge from the interviews. The third author is then presented with the generated themes and offers critique of the early analysis by reference to prior theory. Additionally, a third party is transcribing all interview files. During the early phases of analysis, the research team will thoroughly listen to, read, code, and discuss the data in an iterative fashion (Eisenhardt, 1989). Themes generated through analysis of individual interviews will be identified by tagging text segments with codes using computer assisted qualitative analysis tools. Saliency as an explanatory factor, rather than its frequency, represents our standard for what constitutes a theme (Glaser & Strauss, 1967; Blatt et al, 2006).

### 4.3 Expected Contributions

The findings from this research study have important potential implications for researchers and practitioners alike. From a research standpoint, the domain of healthcare IT is an emerging field requiring more theoretical and empirical research contributions. Our study attempts to make an important contribution in the healthcare IT context, by examining the contextual factors and impacts of RFID-enabled capabilities on healthcare process performance. Future research can build on our research model by seeking to understand the different dimensions of RFID capabilities in different healthcare scenarios such as patient-tracking, and investigating patterns of use of RFID in different healthcare contexts, and their corresponding impacts on adoption. Practitioners can stand to benefit from this research study, by gaining a better understanding of the various internal factors that inhibit their healthcare process performance. By understanding the impacts of RFID-enabled capabilities, practitioners can effectively utilize these capabilities, for improving their information visibility, tracking, and validity, thereby increasing their healthcare performance.

### 4.4 Project Status

Proceedings ECIS 2009
Our research is being conducted in two phases. The first phase focuses on qualitative interviews from individuals within the vendor firm and four hospitals in the United States. Currently, we have conducted interviews with respondents from the vendor firm and one major hospital. Our next steps will be to complete the interviews with representatives from the other three hospitals. Upon completion of this phase, we will move into the quantitative longitudinal analysis of archival data that are currently being collected by the RFID system within each hospital setting. Through this analysis we will triangulate findings generated from the first phase of our study, as well as search for new insights that may emerge from our longitudinal analysis.
References


Would I Use My Personal Blog for Commercial Exchange?

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0047.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>Blogging, Commercial exchange, E-commerce (B2B / B2C / B2G / G2C), Social computing</td>
</tr>
</tbody>
</table>
WOULD I USE MY PERSONAL BLOG FOR COMMERCIAL EXCHANGE?

Tan, Wee-Kek, Department of Information Systems, National University of Singapore, Computing 1, 13 Computing Drive, Singapore 117417, Singapore, tanwk@comp.nus.edu.sg

Tan, Chuan-Hoo, Department of Information Systems, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Kowloon, Hong Kong, ch.tan@cityu.edu.hk

Teo, Hock-Hai, Department of Information Systems, National University of Singapore, Computing 1, 13 Computing Drive, Singapore 117417, Singapore, teohh@comp.nus.edu.sg

Abstract

Despite the prevalent use of blog for various personal reasons, there is a paucity of research examining the commercialization of blog. Even though it is possible to forge strong buyer-seller relationships among blogger and readers in a blog social network through collective social interaction, a fundamental prerequisite towards blog commercialization is the blogger’s willingness to engage in commercial exchange. Consequently, this paper examines the blogger’s decision making process on using his/her blog for commercial exchange. The decision process involves the appraisal of two novel theoretical notions, namely blog-blogger-reader-product similarity (BBRPS) and perceived reader-negative reaction (PRNR). Drawing on the cognitive-motivational-relational theory, we posit that the blogger’s emotional responses to the two appraisal factors affect how the blogger copes with the exchange situation. The coping strategy adopted by the blogger in turn increases or decreases exchange intention. Results from an experiment we conducted provide support for this model. Specifically, when BBRPS is high and PRNR is low, the blogger would be inclined towards using his/her blog for commercial exchange. Even when PRNR is high, it is plausible for blogger to persist in the use of his/her blog for commercial exchange as long as the negative emotional barrier is not excessively high.

Keywords: Blogging, Commercial exchange, E-commerce, Social computing.
1 INTRODUCTION

In recent years, we have witnessed a growing interest among firms in harnessing blogs to reach out to blog readers as potential customers. These firms could engage the bloggers to post static advertisement banners, product reviews, products endorsement and products for sale on their blogs. Despite the increasing number of blogs engaging in commercial activities, there is still a paucity of related empirical research (e.g. Walsh 2007). This paper addresses the issue of how and why the resources embedded within the social network of blogger and readers can be harnessed for commercial exchange, i.e., the blogger exchanging a product for monetary returns with a blog reader.

To this extent, the collective social interaction, as manifested in the posting of contents by the blogger and the comments made by the readers in response to these contents, can lead to the formation of strong ties within the blog social network of blogger and readers (Lento, Welser, Gu & Smith 2006). That is, strong ties are characterized by close and frequent interactions (Nahapiet & Ghoshal 1998). In turn, these strong ties are useful towards forging strong buyer-seller relationships (Frenzen & Davis 1990; Marsden & Campbell 1984). However, a fundamental prerequisite towards the commercial utilization of the blog social network is the blogger’s willingness to engage in commercial exchange. It is thus important to develop an initial understanding of the blogger’s individual perception in this paper. This provides a basis for future research into the commercial viability of the collective blog social network, which includes the buyers’, i.e., the blog readers’, perspective.

With the above in mind, we propose two theoretical notions of blog-blogger-reader-product similarity and perceived reader-negative reaction that would be relevant to a blogger when he/she is deliberating whether to participate in a commercial exchange. Drawing on the cognitive-motivational-relational theory (Lazarus 1991), we model this deliberation as a decision-making process that generates emotive outcome affecting the blogger’s willingness to use his/her blog for commercial exchange.

2 THEORETICAL BACKGROUND

Personal Blog as a Commercial Exchange: A blog serves as a natural user-alike segmentation whereby the blogger and readers converge towards one or more shared interests, which is reflected through the content of the blog postings and comments. This is best represented by two characteristics explicated in the life cycle model of a personal blog (Gurzick & Lutters 2006). First, as a blog progresses from the first stage of non-directed personal storage, i.e., pure posting of daily events and thoughts, towards the later stage of being an established and interactive blog, i.e., readers could post comments and interact with the blogger, the blogger would have developed contents that are targeted towards interacting with the readers, thereby facilitating the formation of blogger-reader relationships. These contents and relationships could serve as the basis to facilitate commercial exchanges. Second, as more bloggers interlink their blogs with the others, they form a blog community (Kumar, Novak & Tomkins 2004). To the extent that blogs provide a conducive environment for the formation of communities (Nardi, Schiano & Gumbrecht 2004), it is posited that they possess even greater potential to expose individual bloggers to larger pool of like-minded bloggers and readers.

With the formation of a blog community, a blogger could pursue the opportunity of exchanging goods, services and money (Balasubramanian & Mahajan 2001). For instance, a philatelic blog could involve the blogger discussing stamp collection; and with the majority of the blog readers themselves being philatelic enthusiasts. Within this blog, both the blogger and the readers could exchange stamps with one another for monetary returns. Indeed, we could see that an economic leveraging attempt through a personal blog is dependent on two contextual factors: 1) the degree to which a blog, its readers and the products are related (i.e., similarity); and 2) the degree to which readers are against commercial activities in a blog.
Blog-Blogger-Reader-Product Similarity: A recurring theme in the personal and relationship selling literature has been the establishment and maintenance of a fruitful long-term buyer-seller relationship (Dion, Easterling & Miller 1995). The seller plays a central role in this relationship by possessing 1) referent power that is built on the perceived attitudinal similarity between the customer and seller; and 2) the ability to attract buyers based on perceived similarity of interests, beliefs, values and personality (Dion et al. 1995). These two factors have been posited to improve sales performance as well as the trust-building process with customers. In relation to our context, the use of blog for commercial exchange could serve as a personal selling process between the blogger and readers. Specifically, the similarity factor within this commercial exchange relationship could provide sufficient motivation for a blogger to participate in the exchange. This view is in accordance with the similarity attraction theory (Byrne 1971) which posits that people prefer to interact with like-minded others. Empirical evidence supporting this paradigm suggests that personality similarity, specifically attitude similarity, could result in positive interpersonal attraction (Singh & Tan 1992). It is reported that peoples are generally more attracted to similar others than to dissimilar others (Chen & Kenrick 2002).

Leveraging on the above line of reasoning, we present the notion of blog-blogger-reader-product similarity (BBRPS). This notion refers to the similarity of the interest-focused product, or category of product (i.e., product), with the interest-focused blog content (i.e., blog) and interest-focused relationship between the blogger and readers (i.e., blogger and reader). The phase “interest-focused” denotes the shared interest or common topic/theme of the blog. In other words, the blog content and the relationship between the blogger and readers should be closely related to the shared interest of the blog. Moreover, the product to be exchanged on the blog must be congruent with the blog’s shared interest. In the presence of BBRPS, we posit that the blogger should be strongly attracted to the readers and thus more inclined to engage in commercial exchange of the relevant products within the virtual social community revolving around a personal blog.

Perceived Reader-Negative Reaction: Prior research has suggested that as competition in the Internet intensifies, consumers are increasingly being targeted with marketing materials using push-oriented mechanisms (Johnson 1996, pp. 16). Such developments have lead to calls to address the issue of commercialism (for instances, see Johnson & Ambrose 2006). An interesting comment is raised by Donnelly (1996) who believes that the Internet should be divided into different areas where advertising and commercial traffic is acceptable and those where they are not. Just how negative could a perception on commercialism be? Some studies have advocated that a consumer’s attitude to commercialization could affect buying behavior (Mortelmans & Damen 2001). For instance, we could encounter a cultural critic which denounces the gift-giving practice during Christmas as “nothing more than materialism and unlimited consumption which are promoted under the veil of traditional values and habits” (Mortelmans & Damen 2001, pp. 157). In relation to our context, we believe that the negative perception of Internet commercialism could impede the use of personal blogs for commercial exchanges by dissuading readers from continuing to read these blogs.

Leading on from this line of reasoning, we posit a second theoretical notion of perceived reader-negative reaction (PRNR) which is defined as the blogger’s perception regarding the negative reaction of the blog readers in response to the use of the blog for commercial exchange. It is unlikely that a blogger would perceive reader-negative reaction as positive. Instead, we are more interested in the strength of the negative perception. Specifically, reader-negative reaction could manifest in the form of negative comments left on the blog suggesting non-purchase intention and a boycott of the blog. If the blogger perceives these comments as sufficient threats to the blog readership, i.e., the virtual social community that his/her blog is painstakingly building, it is conceivable that the blogger would sacrifice any thought of commercialism to retain the community.

Cognitive-Motivational-Relational Theory: Our present study adopts a decisional process model to trace the decision steps taken by the blogger when evaluating whether to continue to exchange. A decision scenario-based experimental approach (commonly used in the decision making literature such as Payne, Bettman & Johnson 1993) was also adopted in which the participants were told that they were currently using their personal blogs for commercial exchange. Moreover, the initial usage of the
blog for commercial exchange may not derive as much benefit for the blogger compared to its continued, sustained usage (Agarwal & Prasad 1997). Thus, to the extent that intention can usefully predict future usage (Davis, Bagozzi & Warshaw 1989), we focus on how the current usage of personal blog for commercial exchange can affect its continued usage intention.

Our process model builds on the cognitive-motivational-relational theory (CMRT) (Lazarus 1991). The CMRT postulates that a factor affecting a decision-maker’s stake in the outcome of a goal, i.e., his/her well-being, will trigger an emotional response (Lazarus 2000; Luce, Bettman & Payne 2001). If the factor facilitates the goal of the decision-maker, i.e., bringing him/her closer to the desired outcome, a positive emotion could be generated. On the contrary, if the factor impedes the decision-maker from achieving the goal, the resultant emotion becomes negative. Finally, the exact emotion elicited is dependent on the type of personal commitment involved. This entire process is collectively known as primary cognitive appraisal (Lazarus 1991). The CMRT further posits that a decision-maker could cope with the emotional encounter in two ways: with a problem-focused or emotion-focused reaction (Lazarus 1991; 2000). Problem-focused coping is associated with a planned action to alter the environment such that the overall emotional situation is improved. Emotion-focused coping attempts to alter the meaning of the events or regulate the expression of the emotion through avoidance or denial behaviors (Yi and Baumgartner 2004). In summary, the appraisal of BBRPS and PRNR as either facilitating or impeding the goal of using a personal blog for commercial exchange incites positive or negative emotion, respectively. The blogger adopts the appropriate coping strategy based on this emotional response, which in turn affect his/her choice of whether to continue exchanging or not (Luce et al. 2001).

**Conceptual Overview:** The theoretical concepts that we have discussed thus far is summarized in Figure 1. In gist, the blogger’s decision making process focuses on the appraisal of BBRPS and PRNR to determine his/her continued exchange intention in accordance with the CMRT. We believe that this is a necessary first step towards establishing the viability of commercializing the blog social network.

![Conceptual Overview Diagram](image)

**Figure 1 Summary of theoretical concepts.**

### 3 HYPOTHESES

**Blog-Blogger-Reader-Product-Similarity:** As emphasized earlier in the similarity attraction paradigm, shared interests, relationship building and transaction are three common motivations for consumers to join an online community (Hagel & Armstrong 1997). Considering these three motivations, an online community, such as a blog, could be perceived as a niche with a very narrow interest relationship that could be used to harvest commercial benefits (Pitta & Fowler 2005). The extant literature in the area of personal and relationship selling suggests that sellers are generally more likely to develop exchange relationships with consumers with similar interests (Wilson & Ghiold 1981). More specifically, similarity in a seller’s attributes such as appearance, lifestyle and socioeconomic status has been found to improve the effectiveness of relational selling (Crosby, Evans & Cowles 1990) and that a seller would sell primarily to prospects who share similar characteristics such as gender and age (Dwyer, Richard & Shepherd 1998). Leading from this, we reason that a high BBRPS should facilitate the commercial exchange goal of the blogger (Smith 1998). In conformity with the CMRT (Lazarus...
1991), we predict that a blogger could develop a positive emotion if the product that the blogger intends to sell is congruent with the interest-focused blog content that he has posted as well as the interest-focused relationship developed with the readers. On the contrary, low BBRPS would impede exchange and lead to negative emotion.

**H1a:** Blog-blogger-reader-product similarity is positively related with positive emotion.

**H1b:** Blog-blogger-reader-product similarity is negatively related with negative emotion.

**Perceived Reader-Negative Reaction:** Intuitively, if the readers react negatively to the commercial exchange activities to the extent that they would stop visiting the blog, it would be impossible for a blogger to achieve the goal of selling his/her products. In addition, extant literature on Internet commercialism suggests that a reader-negative reaction is likely to impede the blogger’s exchange goal. For instance, Internet users, in general, have been found to possess a strong negative attitude towards advertising and its societal effects (Previte & Forrest 1998). Specifically, Internet users develop a negative attitude towards Internet advertising that increases with their years of Internet usage experience. Moreover, the credibility of information on Internet sites has been found to be negatively affected by the presence of commercial advertisements (Burbules & Callister 1998). This indicates, in relation to our context, the possibility that blog readers could form a poor perception of the blog if it is to be used for commercial exchange (Gumbrecht 2004). Consequently, the blogger’s perception of this potential fallout with the readers is likely to deter him/her from selling products. In the light of this, we reason that a high PRNR should impede the commercial exchange goal of the blogger. In conformity with the CMRT (Lazarus 1991), we believe that a blogger could develop negative emotions in view of the perception of negative reader-reaction resulting from readers expressing non-purchase intention or threatening to boycott the blog. On the flip side, low PRNR would facilitate exchange and thus lead to positive emotion.

**H2a:** Perceived reader-negative reaction is positively related to negative emotion.

**H2b:** Perceived reader-negative reaction is negatively related to positive emotion.

**Coping Strategy:** Decision-making often entails the objective of minimizing negative emotions (see Bettman, Luce & Payne 1998). For instance, Yi and Baumgartner (2004) observed that worry, fear and anxiety resulting from undesirable and uncertain events could lead a buyer to adopt emotion-focused coping strategies such as mental disengagement. On the contrary, a buyer who could proactively deal with the situation would adopt problem-focused coping strategies such as planned problem-solving and seeking social support for instrumental reasons. This observation is consistent with the general propositions of Luce et al. (2001) on the effects of negative emotion on decision-making. Specifically, negative emotion may degrade cognitive performance or attention control as well as motivate the decision-maker to avoid making a choice (i.e., adopting the emotion-focused coping strategy). However, negative emotion could also motivate the decision-maker to approach the choice through an increased processing effort, thus suggesting the use of the problem-focused coping (i.e., adopting the problem-focused coping strategy). On the contrary, it has been shown that positive affects can provide support for continued coping efforts, broaden an individual’s attentional focus and increase the available behavioral response, i.e., the adoption of the problem-focused coping strategy (Folkman & Moskowitz 2000). Consequently, a positive affect provides the capability to induce creativity and flexibility in thinking and problem-solving, as well as to process important information (Folkman & Moskowitz 2000). The broaden-and-build theory proposed by Fredrickson (2001) adds that positive emotions such as joy and pride can increase an individual’s physical, intellectual, social and psychological resources.

**H3:** Positive emotion is positively-related to problem-focused coping.

**H4a:** Negative emotion is positively-related to problem-focused coping.

**H4b:** Negative emotion is positively-related to emotion-focused coping.
**Continued Exchange Intention:** Prior research has suggested that problem-focused coping could motivate the decision-maker to exert increased decision effort while emotion-focused coping could induce avoidance behavior (Luce et al. 2001). Furthermore, an individual will attempt to manage the source of a stressful emotional experience with problem-focused coping. From a negative perspective, an individual who opts for emotion-focused coping will alter the meaning of the event or regulate the expression of the emotions (Yi and Baumgartner 2004). Consequently, we believe that problem-focused coping strategies should enable a blogger to explore the potential monetary benefits associated with commercial exchange brought about by BBRPS while simultaneously resolving the threats posed by a PRNR that is caused by Internet commercialism. On the contrary, emotion-focused coping strategies should result in avoidant behavior in which the blogger avoids the perception of threat caused by reader-negative reaction by boycotting the blog.

**H5:** Problem-focused coping is positively related to continued exchange intention.

**H6:** Emotion-focused coping is negatively related to continued exchange intention.

4 **RESEARCH METHODOLOGY**

A 2 (low vs. high BBRPS) × 2 (low vs. high PRNR) full factorial between-subject design was employed. The target population of this study comprises Internet users who are existing bloggers. For each blogger who had responded to our survey, we requested his/her blog address and we painstakingly validated each blog for authenticity. The reason for not including the non-bloggers is that they would reduce the internal validity of the study since they suffered from a lack of the required contextual knowledge and experience. The two independent variables, i.e. BBRPS and PRNR, were operationalized in the form of a treatment blog in the experiment to stimulate the decision-making process (see Figure 2 for a sample blog used in the experiment). In addition, there was a control blog, which was a typical personal blog used as an online diary for a typical undergraduate. It contains 4 postings about the daily life of the undergraduate. The sole objective of the control blog was to orientate the participants to the blog structure used in the experiment. The treatment blog extends the control blog to sell a product, i.e. a shopblog selling Apple iPod Nano MP3 players. Both blogs are structurally similar but with three exceptions in the treatment blog: 1) Posting 1 and the area below the blogger profile are used to advertise the product being sold; 2) Posting 3 is used to discuss the undergraduate’s interest; and 3) Only the treatment blog contains readers’ comments to the postings. However, the participants were not expected to make any recall comparisons that could bias the results since they cannot return to read the control blog and the control blog did not involve product selling.

The low BBRPS was manipulated by portraying the blogger as an avid stamp collector (i.e. blogger similarity) who shared his interest in stamp collecting on the blog (i.e. blog similarity), and which attracted many stamp collectors (i.e. reader similarity). All three aspects are reflected in the blogger profile.
profile with the last one also being reflected in the comments made by the readers. Additionally, Posting 3 was mainly about how to start a stamp collection (i.e. blog similarity). The high BBRPS was manipulated by portraying the blogger as an MP3 player enthusiast who shared his own MP3 purchase and usage experiences with the readers on the blog, and which attracted many MP3 player enthusiasts. Posting 3 mainly contained information on how an MP3 player works. Essentially, the low and high BBRPS was manipulated with a philatelic blog and MP3 player blog respectively (see Figure 3). The low PRNR was manipulated by means of readers’ comments reflecting positive perceptions towards the selling of MP3 players and positive purchase intentions; while the high PRNR was manipulated using readers’ comments reflecting negative perceptions towards the selling of MP3 players (i.e. threats to boycott the blog) and negative purchase intentions.

Figure 3  Manipulation of BBRPS.

MP3 player was chosen as the product since it is a common community product that is gender-neutral and can be considered a search good. It is thus feasible for the readers to purchase MP3 players from the blogger based on what they had read in the blog. While MP3 players can be purchased from other electronic commerce shopping websites, comparing the sales channels is not the objective of this paper. Moreover, there is no convincing reason to believe that the participants had considered sales channels when indicating their response since the controlled experiment made no reference to them other than the blog. Nonetheless, it is necessary to control for the participants’ familiarity with the product, i.e. the MP3 player, and brand, i.e. Apple iPod Nano. This is because, to the extent that consumers generally have a better impression of new product information associated with familiar brands (Kent & Allen 1994), bloggers could be more inclined towards exchanging familiar products or brands versus unfamiliar ones. Finally, we also controlled for the participants’ demographics, Internet usage experience and blogging experience through random assignment of participants to the treatment conditions. As much as possible, the constructs in our research model were operationalized and measured using existing scales that have been validated in prior studies (see Table 1). The remaining scales were developed based on the recommendations of Moore and Benbasat (1991). All scales are 7-point Likert (1-Strongly Disagree to 7-Strongly Agree) scales except for product and brand familiarity.
Problem-focused Coping (PFC) (Yi & Baumgartner 2004)
With respect to the problems that I might encounter in the selling of products in my personal blog; I will:
1. think about how I may best handle the problem.
2. try to come up with a strategy on what I should do.
3. think about what steps to take to resolve the problem.

Emotion-focused Coping (EFC) (Yi & Baumgartner 2004)
With respect to the problems that I might encounter in the selling of products in my personal blog; I will:
1. try to keep my feelings to myself.
2. try not to show other people how I do really feel.
3. try to hold back my feelings.

Continued Exchange Intention (CEI) (Terry, Hogg & White 1999; Agarwal & Prasad 1997)
If I have additional Apple iPod Nano MP3 players available for sale within the next 12 months; I intend to continue to use my personal blog to:
1. sell them.
2. put them on the market.
3. use my personal blog to advertise them.
4. use my personal blog to put them up for sale.

Product Familiarity (PF) (Johnson & Russo 1984)
1. I am not familiar at all – very familiar with MP3 players.
2. I am not knowledgeable at all – very knowledgeable about MP3 players.

Brand Familiarity (BF) (Allen & Kent 1994)
1. I am not familiar at all – very familiar with Apple iPod Nano.
2. I am not knowledgeable at all – very knowledgeable about Apple iPod Nano.

Table 1 Instrument scales used to measure constructs.

The experiment was conducted in April 2008 through the Internet. Invitation messages were posted on local major Internet blogging forums. All respondents were checked to ensure that they had an existing personal blog that had been actively maintained for at least the past 12 months leading up to the experiment. The experiment began with the participants reading the opening narrative, which asks them to assume the role of an undergraduate who maintains a personal blog as an online diary, i.e. the control blog. The participants next proceeded to read the control blog for 5 minutes and answered a series of questions on the blog’s content. The participants could not proceed until all the questions were correctly answered. The participants next read the scenario narrative which tells them that they also blog about their personal interest. More importantly, the participants were told that they would be selling Apple iPod Nano MP3 players on their personal blogs. The participants then read the respective treatment blogs assigned to them for 7 minutes and answered another series of questions on the blog’s content. Thereafter, the participants were told that they were to make a decision on whether to continue to use their personal blog, i.e. the treatment blog shown to them, to sell the MP3 players.

In summary, the adoption of a scenario-based approach for the experiment is appropriate since not all the subjects might have experience on commercial blog (Sheng, Nah & Siau 2006). It also allows us to isolate which treatments have a differential impact on the subjects’ emotional response.

5 RESULTS

A total of 238 participants responded to our invitations. There were 13 incomplete entries (5.46%) and 1 participant (0.42%) who reported a blog address that was ascertained to be doubtful. These were excluded from the data analysis leaving us with 224 participants. Although the number of male participants was lower than the female participants (27.23% versus 72.77%), the univariate analysis of variance (ANOVA) tests indicated that there was no significant difference in the participants’ mean ratings of BBRPS (Mean = 3.99, δ = 1.523, F (1, 222) = 0.108, p = .743) and PRNR (Mean = 3.87, δ = 1.689, F (1, 222) = 1.498, p = .222) across gender. There was also no significant difference in the participants’ Internet usage and blogging experiences across treatment conditions. Interestingly, a sizeable percentage of the participants engaged in some forms of commercial activities using their blogs. Altogether, 46.90% of the participants had previously engaged in at least one form of commercial activity. However, the univariate ANOVA tests indicated that there was no significant difference in the participants’ mean ratings of BBRPS (Mean = 3.99, δ = 1.523, F (1, 222) = 3.146, p = .077) and PRNR (Mean = 3.87, δ = 1.689, F (1, 222) = 0.017, p = .897) across experience in commercial activities.
Manipulation checks were performed to ensure that our manipulation of the treatment blogs reflecting low and high levels of BBRPS and PRNR were successful. Independent samples t-tests using the mean of the respective sets of instrument items indicated that our manipulations were successful: 1) The mean score of BBRPS for the low treatment group (Mean = 3.04, δ = 1.288) was significantly lower than the high treatment group (Mean = 4.94, δ = 1.088), t = -11.924, p < .001; 2) The mean score of the PRNR for the low treatment group (Mean = 4.78, δ = 1.556), t = -9.567, p < .001. In addition, the univariate ANOVA tests on the product and brand familiarity indicated that there was no significant difference in the mean value of these two control variables across treatments: 1) for product familiarity, Mean = 5.39, δ = 1.132 [F (3, 220) = 0.642, p = .589]; and 2) for brand familiarity, Mean = 4.53, δ = 1.591 [F (3, 220) = 0.376, p = .770]. In general, participants across treatments had a consistently high level of product and brand familiarity. Coupled with the consistent Internet usage and blogging experiences of the participants across treatments, it appeared that the random assignment was successful in controlling for differences in participants’ characteristics across treatments.

**Figure 4 The structural model.**

Data analysis was performed with structural equation modeling in LISREL 8.54. The measurement model was first tested using a confirmatory factor analysis. Most of the model fit indices were greater than their respective thresholds (e.g. RMSEA = 0.064; SRMR = 0.053; GFI = 0.85; NFI = 0.92; NNFI = 0.96; CFI = 0.96). Thus the measurement model had a good fit with the data. In addition, the convergent validity, internal consistency, discriminant validity and reliability of the measurement items were assessed to be above their respective thresholds. The correlations among latent constructs did not indicate any multicollinearity problem. Figure 4 depicts the results of the structural model. We observed BBRPS to be positively-related to positive emotion and negatively-related to negative emotion, thus supporting H1a and H1b. The PRNR was found to be positively-related to negative emotion and negatively-related with positive emotion, thus supporting H2a and H2b. Positive emotion was positively-related to problem-focused coping thus providing support for H3. Though negative emotion was positively-related to problem-focused coping; this result is not considered significant. We observed that negative emotion was positively-related to emotion-focused coping. Thus, H4a is not supported while H4b is supported. Problem-focused coping was positively-related to continued exchange intention thus supporting H5. Finally, emotion-focused was observed to be negatively-related to continued exchange intention but this is not deemed significant. Thus, H6 is not supported.

6 DISCUSSION

The integrative decisional process model shows that there are three distinct emotion-coping paths that can lead to either higher or lower continued exchange intention: 1) positive emotion to problem-focused coping; 2) negative emotion to problem-focused coping; and 3) negative emotion to emotion-focused coping. This suggests that consumers may use different coping strategies depending on their emotional responses to online blogs. Understanding these mechanisms can help in designing effective marketing strategies that cater to the emotional needs of consumers.
focused coping. The exact path that the blogger would traverse is contingent on the cognitive appraisal of the two decision factors (Luce et al. 2001). Specifically, high BBRPS facilitates the blogger’s goal of exchanging and positive emotions should be elicited from the blogger (Lazarus 1991). The blogger should then acquire additional coping effort (Folkman and Moskowitz 2000) leading to the adoption of problem-focused coping, i.e. the first path. This enables the blogger to explore options in utilizing the blog for commercial exchange, which can be translated into higher continued exchange intention. For instances, the blogger could create more useful postings on topics relevant to the products or use the blog for after sales communication with the readers.

In contrast, high PRNR impedes the exchange goal and induces negative emotion (Lazarus 1991). On the one hand, if the blogger opts for problem-focused coping, i.e. the second path, he/she could attempt to pacify the readers by making posts designed to project the blog as a one stop location for readers to acquire product knowledge before making purchase. The blogger could also attempt to reply to each negative comment cordially to appease the readers or seek advices from other peers facing similar problem. On the other hand, if the blogger opts for emotion-focused coping, i.e. the third path, he/she is likely to suppress the negative affection associated with the dual threats of non-purchase intention and boycott actions (Yi and Baumgartner 2004) thus increasing the avoidance tendency. For instance, the blogger might refuse to seek advice from other bloggers facing similar situation.

We may therefore surmise that when BBRPS is high and PRNR is low, the blogger would be inclined towards using his/her blog for commercial exchange. Even when PRNR is high, it is plausible for blogger to persist in the use of his/her blog for commercial exchange as long as the negative emotional barrier is not excessively high. This is consistent with the notion that blog is a personal “protected space” (Gumbrecht 2004) and thus the blogger might be more inclined towards retaining the readers at the expense of commercial exchange. In fact, H4a only approaches but did not reach statistical significance. Moreover, the non-significance of H6 could be attributable to the fact that the blogger activated social support seeking, which is one type of emotion-focused coping strategy (Yi & Baumgartner 2004). When the blogger approaches friends for advice or emotional support, it may be perceived by the blogger as self-justification (Bettman, et al. 1998) to minimize the negative emotion caused by high PRNR. Consequently, the blogger could feel vindicated in exchanging commercial products despite unfavorable readers’ comments. In gist, the results of the experiment demonstrate that bloggers could be willing to engage in commercial exchange when contextual factors are favorable. It is thus feasible to leverage blog social network for commercial purposes (see Figure 1).

However, future research must still investigate the decision making process from the blog readers’ perspective (see Figure 1). To this extent, the present process model may be adapted to study the exchange intentions of blog readers. More specifically, the present sellers’ appraisal factors of BBRPS and PRNR may be complemented with other appropriate theoretical constructs to create a set of buyers’ appraisal factors. For instance, the perceived tie strength (Nahapiet & Ghoshal 1998) among other readers with the blogger and perceived trust towards the blogger could be critical factors determining whether a particular reader would exchange with the blogger (Kim & Srivastava 2007).

7 CONCLUSION

This paper has several limitations that readers should be cautioned. First, given that H4a is not supported, future research should examine how and why high PRNR could simultaneously lead to problem-focused and emotion-focused coping. In this regard, whether the blogger perceives a loss of control over the given commercial exchange situation could differentiate the coping behavior (Yi and Baumgartner 2004). Second, this study examines a single product type, i.e. MP3 players, which may not be a suitable one for all the subjects. Finally our model suffers from low $R^2$, i.e. low predictive power. A likely cause could be that the product used might have incited neutral response. Nonetheless, the ability of our model in establishing significant causal relationships between key constructs, i.e. descriptive and explanatory purposes of theory, is a significant one.
Despite these limitations, this study has made numerous contributions. Theoretically, we conceive BBPPS and PRNR in conjunction with the development of a novel decisional process model. These could be extended or generalized to study the impact of commercialism on other social computing tools. Practically, we explicate how personal blogs could be carefully utilized for commercial activities, i.e. capitalizing on the similarity between the blog, blogger, readers and the products. Additionally, the identification of PRNR as a plausible impediment factor implies that individual bloggers should exercise caution when turning their personal blogs into commercial tools.

References


CONSUMERS’ PERCEPTION OF CONTROL OVER ONLINE INFORMATION DISCLOSURE. AN ELECTRONIC FOCUS GROUP STUDY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0273.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
</tbody>
</table>

scholarONE
Manuscript Central

Proceedings ECIS 2009
CONSUMERS’ PERCEPTION OF CONTROL OVER ONLINE INFORMATION DISCLOSURE. AN ELECTRONIC FOCUS GROUP STUDY

Lähteenmäki, Mirella, Helsinki School of Economics, Runeberginkatu 14-16, 00100 Helsinki, Finland, mirella.lahteenmaki@hse.fi
Bragge, Johanna, Helsinki School of Economics, Runeberginkatu 14-16, 00100 Helsinki, Finland, johanna.bragge@hse.fi
Sunikka, Anne, Helsinki School of Economics, Runeberginkatu 14-16, 00100 Helsinki, Finland, anne.sunikka@hse.fi

Abstract

This study focuses on consumers’ perception of control over personal information disclosure on the Internet. Specifically, we examine how consumers perceive controlling their personal data that online companies collect for marketing and customer relationship management purposes. We aim to answer this research problem by clarifying 1) how do consumers express the perception of control over their personal information, and 2) how do consumers perceive controlling their personal information disclosure. Our empirical data is based on four computer-mediated focus group interviews. Our findings show that the perception of control is combined with all stages of personal data utilization: collection, storage and usage. Thus, consumers keep these stages in mind when thinking about their attitudes towards the collection and offering of their personal information. The interviewees also spontaneously mentioned various means with which they control personal data. Perceived trust towards companies, own initiative and permission-based marketing were also combined to the control speech. In summary, the interviewees mostly perceived that they were not controlling their personal data on the Internet. Only when they were talking about control methods of the information disclosure stage, they expressed the perception of control.

Keywords: Privacy, Online self-disclosure, Perception of control, Focus groups.
INTRODUCTION

The development and adoption of information and communication technologies (ICT) among consumers have enabled large-scale utilization of personal data for marketing purposes (Moon 2000). With personal data we mean information that can be connected to an identifiable individual (Culnan & Milberg 1998). This information can be grouped to identifier information (such as an e-mail address), sensitive personal data (information concerning e.g. health), and preference information (e.g. hobbies) (Andrare, Kaltcheva & Weitz 2002). Especially in the Internet, consumers increasingly face situations in which they have to provide personal data for some purpose (Heikkinen et al. 2004). Ordering and buying products, experimenting and adopting e-services, and registering into various loyalty customer programs are typical examples of situations in which submission of personal data is required. The data collection is often well justifiable from the consumer’s point of view; for example an e-commerce site needs customers’ contact information in order to deliver purchases, or to inform customers about possible delivery problems. However, companies collect personal data also for other purposes, for example to be able to advertise their products, or to utilize data in some other way in the future. This is usually acceptable, sometimes even good customer service. However, the customer should be aware 1) that her personal data is utilized, and 2) of the rights granted to her by legislation (e.g. the rights of inspecting, correcting, and prohibiting the usage of personal data). Customers should be able to execute these rights also in practice (Personal Data Act 1999, Velasquez 2006: 290-291).

According to several studies (e.g. Sayre & Horne 2000, Dommeyer & Gross 2003, Chellapa & Sin 2005), most consumers accept the collection and utilization of their personal data for marketing and customer relationship management (CRM) purposes in return for proper benefits. However, privacy questions have received much attention during the last few years (Milne 2000). According to several studies (e.g. Phelps et al. 2000, Eurobarometer 2008, Privacy & American Business 2004) consumers are concerned about their privacy, especially in the Internet. To some extent, the Internet is experienced as an unreliable and unsecure environment (e.g. Knights et al. 2001, Suh & Han 2003, Jensen et al. 2005).

Privacy itself is a multidimensional and ambiguous concept, the definition of which is often based on sociologist Alan Westin (1967, p. 7): Privacy is both control over 1) information disclosure and 2) the environment in which transactions occur. The first dimension is concerned with control over own information, while the latter with the control over one’s own state and of protection from different external intrusions, such as spam. In this study, we concentrate on the first dimension, on control over providing personal data, which, according to Goodwin (1991), can be seen to include also other control aspects than just disclosure regarding personal data utilization. Posner (1981), for example, refers to privacy violation when observing an individual without disturbing her concretely.

Research over personal information control has traditionally emphasized control over the consumers’ environment, for instance the intruding nature of direct marketing (e.g. O’Malley et al. 1997, Evans et al. 2001, Morimoto & Chang 2006). This area has been regulated, e.g., by establishing prohibition registers for marketing measures. Nevertheless, ICT-enabled interactive marketing actions pose more serious threats to consumers’ personal data control. Goodwin (1991) includes the risk of potentially harmful usage to the control over personal data – it is for instance possible to combine separate pieces of personal data and registers together. In that case, the decision whether to disclose or not a single piece of data will be of minor significance with respect to data protection.

The control over disclosing personal data has been examined in connection with several studies (e.g. in Phelps et al. 2000, Sheehan & Hoy 2000, Evans et al. 2001), where it has been found an important factor affecting the provision of data. However, the control over disclosure has seldom been the main focus in privacy studies, apart from a few exceptions. For example, Goodwin (1991) has examined privacy conceptually as the control over personal data and own environment, whereas Olivero and Lunt (2004) have studied whether the consciousness data collection influences the control and trust of consumers. On the other hand, the methods consumers are using to control their personal data have
been widely studied (e.g. Foxman & Kilcoyne 1993, Cranor et al. 1999, Dommeyer & Gross 2003).
The objective of our research is to increase the understanding of consumers’ perceived control over
their personal data that companies operating on the Internet collect for marketing and CRM purposes
(see e.g. Milne 2000, Romano & Fjermestad 2003). Consequently, we restrict our study to this
business-to-consumer context as distinct, for example, from authority transactions or criminal
operations. We aim to answer our research problem by investigating 1) how do consumers express the
perception of control over their personal data, and 2) how do consumers perceive controlling their
personal information disclosure. Next, we will present the research methodology and data, after which
we discuss our results in Section 3. The study finishes with discussion and conclusions.

2 RESEARCH METHODOLOGY AND DATA

A number of research techniques have been developed in order to better understand consumers, and to
assess their desires and preferences. These techniques include surveys, focus groups, statistical
modeling, and video ethnography (Prahalad & Ramaswamy 2004). As control over information
disclosure has been previously studied mostly in connection with privacy surveys (see review in
Kobsa 2007), we felt that an in-depth qualitative study is warranted in order to get a more
comprehensive view on the topic. Thus, we chose focus groups as our research method.

The research data was collected in Finland in September 2006 through focus group interviews that
were conducted electronically via computers in the same-time – same-place mode (Kontio et al. 2007).
That is, the participants were invited to a computerized meeting room to answer structured questions
presented by the researchers, and also to comment on each others’ thoughts. This mode represents
focus group interviews instead of focus group discussions (see Boddy 2005), as we wanted to receive
answers from every participant, and also to prevent the derailing from the issue that easily happens in
more free-form discussions. We utilized a group support system (GSS) called GroupSystems®
MeetingRoom, which has originally been designed to support teamwork and group decision making
(Nunamaker et al. 1991), but which can equally well be used for carrying out of focus group
interviews or discussions (Clapper & Massey 1996, Kontio et al. 2007, Tellefsen et al. 2005). A
typical GSS meeting consists of 10-30 networked computers in the same room, which gives the
facilitator and the interviewees the opportunity to communicate also verbally with each other. It is also
possible to conduct fully virtual meetings via the Internet (Klein et al. 2005, Kontio et al. 2007), but
these require usually audio or web conferencing support in addition. The utilization of different
computer-mediated systems in marketing and consumer research has been steadily increasing (see e.g.

Our study comprised of four group sessions, each lasting 2.5 hours, with 11-14 interviewees in each
group. The sessions were administered by the same main facilitator and two other adjunct facilitators.
The interviewees were posed a small set of related questions at a time, and they were allowed to
answer them in the order they wished, before the facilitator let them proceed to the next bundle of
questions. The interviewees saw each others’ answers, and they were allowed to comment on them or
to further deliberate on their own answers. We opted to carry out the sessions fully anonymously so
that all the comments remained individual. After each session ended, all the interview answers were
immediately available in writing using the automated reporting function of the GSS tool. This was one
important reason for selecting computer-mediated instead of traditional focus groups as the research
method, as our research partner (retail bank) needed to have the preliminary results as soon as possible
in order to plan a subsequent field experiment at its online bank (see Bragge et al. 2007 for details).

The computer-mediated focus group interviews can be seen as an intermediate form between
traditional face-to-face and fully virtual focus groups (Kontio et al. 2007). Some of the strengths of the
method compared to traditional oral interviews are, for example, their efficiency (enabling bigger
group sizes, simultaneous communication, and automated reporting), digital group memory during and
after the sessions, as well as anonymity. The advantages compared to a fully virtual implementation
are the possibilities to detail questions further, to interact verbally, as well as to present confidential or
physical support material (such as product prototypes) during the sessions. The weaknesses of the method compared to traditional groups are certain deficiencies in expression (the lack of gestures and facial expressions in communication), the need for special premises and technical expertise, and the possibility for technical problems. Compared to fully virtual discussions the computer-mediated same-place settings place restrictions on the participants’ geographical recruiting and are more expensive with respect to travelling expenses and venue costs (Klein et al. 2005, Kontio et al. 2007). As the topic of online information disclosure and the control thereof is fairly abstract, we did not consider fully virtual focus groups. In face-to-face settings possible ambiguities can be easier clarified.

Altogether 53 persons participated in our interviews in four groups, which is considered as a good amount for focus groups (Morgan 1996). The pilot group consisted of 14 business university students using online banking services. The other interviewees were randomly selected from among the online customers of a Finnish retail bank. Thus, the interviewees had at least basic computer skills to be able to participate in the electronically mediated interviews, in addition to which they were able to draw from their own experience when asked opinions about online information disclosure. Furthermore, the previous online bank usage made it possible for the interviewees to refer to a concrete context regarding the questions that might otherwise have felt too abstract for them. Thirty of the participants were female and 23 were male, all between the ages of 21 and 49. Occupationally, the interviewees were mainly higher or lower level officials (26), students (20) or workers (11). The majority of the interviewees used Internet daily (46); either at home (33), at work (28) and/or at their study place (11).

The underlying setting was a situation in which a reliable company (for example a bank) collects information online in order to utilize it for marketing and CRM purposes (e.g. for personalized advertising or services). This way we wanted to guarantee that the context experienced by all interviewees would be similar, and that no extreme context or experience (e.g. a foreign e-auction) would affect the answers too much. Furthermore, we were able to inquire their attitude also for the collection of sensitive personal data (concerning e.g. financial situation). Our questioning frame covered the following themes: 1) attitudes towards the collection of personal data on the Internet; 2) the recognition and control over collection and utilization of personal data online; 3) the understanding of the privacy protection concept. Thus, we were able to examine the control and disclosure of personal data both with direct questions and indirectly, without prompting the interviewees too much.

We asked the interviewees’ attitudes also towards the personalization of online services and to the utilization of personal data especially in the banking sector (see Sunikka et al. 2007). These questions are here dealt with only in so far as they are related to the control over information disclosure.

The content analysis of our data started from the question that was presented to the interviewees: “Do you perceive that you are in control over the collection and utilization of your personal information on the Internet?” Utilization, which consumers often understand as usage was brought up in addition to collection, because one can see that the purpose of usage affects the disclosure of personal data (cf. Andrare et al. 2002). First, we developed conceptual labels attached to the interviewees’ verbatim descriptions. Second, based on these conceptual labels, categories were formed at a more abstract level of classification. At the same time, the labels were constantly revised by merging categories expressing similar concepts, or by splitting a category representing different concepts (cf. Olivero & Lunt 2004). The categories thus formed were 1) disclosing, 2) storing, 3) usage, and 4) control methods of personal data. Finally, these main categories were studied thoroughly using the whole data that was collected. This was done by coding the data according to the formed categorization.

3 RESULTS

We will next present the analysis of the empirical findings of this study. Firstly, we will examine what kind of expressions the interviewees used when talking about the control over personal information on the Internet. Secondly, we will examine how the control over personal information in the Internet context is perceived by the interviewees.
3.1 The conception of control over personal data

The control over personal data proved to be an important theme. The interviewees brought it up when asked directly; in 39 of 56 comments the interviewees stated that they do not perceive that they are in control over the collection and utilization of their personal data online. Furthermore, the interviewees spontaneously expressed their concerns about the storage of personal data indicating a fear that third parties could get a hold on the information. This fear was mentioned in 31 (out of 55) comments for the question that inquired the interviewees’ own definition of data privacy protection. Furthermore, 14 comments were connected to the protection from undesirable usage. From the viewpoint of companies, the collection, storing and usage of personal data forms a process of personal data utilization, about which the interviewees thus talked with a negative tone because of their perceived lack of control.

The interviewees talked positively mainly about the methods of control, such as restricting the disclosure of their personal data. Thus, when talking about the control methods, the interviewees expressed the control perception of their data. The methods of control were connected mainly to the offering stage, and the main method was the decision not to disclose personal data. In most answers, the methods of control were blended with expressions of the control perception. In a couple of cases the interviewees, however, expressed both the lack of control and the deployment of control methods somewhat conflictingly at the same time: “I do not perceive well enough [that I am in control over the collection and utilization of my personal information on the Internet]. On the other hand, a huge amount of issues can be found out, but one must be critical herself and choose whom to believe”.

The interviewees also talked about own initiative and permission-based marketing, which can be seen as one type of control manifestations (see Milne & Gordon 1993, Evans et al. 2001). With own initiative the interviewees referred mainly to the desire to make the first move themselves in instigating a customer relationship (pull approach), and not just being a passive target for advertising. If customers perceive themselves as objects of push marketing, personal data is not willingly disclosed: “…one must still always be careful, where to provide personal information. It is always better if you have created the contact yourself, than if it would be an impulsive answer to some advertisement.” The motives behind taking own initiative were the control over own environment (e.g. restricting targeted advertising), and control over own information: when consumers take the initiative themselves, it is easier for them to act deliberately. This view is compatible with the contemporary conception of an active consumer (Pavlou et al. 2002).

Regarding permission-based marketing the interviewees expressed their desire to make a conscious decision of whether to accept the utilization of personal data, and to control their own information this way. The benefits gained from providing personal data were emphasized. “It is rather understandable [that companies collect information from and about you], but the protection of privacy is, of course, put to the test. To become a victim of, for example, advertising or spamming, is slowing down the computer and thus really strenuous. If the information I receive is related to my own interests, and approved by me, then it is quite OK”. A few interviewees (8/55 comments) associated the permission-based marketing system with their perception on privacy protection. In general, the fact that permission is asked before collecting and utilizing data for marketing purposes was appreciated.

Finally, the trustworthiness of a company (also its familiarity, or an existing customer relationship) added the willingness to disclose personal data. In addition to personal experiences, the interviewees’ trust was increased if the company is known or domestic, in which case the reputation of a company and the legislative environment (domesticity) increased the feeling of safety. A similar typology of trust was presented also by Milne & Boza (1998) regarding consumers’ views on utilizing personal data in marketing. In their study, the trust was manifested in earlier experiences of the company, in the reputation of the company, and in the confidence which is based on regulation or agreement.

Trust in a company was often mentioned as a prerequisite for the disclosure of personal data. We could interpret from the interviewees' comments that trust increases the perception of control, in other
words trust itself is a control method, which is also a prevailing view in the literature (see for example Cranor et al. 1999): “The disclosure of data to small, unknown companies nearly always raises doubts about where the information is used. If a piece of information seems irrelevant to the issue at hand, I will not usually disclose it.” On the other hand, trust compensates the need for control on the pursuit of privacy protection, especially on the Internet (e.g. Milne & Boza 1998, Olivero & Lunt 2004).

The different manifestations of the control speech are presented in Figure 1. The interviewees expressed the control perception of their personal information mainly through control methods, such as restricting the disclosure of personal data. The control perception was manifested also by the interviewees’ own initiative, by permission-based marketing, and by the trust in the company. The methods of control were applied especially to the disclosure of data, which was otherwise experienced to be poorly in control. Merely with the decision whether or not to disclose data the interviewees control their own data and their privacy. On the other hand, the interviewees did not perceive that they were in control over the storage or usage of personal data, not even when utilizing control methods.

![Figure 1. The manifestations of the control speech among the interviewees.](image)

In the following four sections we will deliberate on the four main classes we found when analyzing our interviews: (1) disclosure, (2) storage, (3) usage and (4) control methods of personal data.

### 3.2 Avoidance of disclosure of personal data to companies on the Internet is difficult

The interviewees experienced that it is difficult to avoid the disclosure of personal data on the Internet since personal data is collected everywhere, even when the consumers do not notice it. In the interviewees’ opinion, companies can too easily access the personal data of consumers. Additionally, there are more and more situations in which information must be disclosed, even though one would not want to. According to several interviewees, the companies in general collect too many personal data. Similar results were obtained also in a study by Statistics Finland (Muttilainen 2007), in which 28 per cent of the respondents experienced having given too much information of themselves. Alike thoughts were presented already in the 1970's in opinion polls (see e.g. Katz & Tassone 1990). Thus, the attitudes seem to have remained comparable in spite of the increase in the digital data collection.

The interviewees often experienced that the data collection is not always necessary. “I understand the objectives of companies to get the best possible picture and profile from the customers, but there should be sensible limits to it. One could think that the information on what are my hobbies or which country I have last visited would hardly be useful to anyone.” The interviewees' views are supported by numerous studies (e.g. Sayre & Horne 2000, Dommeyer & Gross 2003, Chellapa & Sin 2005, White et al. 2008), according to which consumers want to gain more in exchange for the disclosure of their personal information. Furthermore, unnecessary inquiries are regarded as time-consuming and
irritating. In our study, the perceived purposelessness of the data collection was the most common reason for not disclosing personal information.

The data registers were regarded as problematic: a few interviewees felt that the storing of personal data in registers is uncomfortable. Furthermore, some felt it difficult to remember afterwards what information they had given and to whom. The interviewees commented that they did not know any more what the companies knew about them, or whether this information was up-to-date. They hoped for the possibility to access their own data and also to edit or delete it, when wanted. Thus, the interviewees were not particularly aware of their own rights (e.g. to inspect personal data in registers), or the exercise of the rights was regarded as difficult in practice. On the other hand, the online data collection was regarded as a more reliable method of acquiring data than other means, as the data did not disappear or change when processed (e.g. when using traditional forms). In the opinion of a few interviewees, the privacy protection requires also the consumers’ own activity and carefulness: “Firms get a lot of data easily; it is given without thinking about it too much.” On the other hand, they admitted that they are sometimes careless themselves and disclose data without consideration.

Several interviewees experienced that a lot of personal data is collected without telling it to the consumers, of which they disapproved: “In a way I perceive that I am in control over the collection of information. However, the information can be gathered in a way that one does not notice it, one does not even always understand how the collection of information on something might be significant ...”. In this comment, the possibility of disclosing information accidentally emerged; consumers do not always understand the consequences of information collection. Also the risk of information acquisition by combining data from several registers was mentioned. Furthermore, the interviewees were worried about data collection that is enabled by recording and analysing click-stream data, or using other methods. In general, the interviewees felt uncomfortable being observed.

The majority of the interviewees (46 of 57 comments) were not sufficiently aware of the collection of personal data online or the usage of the information when asked directly. Especially the purpose of use of personal information puzzled most interviewees. It was also challenging to stay ahead of the technological development; for example cookies were experienced as threatening. The reasons for the unawareness were both consumer-driven (e.g. lack of interest and passiveness) and company-driven (e.g. the difficulty of getting information when one needs it).

Finally, several interviewees commented that one is occasionally compelled to give information to service providers. Companies might request information that is essential for the functionality of the service (e.g. alert services), and then the data collection was regarded as justified. Often the interviewees, however, experienced that they had to give information of lesser importance or of sensitive nature to be able to conduct transactions, even if they would not have wanted that: “If I took the initiative myself, then I would react neutrally to the fact that a company needs information e.g. for granting a loan decision. If, on the other hand, the company has phished my contact information, or if I am otherwise unable to proceed on a company’s website without answering all required questions, that is really irritating. I don’t like it either if a really long form with many compulsory fields must be filled in (especially when asking home address, social security number, children’s names etc.) in order to be able to register as a user to some service.” In this example the usability of the answer form (e.g. the plenitude of compulsory questions) has an effect, too. However, the interviewees said that they are disclosing their personal data to companies even when compelled, but still with certain terms only.

3.3 Controlling third party access to stored own data is challenging

With regard to the control over personal information the interviewees were worried over the collection of data and what happened to the data after it. The majority of these instances were related to a third party’s access to own data when it already was in the possession of the company. The worries included the forwarding of data to external parties without permission, the uncontrolled spreading of personal data online, and to whom one’s personal information eventually ends up on the Internet. Some of the observations made were connected to data security issues. Firstly, the interviewees were uncertain
about where their data eventually ends up after its disclosure, and whether some third party can have access to this data: “I do not [perceive that I am in control over the collection and utilization of my personal information on the Internet]. Every now and then I fill in for example various surveys, but I always have to think where my information eventually goes.”

The interviewees were especially worried that their information might end up ‘in the wrong hands’, for example to conmen. They hoped that the companies informed them better of who has access to their data; they wished more openness and honesty. On the other hand, the interviewees did not necessarily trust the information provided by companies even though the mere existence of data protection statements already aroused the confidence in a few interviewees (see discussion also in Kobsa 2007). The data protection statement is a report given by the registrar about how it acts with the personal data it has acquired in its possession, what are the data protection rights of the person being registered, and how the rights are implemented in practice (Data Protection Ombudsman 2008).

Several interviewees also experienced that personal data spreads easily on the Internet. When own information is disclosed online, it is perceived to spread everywhere without any means of controlling it: “Frighteningly poorly [I perceive that I am in control over the collection and utilization of my information on the Internet]. One can only be amazed by the ways with which information travels online, and suddenly, for example, your e-mail is flooding with spam.”. Whereas the uncontrolled spreading of information was connected especially to the Internet, the disclosure of personal data or selling that systematically to third parties, usually for advertising purposes, was seen as a problem also outside the Internet realm: “Yes, most of the time [I perceive that I am in control over the collection and utilization of my personal information on the Internet]. I always consider carefully where I disclose my information. The problem is naturally the fact that companies can pass on or sell the information, if, e.g., the forwarding of information for marketing purposes has not been forbidden.”

The deliberate forwarding of personal data was strongly disapproved, especially in the financial sector. In general, the interviewees trusted the banking sector actors with regard to the use of personal data. This is in line with the latest Eurobarometer (2008), according to which 90% of the Finns trust banks and financial institutions in this respect, which is 24% more than the EU-27 average. However, the recent value network developments (e.g. increased co-operation between banks, insurance companies and real estate agencies) caused some hesitation. Finally, what comes to third parties’ possibilities to get a hold on personal data, several interviewees expressed their concern about data security: “In e-commerce it is handy to use your credit card. However, the problem is the risk that the credit card number can be stolen at the data transmit stage. The site must be reliable.” The interviewees were willing to disclose sensitive personal data only in secured connections. In spite of the cautious attitudes many interviewees said that they had given their credit card number on the Internet also in unsecured connections, e.g. when booking a hotel reservation. However, the events had troubled them afterwards. Regarding information security the media appears to have an effect on the consumers; a few interviewees mentioned cases that had been reported recently by the media, for example fake websites with which information is phished for criminal purposes.

3.4 Control over the usage of personal information is difficult

The interviewees discussed in several occasions the usage of their personal data. When making a decision whether to disclose information or not, they felt uncertainty about the purposes of their personal data usage, and considered it difficult to control the usage. The interviewees did not perceive being sufficiently aware of how and to which purposes personal information is utilized: “I do not [perceive that I am in control over the collection and utilization of my personal information on the Internet]. I disclose information very randomly, depending on which sites I happen to end up. I can of course decide where and what kind of data to provide, but later on I am not able to remember or identify what I have given. Especially the usage of information can remain totally murky.” In these uncertain situations, a few interviewees were unwilling to disclose their personal data to companies. They wanted to know clearly how their information would be used. On the other hand, a few admitted
that they were too unconcerned regarding the usage of their personal data. However, several interviewees thought that personal data is utilized without informing the consumers properly, and even though companies would give information about how they use consumers’ personal data, most consumers would not believe that the companies would do as they say.

Secondly, the interviewees felt that it is difficult to control the usage of personal data; when information is disclosed, the control over it is easily lost. In any case alertness is required from the consumer with the control over personal information. Thus, in practice, the possibility to control personal information is connected only to the disclosure of data: “It is quite impossible to control fully [the collection and utilization of my personal information on the Internet] except for the information you provide. On the other hand the information is obscured into the data mass so it will not necessarily hurt you personally.” Once again, the control perception of the consumers’ personal information and their actual rights do not seem to meet each other. Thus, the consumers should be more informed of their own data protection rights.

3.5 Methods of control regarding personal data

The control methods of personal data utilized by consumers can be divided into active (invoking options) and passive (avoiding practices) means (see e.g. Givens 1997). The active means require deliberate actions from consumers, e.g. signing up to a prohibition register of direct mail advertising. Conversely, the passive means denote that something is not done, that is, restricting the disclosure of personal data in different ways, for example, by avoiding shopping on the Internet altogether.

The methods of control mentioned by the interviewees were limited mainly to the passive means, which, according to previous studies (e.g. Dommeyer & Gross 2003) are also best known by the customers. A critical attitude towards information disclosure was mentioned as the main passive means. In addition, awareness of companies’ actions was emphasized. Based on the consumers’ ability to distinguish reliable marketers from suspicious ones, consumers can decide to whom personal information is disclosed and what services are used. Many interviewees actually restricted the disclosure of their information in some way. A part of the interviewees avoided the disclosure of certain information (e.g. own name), whereas a few provided as few information as possible: “The less information you disclose, the better you are able to control that. But the information that I already have disclosed, I am not able to control that, but I trust that the firms use it confidentially.” Some of the interviewees conditionally disclosed their information in some ways, for example by forbidding the forwarding of their personal data to other parties. The majority of the interviewees disclosed information only to companies that they considered reliable.

Regarding the active means, only lying and making up false details were mentioned: “With surveys to some extent [I perceive that I am in control over the collection and utilization of my personal data]. Sometimes a service which I want or must use is asking to fill in compulsory information that I would not like to give. Within certain limits you may give something else than your truthful information.” This was expected to some extent as most control methods that require actions are in fact connected to the control over own environment rather than that of own information. For example, the marketing prohibition registers protect the consumer from direct mail advertising, not from the processing of personal data. On the other hand, not a single interviewee mentioned using a commercial e-mail account (Hotmail etc.) for the control over own information even though their usage is quite common.

4 DISCUSSION AND CONCLUSIONS

This study contributes to the area of privacy protection by offering a comprehensive view on how consumers perceive controlling their personal information disclosure on the Internet regarding marketing and CRM contexts. This topic has not been previously studied separately, but rather in connection with related research questions; thus our study produces valuable knowledge to the discussion of consumers’ control perceptions in the online context. Our categorization of the control
perceptions experienced by online consumers to four separate themes - or problem situations and control methods - supplements especially the work of Goodwin (1991) regarding the control over personal information disclosure.

The findings of our study show that the consumers perceive the control over their information relating to data collection as a part of a wider whole; when considering whether to disclose information, most consumers think about its sensible and safe storage and purpose of use. On the other hand, it is clearly difficult to control, or even grasp, the utilization of personal information in its entirety. Consumers look at the utilization of personal information from different viewpoints: one emphasizes data and channel security, and the other ponders on the reliability of the company that utilizes the personal information. Even though the interviewees said that they were not in control over their personal information, they, however, demonstrated the perception of control by utilizing control methods suitable for themselves. The lack of control, however, did not necessarily disturb all the interviewees. In future research, control should thus be examined more thoroughly paying attention to the different approaches employed by consumers.

One of the central themes of the interviewees’ control speech was trust, which can also be considered on the basis of our data as one method of control, or as a factor that increases the perception of control. Alternatively, trust can be regarded as a compensating factor for the need of control, as preliminarily presented by Olivero and Lunt (2004). According to our qualitative data, both assumptions are supported. On one hand, several interviewees mentioned trust as a prerequisite for disclosing personal information, especially sensitive one, thus referring to a method of control. On the other hand, the interviewees who did not generally trust the Internet, however, disclosed their information using other methods of control, e.g. restricting the disclosure of information with certain conditions. The significance of trust in the control over the consumers’ personal information is definitely an important topic for further research.

As for the limitations of this research, we should note that there might be differences between the stated and actual behaviour regarding online information disclosure (see discussion in Kobsa 2007 and Jensen et al. 2005). However, we believe that a survey would be more prone to this distortion than anonymous focus group discussions, as many of our interviewees admitted that they had disclosed information in situations where they in principle would not state to do that. Furthermore, the focus group interviews were carried out only in Finland, and with relatively homogenous sample of participants (online banking customers of age 21-49) which might lower the generalizability of our results. In Finland, online banking is the third most popular application on the Internet, thus almost three out of four Internet users are online banking customers (Statistics Finland 2008). According to the latest Eurobarometer (2008) survey, Finns (as well as other Nordic citizens in general) have higher confidence than the 27 EU countries on average to all market actors and public authorities regarding the usage of their personal information in an appropriate way. Only the market and opinion research companies are trusted equally little in Finland as in the other EU countries on average (33%). Thus, future research might examine the perception of control over online information disclosure in some other parts of the Europe or in the USA, and compare the results to the present study. Moreover, additional research is warranted that studies control in relation to personal information that is provided in consumer-to-consumer relationships (e.g. in social networks or via google mail), as companies exploit increasingly also this information in their marketing efforts.

References
Boddy, C. (2005) A rose by any other name may smell as sweet but “group discussion” is not another name for a “focus group” nor should it b. Qualitative Market Research, 8 (3), 248-255.


File-sharing – a threat to Intellectual Property Rights, or is the music industry just taking us for a spin?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0406.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
FILE-SHARING – A THREAT TO INTELLECTUAL PROPERTY RIGHTS, OR IS THE MUSIC INDUSTRY JUST TAKING US FOR A SPIN?

Andersson, Bo, Lund University, School of Economics and Management, Department of Informatics, Ole Römers väg 6, 223 63 Lund, Sweden, bo.andersson@ics.lu.se

Lahtinen, Markus, School of Economics and Management, The Institute of Economic Research, Lund University, P.O.Box 7080, SE-220 07 Lund, Sweden, markus.lahtinen@ics.lu.se

Pierce, Paul, School of Economics and Management, The Institute of Economic Research, Lund University, P.O.Box 7080, SE-220 07 Lund, Sweden, paul.pierce@ics.lu.se

Abstract

File-sharing has become synonym with the “digital economy” where large music conglomerates as well as certain artist voice strong concern over the impact on their bottom line. This research analyzes the music industry, which has been heavily impacted by a major technological shift i.e. the invention, and rollout, of the Internet. We look to the technological shift that has enabled the downloading phenomenon, as well as analyze the uniqueness of the music industries situation. By comparing the book publishing industry as well as the software industry, which are arguably also influenced by intellectual property rights and plagiarism, we try to find similarities as well as dissimilarities with the music industry. We find that the music industry has used alliances as well as Mergers and Acquisitions in order to consolidate their positions in an attempt to slow down change. There is no consensus on the exact extent of ill effects of filesharing. We point to an unwillingness to achieve convergence of purpose between the IT-community and the much of the music industry. Finally we point to the historical fact that consumers always get what they want in the end, which should indicate a need to find a viable e-commerce solution.

Keywords: Music Industry, piracy, Intellectual property rights, alliances, Internet
1 INTRODUCTION

A lot of public attention has been focused at the economic impact on the music and film industry by filesharing. If we have not tried it out ourselves, most of us know of someone that has tried downloading. What has enabled this change is the more profound technological change that the spread of the Internet has brought about.

Drucker (1994) discusses how a company’s business is based on a number of assumptions. These assumptions relate to market, customer, technology and strengths and weaknesses in them. The assumptions that companies have on the three topics must correlate with reality as well as with what the companies’ core business is. Drucker (ibid.) argues that the changing external environment forces the company to always question its business assumptions and this is something that we argue that the music industry has failed to do.

One way of understanding what Drucker refers to as the external environment (macro-environment) is what in market analysis terms are called ‘PEST’-factors; referring to Political, Economic, Social and Technological (cf. Armstrong, 2006), the Internet is an example of such a technological change.

The rise of the ‘digital economy’ or ‘new economy’ came about during the 1990’s. Scholars like Tapscott (1996), Brynjolfsson & Kahin (2000) have used the term to denote an era where IT is assumed to have revolutionary impact on society. Rayport & Sivokla (1995) suggested that IT is an enabler to endless replication of transactions. Tapscott (1996) suggest that in the digital economy old rules, norms, and customs will no longer apply and this will cause both relationships as well as economic activity to change and in some cases be even inappropriate.

These are all ‘voices’ from the past decade; today we know differently. The dot-com era, in most cases, did not deliver the share-holder value promised, several companies went bankrupt, the much often referred ‘easiness’ of the Internet actually gave rise to challenges not projected in the past. But at the same time, there has been an ever-increasing spread of Internet nodes, which in turn gives rise to an exponential increase in interconnectivity across the world.

According to the Internet World Stats (2007) there were 1,244 Billion users of the Internet per September 30, 2007; which results in a world penetration of nearly 20%. The Internet penetration displays the highest numbers in North America (70%), Australia (55%) and Europe (42%). Hence, downloading is by and large a phenomenon represented by the demand of the Western-oriented societies, due to its relatively large penetration.

The possibility for connectivity to a great number of people as a consequence of the compression of geographical distances due to the Internet, the speed of change as well as the need for change has also increased because of the flexible nature of the new system.

Returning now to the music industry we are interested in how this industry is changing compared to other, but similar, industries. Can it be the case that other industries are experiencing the same kind of claimed sales drop as the music industry, or put more as a general question: what is unique with the music industry?

1.1 Purpose

The purpose of the paper is as follows:

*To describe and analyze an industry affected by modern piracy mediated by a major technological shift in form of Internet technology.*

The industry at hand is the music industry. And in order to reach this purpose, we are interested in answering the following questions:
• What is the technological shift that has enabled the piracy phenomenon?
• What is unique with what is happening in the music industry?
• Are similar industries experiencing the same kind of transformation?
• Analyze the differences by contrasting the music industry with these references industries.
• And finally, what are the future implications for the music industry?

1.2 Structure of the paper

The approach to the topic is critical realism and a mix of qualitative and quantitative methods. In section 2 the mediating technology change and related phenomena is described. In section 3 the analysis of the music industry presented. Followed by section 4 where possible future directions for the music industry are elaborated upon. Finally, section 5 presents our concluding remarks.

2 TECHNOLOGY CHANGE: INTERNET

The Internet as we know it is a worldwide network, that is built on, or by, a multitude of interconnected computer networks that are for the most part publicly accessible. All the computers on this network share the same common language of communication, namely Internet Protocol (IP). The different networks are built up from businesses, academia, government and other players, who together move information such as mail, chat, music, documents, file transfers and interconnected web pages, which all together form the World Wide Web (www). The Internet Protocol can be said to be the backbone, or rulebook, of the WWW. The authors of this article are especially interested in if, and how, file-transfers or more accurately file sharing has affected the music industry.

The company most often “accused” of starting the file-sharing craze is Napster. This was a Peer-to-Peer (P2P)\(^1\) service, that in reality paved the way for other similar programs such as Kazaa, Morpheus, DC++, BitTorrent and many more that are still operating today even though the original Napster was closed down as early as July of 2001.

2.1 The music industry

The music industry as such is a $35 billion industry world wide (IFPI 2004a). Music plays, and has played, an important part in our everyday life. As a reference, the software industry was worth $380 billion according to Software Magazine (2006). Yet, music matters greatly in our everyday life. Music is available in your phone, in most stores, as background when you see a movie, in the car, as a leisure activity and so forth.

Currently there are four big players – in the music industry, representing between 70-75 % of the total market. Sony BMG is the result from a merger between the music division of Sony and BMG in 2005\(^2\). Sony owns 50% of the shares and Bertelsmann Group the other 50%. The three other players are Universal Music Group (owned by French Vivendi), Warner and EMI.

In an attempt to get a more detailed and exact understanding of the music industry sales, the annual reports from the five companies were examined. The purpose of the comparison (table 1) is to see how

---

\(^1\) P2P file sharing is distinct from file trading in that downloading files from a P2P network does not require uploading, although some networks either provide incentives for uploading such as credits or forcing the sharing of files being currently downloaded.

\(^2\) References to the big five, refers to the pre-SONY BMG period.
sales have been affected in each of the individual cases. Therefore, the sales numbers have not been adjusted to a common currency.

In four out of five cases the sales have dropped. Comments from senior management from these companies suggest that some of the lost sales of physical music have been replaced by digital sales, but not to the extent that it makes up for the lost sales on the physical side. Consequently, a more general remark about the industry is that the market as such is shrinking.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Sales comparison – The Big Four/Five</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2000</td>
</tr>
<tr>
<td>Company</td>
<td></td>
</tr>
<tr>
<td>Sony Music</td>
<td>707 ¥</td>
</tr>
<tr>
<td>EMI (Rec. Music)</td>
<td>2 033 £</td>
</tr>
<tr>
<td>UMG (Music)</td>
<td>N/A</td>
</tr>
<tr>
<td>BMG (Music division)</td>
<td>N/A</td>
</tr>
<tr>
<td>Warner (Rec. Music)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The prime reason for dropped sales are new distribution forms using the Internet and mobile phones. The Internet functions as an enabler as well as a threat simultaneously. Digital pirates constitute a threat, but also industrial piracy, e.g. counterfeiting where illegal physical copies of CDs are made; often in smaller industrial ‘labs’. This problem is not very widespread in the Western Countries, but more often reported from Latin America and Asia.

Looking then into more detail to some of the remarks from top management of the largest companies:

*Piracy, both organized and casual, continues to take a large proportion of potential earnings from the pockets of our shareholders and artists. Protecting our intellectual property remains a top priority for EMI.* (Annual Report 2006, EMI, p. 3)

*The industry and UMG are increasing their anti-piracy activities with a multipronged approach focusing on legal action, including participation in industry legislative efforts, public relations and education, and technical countermeasures while offering consumers new products and services.* (Annual Report 2006, Universal Music Group, p. 46)

*SMEI is also continuing its work with the Recording Industry Association of America (RIAA) and the International Federation of the Phonographic Industry (IFPI) on the industry’s anti-piracy efforts, and has led the industry in the development of an innovative initiative known as the Campus Action Network (CAN). CAN is designed to help colleges and universities across the country launch legal online music services and combat illegal filesharing on campus.* (Annual Report 2004, Sony, p. 50)

Judging from the comments in the annual reports, file-sharing is a top-management concern. It is also a kind of battle that needs to be fought on several arenas, e.g. through education as well as by lobbying through the different national legal systems.

However, looking purely at the music industry it can be concluded that piracy has affected their business to a significant extent, which confirms the conclusions by Peitz and Waelbrock (2006), Liebowitz (2004) and the music industry as well. Consequently, it makes sense to say that the music
industry reacts to lost sales by pursuing both the offenders and also changes the way they make their money, i.e. re-configuring their business model.

2.2 Drivers for alliance building

The software industry is potentially threatened by the risk of piracy related to use of software that is not licensed to the user. Being a digital kind of product it lends itself to be easily distributed through the Internet. The interesting case of the book publishing industry is that it reassembles the music industry in certain important aspects. One of the important features of the book publishing industry is the ownership, i.e. TimeWarner and Bertelsmann Group both stand behind two of the world’s largest publishing houses; DC Comics and Random House.

2.2.1 The book publishing industry

According to Hunt (2007) one of the first recorded uses of piracy on intellectual property rights was in 1701 when Daniel Dafoe published the poem “The true born Englishman” which was promptly copied and sold in the streets of London without Dafoe’s permission. Although the first copyright act appears first in 1709 this could be regarded as a copyright infringement.

The interesting thing about this was that Dafoe himself was not very upset by the plagiarism, since the wide circulation of his work had made him famous and enabled him to sell more of his coming works. In this case there were definitive gains from the circulation of the material.

Another example where the gains have not been so evident is the Harry Potter books. According to March (2007) legitimate book sales in India where down an estimated 50% due to pirate prints. In China there where “new” sequels being released in 2002 where the content was made up from other books where the character names had been swapped to match with Harry Potter characters. J.K. Rowlings lawyers have pursued all known cases of copyright infringements. The interesting aspect here is that the publishing industry probably has tried to stem the flow of piracy longer than any other business, where new technology has continually threatened the industry of more than 300 years and yet they are still here even though there are still pirate prints around.

Having gone through the annual reports dating back to 2002 for the book publishing industry for the 4 biggest book publishers of English-speaking books, i.e. HarperCollins, Penguin, Random House and Simon & Schuster, we can conclude that the none of these annual reports comments upon being effected by illegal downloading through the Internet; nor are there any reports on dropped sales as a consequence of file-sharing.

2.2.2 The software industry

In 2004 the director of research at IDC John Gantz claimed (in Locklear, 2004) that the piracy loss figures that are reported by the software industry are highly misleading. According to their study, done on the behest of the “Business Software Alliance” (BSA), only one out of ten unauthorized copies of software is really a lost sale. The correct term to be used when discussing loss figures should be “retail value of pirate software” according to John Gantz, since many of the copies out there are in developing nations where the users just can not afford to buy it.

According to Davis (2007) this is something that was backed by Romania’s president Traian Basescu, who claimed that “piracy rocks” no later than in February of 2007. He went on to thank Bill Gates for opening a new global technology support centre in Bucharest, which ironically was possible due in large measure to the widespread piracy in the country, which has enabled the employees to be familiar with the wares they need to give support on. According to the president “Piracy helped the young generation discover computers. It set off the development of the IT industry in Romania.”
Hunt (2007) points out that Jeff Raikes, president of Microsoft’s business group has admitted to the fact that pirated Windows is easily the best marketing tool they have ever had and was quoted to saying: “If they’re going to pirate somebody, we want it to be us rather than somebody else.”

Similarly for the software industry, the annual reports of Microsoft and Adobe were examined back to 2002. Microsoft’s operating system and its Office suite are common products subject to illegal downloading. Adobe is the software house that develops some of the most important tools for the graphic- and image industry and the suite of applications offered by Adobe are often subject to downloading and copyright breaches.

We may not be able to protect our intellectual property rights against piracy, infringement of our patents by third parties, or declining legal protection for intellectual property. We defend our intellectual property rights and combat unlicensed copying and use of software and intellectual property rights through a variety of techniques. (Microsoft, Annual Report 2006, p. 15)

Although we defend our intellectual property rights and combat unlicensed copying and use of software and intellectual property rights through a variety of techniques, preventing unauthorized use or infringement of our rights is inherently difficult. (Adobe, Annual Report 2006, p. 34)

Piracy is seen, both by Microsoft and Adobe, as just one type of threat among many others, and also a threat that they acknowledge being difficult to fight. That is a different view than the one conveyed by the music industry. On the other hand and as indicated by table 2, neither Microsoft nor Adobe has suffered any drops in their sales since 2002.

<table>
<thead>
<tr>
<th>Year/Company</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adobe</td>
<td>1,165</td>
<td>1,295</td>
<td>1,667</td>
<td>1,966</td>
<td>2,575</td>
<td>3,158</td>
</tr>
<tr>
<td>Microsoft</td>
<td>28,365</td>
<td>32,187</td>
<td>36,835</td>
<td>39,788</td>
<td>44,282</td>
<td>51,122</td>
</tr>
</tbody>
</table>

2.3 Intermediary conclusion

The music industry has experienced a considerable drop in sales

Addressing the piracy and illegal downloading as the main threat has been a central issue to the top management of the music industry when communicating with the shareholders.

The book publishing industry does not actively recognize downloading as a strategic threat to their operations and sales.

The software industry is affected by piracy of different kinds. However, a more complex pattern of more general business threats exist in that industry.

3 ANALYSIS OF THE MUSIC INDUSTRY

As indicated above, it can be stated that the music industry informs their share-holders about the effects of illegal downloading. Neither the book publishing industry nor the software industry makes such a strong case about piracy to their shareholders. How are we to make sense out of the differences observed between the different industries?
Departing from the five forces-model suggested by Porter (1980) a division is made about several competing forces that needs to be taken into consideration when analyzing an industry; the supplier’s bargaining power, bargaining power of the customers, internal competition, threat of substitute products and, finally, the threat of new entrants.

Looking into more details to the category of threat from substitute products where switching costs become a central theme; the possible migration from one product to another in this case is relatively low when it comes to music due to the ‘low-cognitive’ nature of the product. Music and films are ‘consumed’ in a less complex manner than for example books and software, which creates a situation of low switching costs as regards substitute products. However, music works complementary to movies, mobile phones, MP3-players etc. and that is where the future revenue streams exist in the future. The quotation from the EMI Annual Report of 2006 confirms this view:

*We continued to broaden our digital distribution channels globally by entering into agreements with partners who will make our music available on their platforms. During the year, we entered into partnerships including: a multi-territory deal with Apple; regional partnerships with MTV, Yahoo, Last.fm and Sony; and national agreements with Amazon in the US, Baidu in China, Napster in Germany and Playlouder and BT Vision in the UK.* (EMI, Annual Report 2006, pp. 18-19)

Although piracy holds a threat to the software industry as well as the book publishing industry the nature of the products is somewhat different, e.g. a piece of software or book has an overall longer ‘consumption durability’ that makes them less vulnerable to the effects of piracy. There are exceptions of course but a large amount of today’s music have short ‘consumption durability’. This is has worked against the music industry in the sense that, for instance, music is consumed once or many times over the radio, where the music is free to the consumer and on top of this it is legal to record the radio music and listen to again i.e. there is already a low threshold towards using the music freely.

3.1 Bargaining power of suppliers

Assuming that artists are as numerous and fragmented in the music industry as in the book publishing industry; the software industry is somewhat different in the sense that in what is referred to as the ‘industry’ the companies are not only intermediaries between the artist and the consumers, but rather both the artist and the intermediary at the same time. Hence, the upstream threat visible in the music industry, where artist may take over responsibilities that the music industry served in the past, is not present in the same way in the software industry. Although a reservation here should be made about conflicts with patent holders to base technology used in assembled end-user software or technology. The typical large software house controls more of the value-creating process.

The music industry has lived strongly on the asymmetrical power relationship that exists between the industry and the music creators. One could even claim that it has a strong institutional character where bands artists long to be ‘signed’ to a large record company or label. Having been the cash cow for decades, the relationship is being questioned by the music creators to a larger extent.

3.2 Bargaining power of customers

In the case of the industry software, end-users are often corporations and consumers. The existence of a clear corporate customer works to the software industry’s benefit. Copyright protection can be more easily upheld in the case of a corporate counterpart. A more detailed documentation of the customer together with legal accounting principles works to the software industry’s benefit. The wide demand from both consumers and corporations also makes the end-user fragmented in many instances, which favors the industry.

The fact that there is a limited amount of available alternatives of operating systems and the bundling with the office suite of MS Office has rendered Microsoft a very strong bargaining position vis-à-vis its customers.
The music industry and the book publishing industry on the other hand have potentially much less loyal customer base than the software industry. Customers can easily change their consumption patterns by putting their next potential spending on a different kind of entertainment media than music, e.g. movies, Internet, TV or even illegal downloading.

3.3 Internal competition

All industries studied are subject to internal competition. But what is more interesting to see is the internal cooperation in the music industry, where forces are joined in order to limit illegal downloading.

Both the music publishing industry and the book publishing are somewhat similar; ‘Imprints’ is the corresponding term to the music industry’s ‘record labels’ and many of the larger publishing houses maintain a large number of imprints.

3.4 Threat of new entrants

Of the three industries studied the software industry is probably the one that has the lowest barriers of entry. One example of this is the open source-movement driven by Linux as a kind of prime example on an operating systems that threatens the market position of Microsoft Windows. Getting access to established sales channels of books and music is a difficult task compared to how easily software can be distributed through the Internet.

3.5 Summary of findings

Concluding the industry analysis, the following points can be made:

In the past, the music industry has thrived on the relationship between the music industry and the suppliers of music, i.e. the artists, but also on controlling the distribution forms to their customers.

New technology-enabled distribution forms have loosened the music industry’s control of the distribution and also enabled illegal downloading.

Artists see possibilities of by-passing the music distributors.

Music and songs per se, are easily consumed and subject to low switching costs to complementary means of entertainment in comparison to the publishing industry and the software industry.

4 FUTURE DIRECTIONS FOR THE MUSIC INDUSTRY

As discussed earlier, the use of Music is something that affects all of us whether or not we listen to this music of our iPod, a CD or sitting in the outback listening to a Didgeridoo. Because of this widespread use of music the industry as such is well adapted at alliance building as well as M&As, which have been used to form partnerships and alliances worldwide. Lorange and Roos et al. (1992) describe how alliances can be used offensively, by creating or accessing markets in order to define and/or set industry standards, or defensively in order to protect current market positions, gaining economies of scale or sharing financial risks.

It would seem as if the music industry to date has used alliances and M&A to consolidate their position on the market, being unwilling to change and adapt. Considering that it is very hard to research the total number of music sales today, since we have music in phones, computers, and shopping malls as well as on CDs and in Movies, it is hard to pay heed to talks about the industry being on the verge of collapse. Rather there is an issue of an unwillingness to change and adapt to the new rules of engagement that the Internet and electronic sales have brought to the negotiating table.
4.1 Other values from file-sharing?

Oberholtzer-Gee et al. (2007) debate the promotional value power of file-sharing, but other authors such as Hunt (2007), Locklear (2004), SIIA (2007), and Wilen (2007) also carry discussions in and around this area. The music industry’s point of view has been that all forms of file-sharing hurt their sales, but there has been proof of other views, most recently Radiohead released its new album only online and let their audience decide the price they are willing to pay. This is interesting since Radiohead was also one of the first bands to take a different view of file-sharing. In 2000 their album Kid A came into circulation on the internet before it was actually released as a CD, but contrary to regular beliefs and opinions of file-sharing the result for Radiohead was to hit the #1 spot on the billboard 200 sales chart its very first week in real circulation.

4.2 A changing landscape

It seems as if many artists as well as musicians of today actually embrace the Internet as a tool and platform they can use both to create and sell their work, and more importantly it is a very powerful channel for promotion. According to Madden (2004) the big issue is the actual impact of the free file-sharing and the effect this has on copyright issues.

There is no doubt that illegal downloading of files has become more common as the Internet diffusion has grown. The ongoing debate seems to be centered on copyright issues in the music industry and whether or not file-sharing has helped or hindered sales of CDs. Most studies – including ours – and academic articles show that file sharing affects CD sales in a negative way. Peitz & Waelbroeck (2006) argue that the possibility to download songs and text at any time from anywhere over the Internet on the one hand increases the value of the product, but on the other hand that rights management tools, e.g. DRM-protection, can restrict and even lock the use of the digital media.

This unwillingness to achieve convergence of purpose between much of the music industry and the IT community is maybe what lies at the heart of the problem. There are 4 conglomerates, who control and own much of the music labels. EMI, Sony BMG, Time Warner and Vivendi Universal controlled 84% of the US market and 79% of the European music market in 2002. The fact that some of them also have a diversified product portfolio seen from a corporate view, i.e. they control many other parts of the different verticals of the industry, such as hardware for playing different media, e.g. CD, DVD, projectors etc. creates a situation of almost total control. The typical scenario when you have an oligopoly is to have an unwillingness and resistance to change, or even to participate in any form of partnerships outside of the small group (INFO, 2006).

According to Dr Jo Bryce, University of Central Lancashire “The purchase of counterfeit goods or illegal downloading are seen as normal leisure practices” In the same article by Hermedia (2005) “Teenagers are being tactical spenders,” said Dr Bryce. “The money saved lets them spend more on mobile phones, going to the cinema or eating out.” This is an interesting observation, since it would insinuate that the money that is not taken in on selling an actual record is taken in from movies or by downloading ring-tones to the mobile phone both of which are markets that are heavily controlled by many of the same players that control the music scene.

5 CONCLUDING REMARKS

It would seem as if the “pirates” are here to stay and the best way to handle this is to embrace the change. History has shown that the consumers always get what they want in the end. Consider the upheaval the film industry had towards VCRs when they first came to market or the music industries cry when cassette tapes where introduced. Over the years we have seen a number of new technologies being introduced in the music industry as well as the movie industry and most of the time there have been problems and downturn in sales of the current predominant media channel. It would be
interesting to do a study of the dips in sales after external events affect the sales of an industry, i.e. how were sales affected after the Disco era was over? How were LP sales affected by CDs, how were CDs affected by DVDs and maybe mp3s? How was cinema attendance affected by Cable TV? The list goes on and on.

We arrive back to the statement of complexity. The new market is extremely complex where the big musical giants are also the big movie giants, but also giants in producing hardware for both markets. When you are both vertically and horizontally integrated it is had to know where one market ends and another begins and the real overall loss or gain of the full market is something probably only the big four themselves can truthfully answer.

6 REFERENCES


Locklear, F. (2004) ‘IDC says piracy loss figure is misleading’ ARS Technica, the art of technology, July


Journal of Political Economy, vol. 115, no 1., The University of Chicago


Free Press, New York


Mcgraw-Hill

Reports:
IFPI 2004a
Annual Report 2006, EMI
Annual Report 2002-2007, Microsoft
Annual Report 2006, Universal Music Group/Vivendi
Annual Report 2004, Sony
THE ROLE OF BOUNDARY OBJECTS AND BOUNDARY SPANNING IN DATA WAREHOUSING – A RESEARCH-IN-PROGRESS REPORT

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0114.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Communities of Practice, Data Warehousing, Design Science, Behavioural theories</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
THE ROLE OF BOUNDARY OBJECTS AND BOUNDARY SPANNING IN DATA WAREHOUSING – A RESEARCH-IN-PROGRESS REPORT

Vranesic, Helena, Goethe University, Grüneburgplatz 1, RuW Building, 60323 Frankfurt, Germany, vranesic@wiwi.uni-frankfurt.de
Rosenkranz, Christoph, Goethe University, Grüneburgplatz 1, RuW Building, 60323 Frankfurt, Germany, rosenkranz@wiwi.uni-frankfurt.de

Abstract

Data warehouse projects bring together different communities of practice, with the primary objective of producing one body of information which is capable of comparative advantages in business analysis. Due to the number of involved communities and the complexity of their collaboration, data warehouse projects are costly. In this paper we give a closer look at communication problems on boundaries between participating data warehouse projects’ communities. Our analysis enlightens the potential relation between the early creation of language communities of the involved communities and lowering data warehouse project development costs. As today, there is hardly any methodology available for analyzing and aligning mutual understanding between data warehouse project participants. In this paper, we propose a data warehouse development scheme for project improvement based on our discussion as a first step in a design science project.

Keywords: Data warehouse, Boundary Objects, Community of Practice, Design Science.
1 INTRODUCTION

Since business environments and information technology (IT) are now changing at an extremely fast pace, management information systems (MIS) become a crucial prerequisite for the long-term efficiency of any organization. Well-informed actions (i.e., those based on true beliefs) are more likely to achieve desired ends, and information is valuable insofar as it helps individuals form true beliefs which, in turn, promote effective, goal achieving action (March and Smith 1995, p. 251). The central purpose of MIS is to provide information for the management activities carried out within an organization (Curtis and Cobham 2005, p. 27 f.). Thus MIS are systems that assist management in decision making and in forming true beliefs. This includes both, executive information systems (EIS) and decision support systems (DSS) that allow ad hoc queries, and analytical reporting systems. Those IS are among one of the most prominent subjects in IS research (Claver, Gonzalez and Llopis 2000, p. 183).

It is general consensus that state-of-the-art MIS are based on data warehouses (DWH) and online analytical processing (OLAP) (Codd, Codd and Salley 1993; Devlin 1997; Inmon 2005). DWHs support the management perspective on business processes technically, but their implementation is extremely complex and costly (Vassiliadis 2000). The complexity manifests itself in the high failure rate of projects and increasing development costs (BCS 2004; Kimball and Caserta 2004; SGI 2001). Therefore, a clean design and specification of the system is of great importance. Consequently, the involvement of the management users and management’s support are deemed key factors for MIS quality and system success (Wixom and Watson 2001, p. 35 f.). One of the major issues in the conceptual specification of management views relates to the bridging of the so-called “communication gap” that originates from differences in semantic understanding (e.g., homonyms and synonyms) (McDavid 1996). This gap exists between business units using a MIS and computer departments designing and implementing the MIS within an enterprise (Alter 2001; Peppard 2001). As Alter (2001) puts it: “Of course most IT professionals know more about computer hardware and software, but the communication gap is about the difficulty business and IT professionals have in establishing mutual understanding that helps them communicate in both directions about their views and concerns” (p. 46).

In DWH projects, IT professionals developing DWHs (DWH professionals) are confronted with communication gaps on boundaries towards two communities of practice (Wenger 1998): business experts in decision-making fields (BEDF), who define business requirements, and operative system professionals (OSP) in charge of data required for fulfilling those requirements. According to several authors (Brown and Duguid 2001, p. 209; Star and Griesemer 1989, p. 393; Wenger 1998, p. 105), there are two forms of boundary connections: (1) boundary objects and (2) brokering. Therefore, DWH professionals in these circumstances have brokering roles in closing the communication gap on both boundaries. Moreover, we argue that boundary objects mediate the creation of language communities (Kamlah and Lorenzen 1984, p. 57) that afterwards enable the transfer of knowledge across boundaries between communities of practice. We argue that the creation of a language community between DWH professionals and both BEDFs and OSPs in early stages of DWH development plays an important role for the success of the final product. We suggest that early creation of language communities lowers DWH development costs by helping DWH professionals to construct extraction, transform and load (ETL) processes faster and more correctly, and by synchronizing decision-makers’ requests with sources of data. Consequently, we propose that the development of IT artifacts should not start sooner than a language community is created. The fundamental question that arises from this argument is how we are able to determine whether the creation of a language community was successful or not.

The structure of the remainder of the paper is as follows. In the next section, we discuss the motivation for our research. Afterwards, we present our research objectives. Following this, we review related work and possible theoretical foundations. We present an overview on our research-in-progress project and summarize the findings and limitations so far, giving an outlook on further research.
2 MOTIVATION FOR RESEARCH

Because DWH projects often fail or significantly exceed budgets, existing research has concentrated on quantitative or qualitative analyses of success factors as well as contemporary best practices for building DWH (e.g., Herrmann and Melchert 2004; Hwang, Ku, Yen and Cheng 2004; Hwang and Xu 2007; Watson, Fuller and Ariyachandra 2004; Weir, Peng and Kerridge 2003). March’s and Hevner’s (2007) thorough literature research warns how links between data warehousing, strategic decision-making and evaluation are under-researched. They present four overarching objectives for DWH support of management decision-making processes and influences of those objectives on the successful use of decision support in business organizations: integration, implementation, intelligence and innovation.

Integration, the first objective, is concerned with the consolidation of data from disparate sources into one consistent body of data. Integration lies at the heart of decision-making tools, with content management as one of the biggest challenges of DWH designers (March and Hevner 2007, p. 1036). March and Hevner (2007) warn that effective integration decisions can only be reached by experts having intimate acquaintance with knowledge domains pertaining to business internally and to the context in which the business that the MIS was built for is immersed (March and Hevner 2007, p. 1041). Therefore a pre-sequence that prefigures an integration process should be DWH professionals’ affiliation with the community of individuals (BEDFs) that are familiar with the business logic and the business environment (internal and external business ontologies) and the community of individuals (OSPs) familiar with operational source data (internal technical ontologies).

The next objective, implementation, addresses DWH development alignment with existing business processes in the organization, overcoming the current role of DWH as a repository of historical data which addresses passed organizational transactions (March and Hevner 2007, p. 1038). This is only possible when DWH professionals understand the business logic and can introduce learned business logic into the DWH.

The third objective, intelligence, refers to a right fit between DWH users and different type s of tools (March and Hevner 2007, p. 1041). Once the fit is found users receive appropriate training and support from DWH professionals. Analytical and data mining tools provide interfaces between decision-makers and data in the DWH. The users as decision-makers use the data to generate meaningful information on which they act. The last objective, innovation, deals with change management, specifically how changes in DWHs propagate to MIS. The DWH professionals’ job is to design the DWH to be ready for change so that every change is transparent to decision-makers (March and Hevner 2007, p. 1041).

We argue that only when the first two objectives (integration and implementation) have been successfully realized, decision-making can use the DWH to engage into the objectives of intelligence and innovation. Therefore we suggest that research should focus on concepts and theories that help DWH professionals to successfully accomplish both objectives of integration and implementation. In this context, the concepts of convergence and divergences between communities of practice and information artifacts offer a useful starting point for further research.

Star et al. (2003) introduced the concept of convergence to describe processes of mutual constitution between communities of practice and information artifacts they use: “information artifacts undergird communities of practice, and communities of practice generate and depend on the same information resources” (p. 244). Basically, DWHs are shared information artifacts between OSPs and BEDFs. Given the term convergence, the alignment of practice with information artifact increases when use and practice fit design and access (Star et al. 2003, p. 224). Therefore, DWH professionals’ main concern should be to build a DWH that will satisfy both the BEDFs’ needs for information and OSPs’ needs for easy integration. The complex processes at the boundaries between the two communities of practice include the discussion of differences between community insiders and outsiders. Members of
one community are called insiders of that community, whereas all others are outsiders. The thin line between insiders and outsiders is not easy to draw (Star et al. 2003, p. 262).

What follows from this discussion is that shared IS such as DWHs are a kind of boundary object, and DWH professionals are brokers between multiple communities of practice (Pawlowski, Robey and Raven 2000, p. 332). Apart from IS as such, entities pertaining to IS (e.g., system documentation, user training materials, standards, policies) are also considered boundary objects that are shared by the set of communities (Pawlowski et al. 2000, p. 334). Pawlowski et al. (2000) pointed out a major problem IT professionals (such as DWH professionals) face while supporting IS shared by several communities of practice: they have to learn the practice of each community and have to become insiders of both communities (not only learning about these practices, but also entrée to practice-enough legality) (Pawlowski et al. 2000, p. 335). In this context, the process of brokering means to get members of the one community (insiders) to understand and accept changes in their work necessitated by changes in the shared system that are needed by the other community (outsiders).

Based on these suggestions, we propose a solution for lowering the development costs of DWH projects and for creating business value. Creating mutual understanding between DWH professionals and communities of practice involved in data warehousing (i.e., OSPs and BEDFs) is a precondition for building successful and cost-efficient DWH-based MIS that add business value. We describe “evolution of a shared IS” (the DWH-based MIS) as developing a boundary object between OSPs and BEDFs based on various artifacts. Moreover, these artifacts also exist as boundary objects between DWH professionals and OSPs, and between DWH professionals and BEDFs. We propose that investigating how these artifacts influence the creation of language communities between the involved communities of practice should yield interesting and useful new insights for research. We are especially interested in exploring which phenomena contribute to the coherence of this language community, and how a language community, once created, enables DWH professionals to overcome the previously mentioned problems during the development of decision-making systems.

3 RESEARCH OBJECTIVES AND QUESTIONS

To summarize, the importance of communication in DWH projects has been widely recognized and studied (e.g., Wixom and Watson 2001). Research on impact of communication on DWH project success is limited, but is growing (e.g., Behrmann and Räkers 2008; Gallivan and Keil 2003; Joshi, Sarker and Sarker 2007; Ko, Kirsch and King 2005). As yet, little is known about how communication in DWH projects is influenced by boundary objects. While previous studies on boundary objects have, for instance, dealt with their influence in traditional ISD projects (e.g., Sapsed and Salter 2004), there is a lack of valid, reliable measures for assessing the boundary objects used in communication in DWH projects. This research-in-progress paper focuses on the role of boundary objects for communication in DWH projects with the aim of understanding how boundary objects influence the success of such projects. There are four objectives to this research:

- Which boundary objects are used in DWH projects to connect different communities of practice?
- How do the boundary objects (both formal and informal) evolve during a DWH project?
- How does selection or availability of adequate boundary objects or/and brokers in early phase of requirements analysis affect DWH project success, and how do these boundary objects affect DWH project success from the perspective of the different communities of practice involved?
- Can the development of boundary objects and the process of brokering be regulated or influenced in order to achieve better project results?

This research hopes to contribute to previous work on boundary objects in IS development projects by investigating the evolution and use of boundary objects in DWH projects and by developing and validating artifacts for leveraging such boundary objects. We also hope to provide a practical contribution by developing a framework that project managers and project teams can use to measure and assess the effectiveness of specific boundary objects at various stages throughout a DWH project.
4 THEORETICAL FOUNDATIONS

Communities of practice are characterized as shared histories of learning (Wenger 1998, p. 86; Wenger, McDermott and Snyder 2002). Such histories are known to create discontinuities between the ones who participate in the community work and the ones who do not (Wenger 1998, p. 103). These discontinuities are also revealed in development of DWHs. We distinguish two major communities of practice which confront each other in DWH projects: (1) OSPs with knowledge of the legacy and source systems, and (2) BEDFs. OSPs are in charge of maintenance and further development of operational systems; whereas BEDF influence the course of actions the company needs to take in order to stay competitive in the environment. Connection between these two practices is made by introducing a third practice, DWH professionals who are responsible for the development of the DWH.

Wenger (1998) uses the term negotiation of meaning as “the process by which we experience the world and our engagement in it as meaningful” (p. 53). Thus, our personal meaning of the world around us is located in the process of “negotiation of meaning” which involves an interaction of two constituent processes, namely reification and participation (Wenger 1998, p. 52). Participation involves all kinds of relations. In social communities participation shapes our experience, and also shapes those communities. Trough the process of reification, aspects of human experience and practice are congealed into fixed forms and given the status of objects. Participation and reification form a duality that is inherent to human perception of meaning and hence reflected to the nature of human practice (Wenger 1998, p. 52).

The two constituent processes of participation and reification have to be in balance (Wenger 1998, p. 56). When one of the processes predominates, negotiation of meaning is negatively affected or disturbed. If the process of participation prevails, negotiation of meaning lacks real, physical arguments that could help participants coordinate their argumentation lines and solve the differences. If on the other hand, the process of reification predominates, very little interaction and experience exchange remains. In these conditions, reconstruction of meaning buried in the products of reification (reified objects), could be very difficult (Wenger 1998, p. 56). Participation and reification can both contribute to the discontinuity of boundaries (e.g., a specific jargon is understood only by participants of a community where it is used), but can also create continuities across boundaries. Thus, the process of reification produces objects that can move across boundaries (Wenger 1998, p. 105). Likewise, members of one community of practice can participate in multiple communities of practice at once and help to introduce those reified objects from one community of practice to the others, when such translation is needed. Therefore two forms of boundary connections exist, as introduced before: boundary objects and brokering (Brown and Duguid 2001, p. 209; Wenger 1998, p. 105).

Star and Griesemer (1989) define boundary objects as “both plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites” (p. 393). Boundary objects may be abstract or concrete, are weakly structured in common use, and become strongly structured in individual-site use. In different social contexts (i.e., communities of practice), they can have different meanings; however, their structure is common enough across contexts to make them recognizable by means of translation. Consequently, the creation and management of boundary objects is especially important in developing and maintaining coherence across intersecting social contexts (Star and Griesemer 1989, p. 393). In the context of DWH projects, the artifacts (e.g., shared documents, tools, business processes, objectives, schedules) exchanged between communities of practice (Brown and Duguid 2001, p. 209) can potentially become boundary objects if they should belong to (at least) two different practices. These artifacts then form a nexus of different perspectives (Figure 1), which then needs to be coordinated (Wenger 1998, pp. 107-108).
Pawlowski et al. (2000, p. 331) emphasize the role of legitimacy in the jobs of IT professionals such as DWH professionals that helps them to influence the development of practice, to mobilize attention and to address conflicting interests. DWH professionals have a coordinative role as brokers in DWH projects. They create participative connections between the two practices of OSPs and BEDFs:

- On the one hand, gathering requirements in DWH projects means not only gathering reports and documents created, for example, by the BEDF community (boundary objects), but also meeting with community’s representatives. Reports do represent and are related to some of the more tactical decision-making requirements (Kimball and Ross 1996, p. 15), but they are purely reifications of them. Only through interaction with members of the BEDF community, where those objects were reified, can a full understanding of the requirements be reached. These objects then form the nexus where different perspectives of BEDFs and DWH professionals through processes of participation align and hence form DWH professionals’ understanding of the objects. This suggests that communication between both practices is of paramount importance for successful DWH projects.

- On the other hand, in interaction with members of the OSP community, DWH professionals concentrate on the main process of DWH development: implementation of extraction, transformation and loading (ETL) of operational data with the goal of creating more appropriate data-sets matching the BEDFs’ requirements. Compared to the overall costs of DWH projects, the costs resulting from the ETL process have a major share (Kimball and Caserta 2004). Extraction of operational data implies knowledge of what operational data exactly needs to be extracted to meet the requirements and business demands, and from which operative system this data needs to be extracted. In this context, we suggest that communication between OSPs and DWH professionals is a profound prerequisite for aligning different perspectives of exchanged boundary objects (e.g., database schema, database tables, excel tables, text files etc.) between those two communities. Only in close cooperation with OSPs, DWH professionals can extract and interpret operational data so that it matches the requirements of the BEDFs.

We propose that DWH professionals generate a mutual understanding of artifacts on the nexus of perspectives (i.e., boundary objects) through communication with members of both OSPs’ and BEDFs’ communities. In accordance with Ågerfalk and Eriksson (2004), we argue that traditional research on requirements analysis has focused too much on the syntactic and semantic aspects of conceptual modeling languages, and too little on the pragmatics of the requirements process. But where Ågerfalk and Eriksson (2004) use speech act theory as a theoretical foundation, we suggest to employ a different theoretical foundation from philosophy of language. Language critique (Kamlah and Lorenzen 1967; Kamlah and Lorenzen 1984), a branch of constructivist philosophy known as the “Erlangen School”, provides useful insights and backup for our understanding of DWH projects, and has been successfully applied to IS research before (e.g., Becker, Niehaves and Pfeiffer 2008; Holten 2007; Holten, Dreiling and Becker 2005; Holten and Rosenkranz 2008; Rosenkranz and Holten 2007; Rosenkranz, Laumann and Holten 2009).

Kamlah and Lorenzen (1984, p. 47) argue that language as a system of signs promotes mutual understanding as a “know-how” held in common, the possession of a “language community”. A new term is introduced by explicit agreement between language users with respect to its usage (first agreement) and meaning (second agreement) (Kamlah and Lorenzen 1984, p. 57). This agreement leads to a relation of concept and term, and is shared by a language community as the knowledge of using this term. Accordingly, if members of a group of people communicate, and each has an aligned semantic and pragmatic dimension of a symbol (or term) in mind, then this group of people forms a
language community. All members have the same concept in mind if they are confronted with a symbol of the language and vice versa.

According to our understanding, boundary objects play a significant role in making language communities explicit: if members of communities of practice have a same concept in mind when confronted with a same artifact (boundary object), they belong to the same language community (Figure 2). It is difficult for non-members of a language community to understand the local artifacts in a proper way. An individual has to revise his or her understanding of the given language signs with respect to the ones of the language community in order to become its member, and thus gain the possibility of using the same terminology in discourse with the other community members (Holten 2007, p. 3). One can readily say that an alignment of a community’s terminology is essential for its viability.

We argue that DWH professionals have a unique and moderating role: since they simultaneously belong to both communities of practice, and hence to both language communities, they are aware of the different perspectives that OSPs and BEDFs have. Their job is to moderate and manage the mediating process of translation, coordination and alignment between the two perspectives. Therefore it is essential that DWH professionals have a free hand to influence the development of practice and to mobilize their attention to address conflicting interests (Pawlowski et al. 2000, p. 331). In order to achieve a successful implementation of the developed DWH, both involved communities of practice need to fully participate in the development process. Therefore, we suggest that the job of DWH professionals is to make the developed DWH to be accepted and actively used in both practices (brokering). This implies that OSPs know why they prepare source data for the DWH, and BEDFs to comprehend the benefit of the reports generated by the DWH solution. As a result of DWH projects, a DWH is constructed, which represents a new boundary object between OSPs and BEDFs as a new nexus where community’s perspectives meet and align trough the process of negotiation of meaning (Figure 3).

We argue that the brokering role of DWH professionals remains a significant success factor even when a DWH is finally implemented. Up until now, maintenance of DWHs often relies entirely on the
knowledge base these high-skilled professionals prepare in form of exhaustive documentation or as consultancy service they offer (Rükers and Rosenkranz 2008, p. 10).

5 RESEARCH METHODOLOGY

5.1 Design Science Research

We propose to adopt design science as a research framework, which is recommended for scientific studying in the field of IS when artificial, human-made phenomena, such as DWHs, are examined (March and Smith 1995, p. 253). Developed and evaluated IT artifacts that are useful and capable of solving research problems in the IS realm constitute the cornerstone of this approach (Gregor 2006, p. 62; Hevner, March, Park and Ram 2004, p. 78). That is, rather than producing general theoretical knowledge, we aim to create efficient IT artifacts by design science research. In contrast to behavioral science, design science research aims to provide four general outputs: (1) constructs, (2) models, (3) methods, and (4) instantiations (March and Smith 1995). Therefore we aim to develop and provide useful methods and instantiations of IT artifacts that DWH professionals can utilize for their role as boundary spanners and brokers. We argue that research has to support this role explicitly in order to support the integration objective. We plan to follow the design science research methodology as suggested by Vaishnavi and Kuechler (2008, pp. 19-22). In order for DWH professionals to acquire the necessary understanding of the involved communities of practice (OPS and BEDF), and thereby to develop cost-efficient DWHs which add business value, we propose the following DWH development scheme for DWH project improvement as a first step in our design science project (Figure 4).

![DWH Project Improvement Scheme](image)

Figure 4. DWH Project Improvement Scheme

Furthermore, the improvement scheme acts as a framework for our research. We envisage that project managers and project teams can use a more elaborate version of the scheme for identifying potential problem areas and for planning and coordinating their requirement analysis processes (see Figure 5 for a first tentative proposal). Initially, it is necessary to discover community’s representatives (BEDFs and OSPs). During the process called negotiation of meaning, BEDFs and OSPs together with DWH professionals discuss the meaning of BEDFs’ set of business requirements and OSPs’ operational data sources. We suggest that this process should be supported by a kind of Analysis Panel. Artifacts such as business documents, source system documentation, source system data or interview notes will be gathered and stored in an Input Data Storage. The Analysis Panel is by no means an ETL tool. Whereas ETL tools try to map source systems and DWHs unequivocally, we consider ambiguities in both requirements and sources as opportunities for negotiation and creation of new understanding. The Analysis Panel then serves as a foundation for engaging into negotiation of meanings between DWH professionals and the other stakeholders. As a result of this process, DWH professionals produce, vary and store a set of terms in the so-called Term Storage. Terms in the Term Storage represent the
explicit definition and meaning of the domain language that DWH professionals have developed while negotiating the meaning of the information artifacts with the help of the Analysis Panel (and stored in the Input Data Storage). The sketched feedback loops suggest how the process of negotiation affects all participants and changes their preliminary understanding of the discussed artifacts. Each time participants meet, the process of negotiation influences the DWH professionals’ set of terms. Consequently, the Term Storage should support versioning of each single term. When no novel versions are produced, the meaning of a term is stable and ready for usage (i.e., a language community is created when all necessary terms are stable). In later phases of a DWH project, these terms should enforce further business requirements and operational data source collection as well as the development of the final DWH.

Figure 5. **DWH Project Improvement Scheme – Detailed Version**

For example, consider the following scenario: company X is in charge of DWH development for an insurance company, in our case called company A. During the first phase of requirements gathering, companies’ representatives, DWH professionals from company X, BEDFs and OSPs from company A, meet on regular basis. Most of the meeting sessions include discussions over company A’s legacy systems specifications as well as company X’s DWH solution proposition. Those documents represent the foundation for negotiations of the meaning for concepts and terms involved (e.g., differences in the meaning of the terms “hospital” and “clinic”, that DWH professionals are unaware of, is clarified and both terms are afterwards stored in the Term Storage). Our proposed DWH improvement scheme could act as a framework that facilitates systematic investigation of hidden, weak and unexpected relationships between language terms from the perspective of visual exploration and analysis of discussed documents.

5.2 **Proposed Data Collection & Further Course of Project**

The goal of design science is utility. For evaluating a proposed artifact, its utility for solving specific problems must be assessed. In presenting the development of methods and instantiations used to support the presented improvement scheme, we will firstly explore and introduce theoretical constructs and express the relationships among constructs to propose “how things are” in DWH projects. By building a logical chain of arguments for our reasoning, we will discuss why artifacts that we propose and build during our project will provide a suitable basis for addressing those problems.

First, in extending our already thorough literature review, we will conduct a second review in the hope of finding existing and related approaches in other areas in order to detail and refine our improvement scheme. Second, we plan to conduct expert interviews with DWH professionals in DWH departments of software development companies. The interviews should revile a set of communication problems and their probable causes in DWH projects. We will then match the results (in form of cause-problem pairs) with our scheme for DWH project improvement and the suggestion that the early creation of language communities between involved communities of practice, using boundary objects as nexus of different perspectives, will lead to alignment between participants. Furthermore we will conduct an
additional round of interviews with a second group of experts in other software development companies to test if they experience the same set of problems as the first group. This should show if the perceived practical problems can be linked to the theoretical foundation and the suggested improvement scheme. Then we plan to generalize our project improvement scheme and test it against other shared IT projects in the companies, conducting surveys. Third, we plan to explore how specific IT artifacts (i.e., both methods and instantiations) can be used (1) to support the creation of boundary objects and the process of brokering in general, and (2) implement the proposed improvement scheme. This will include several proposals for IT artifacts, which will be based on a more detailed improvement scheme and aim to solve issues arising from this scheme.

By conducting field studies to evaluate the usefulness of our improvement scheme and the to-be developed IT artifacts, we hopefully will show significant improvements in language community creation, reduction of DWH development costs, and added business value. Since design science is characterized by an iterative search process, during these “micro evaluations”, minor redesigns will probably occur on several occasions, which is a common occurrence in design science research (Vaishnavi and Kuechler 2008, p. 25). Despite the problems inherent in the subjectivity of these approaches, we strongly believe that there is a need for field studies to examine the usability and actual usage of proposed IT artifacts.

5.3 Current Status of the Project

The project to date has reviewed the literature on boundary objects in many different domains. Literature relating to DWH projects has also been reviewed. At this point the motivation for the research is clear, the overall research objective has been identified and the research questions are defined. Currently, various possible artifacts that relate to the improvement scheme are being identified and critiqued for their suitability and applicability to this study. It is hoped that at least some of the constructs already detailed in the literature can be adopted, but the expectation is that some new constructs will need to be developed that relate specifically to boundary objects in DWH projects. Once the improvement scheme is defined and pilot tested, IT artifact construction will proceed.

6 CONCLUSION

In this paper we argue that the creation of a language community between DWH professionals and both BEDFs and OSPs in early stages of DWH development lowers DWH development costs by helping DWH professionals to construct ETL processes faster and more correctly, and by synchronizing decision-makers’ requests with sources of data, thus allowing for effective and efficient integration. By using boundary objects as means to create mutual understanding between participating communities, we proposed an initial scheme for DWH project improvement as a first step in our design science project. According to Kimball and Ross (1996), the combination of understanding the business requirements and the available data is necessary as a first step toward the design of a DWH (Kimball and Ross 1996, p. 33). From our perspective, only DWH professionals belonging to both BEDFs’ and OPSs’ language communities can live up to these design expectations.

7 ACKNOWLEDGMENTS

We would like to thank our reviewers for their recommendations on our work. Furthermore, we would like to thank the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung), which funded this work under record no. 01FD0611.
References


EFFICIENT COMMUNICATION IN FINANCIAL DATA WAREHOUSING PROJECTS – INSIGHTS FROM A MULTIPLE CASE STUDY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0387.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Face-to-Face Communications, Data Warehousing, Case Study, IT Project Management</td>
</tr>
</tbody>
</table>
EFFICIENT COMMUNICATION IN FINANCIAL DATA WAREHOUSING PROJECTS – INSIGHTS FROM A MULTIPLE CASE STUDY

Räkers, Marc, zeb/information.technology, Schlossstr. 22, 48455 Bad Bentheim, Germany, marc@raekers.com

Rosenkranz, Christoph, Goethe-University, Grüneburgplatz 1, 60326 Frankfurt am Main, Germany, rosenkranz@wiwi.uni-frankfurt.de

Abstract

Data warehouses play important roles in the IT landscape of the financial industry. Banks have to deal with complex communication issues in financial data warehouse projects. Especially the creation of extraction, transformation and loading (ETL) processes depends on the project team’s communication ability and given communication barriers. We briefly present a theoretical efficiency model based on philosophy of language and the conceptualization of information systems development projects as language communities. We test the model against three case studies to derive evidence for and confirm three propositions leading to suggestions for project management of financial data warehouse projects.

Keywords: Financial Data Warehousing, Case Study, Project Management, Communication Efficiency.
1 INTRODUCTION

In order to cope with proliferating complexity (Courtney, Merali, Paradice and Wynn 2008), information systems (IS) are often developed in the form of structured approaches (Hirschheim, Klein and Lyytinen 1995, p. 33). IT projects are a structured set of activities concerned with delivering a defined IT capability to an organization based on agreed schedule and budget (Ribbers and Schoo 2002, p. 45). Complex IS development projects are large-scale IT projects with a significant and complex software component (Xia and Lee 2005). IS development projects are inherently complex because they must deal with both the technological issues and the organizational factors that, by and large, are outside of the project team’s control (Xia et al. 2005, p. 46). Consequently, the complexity of IS development manifests itself in the high failure rate of IT projects (SGI 2001, BCS 2004). This is especially true for large financial data warehouse (FDWH) projects (Behrmann and Räkers 2008, pp. 7-9);

1. FDWH projects have to deal with several business domains and fields of knowledge. To build up a centralized and integrated system a common understanding between all involved departments and subsidiaries has to be reached.
2. FDWH projects are characterized by high semantic complexity. Due to the different fields of knowledge a high number of context-dependent homonyms (e.g., limit, facility, book value, and market price) and synonyms (e.g., debt security, bond, and obligation) can be observed.
3. Specification-based approaches are not sufficient in FDWH projects. Beside the specification, additional methods are required, for instance, face-to-face communication.
4. FDWH projects require knowledge transfer methods and strong interaction. For a successful implementation a common understanding between all involved project members has to be ensured.

Therefore, the imperative of this paper is to develop and test an understanding of how communication influences decisions of (FDWH) project managers. The following research questions are of special interest to us: (1) How do managers deal with the efficiency and effectiveness of a (FDWH) project? (2) How do identified communication barriers influence management decisions?

The structure of the paper is as follows. In the second section, we discuss related work and the theoretical basis for the later sections. Then, in section 3, we introduce three case studies we extended to find evidence for our propositions. Afterwards, we discuss and generalize our findings in section 4. Finally, section 5 summarizes this paper and draws conclusions for further research.

2 RELATED WORK AND THEORY

2.1 Communication in IT Projects

It has been known for complex IT projects that coordination between the various stakeholders involved becomes a necessity (e.g., Gallivan and Keil 2003, Ko, Kirsch and King 2005, Joshi, Sarker and Sarker 2007). Communication is considered to be a prerequisite for organizational coordination (e.g., Malone and Crowston 1994, Nahapie and Ghoshal 1998, Pondy 2005, Quinn and Dutton 2005), and successful communication between involved stakeholders is deemed to be one of the main drivers for project success (e.g., Ribbers et al. 2002, Gallivan et al. 2003, Vlaar, Fenema and Tiwari 2008). Moreover, as organizations grow, they differentiate and specialize – a single company has several departments, a project has several specialized sub-projects. Each of the specialized units generate its own values, terms and coding schemes for information processing, which leads to a mismatch and a communication boundary hindering the information flow (Tushman 1977). From an information processing view, organizations information flows to function, and strive to create efficient information flows to be effective (Tushman and Nadler 1978, Jin and Levitt 1996). This perspective implies that project managers should first consider the tasks, composition and structure of subunits, and then
consider appropriate mechanisms for linking those units together, using a range of communication tools (e.g., electronic mail, fax, phone, management IS, etc.) (Levitt, Thomson, Christiansen, Kuntz, Jin and Nass 1999). Therefore, an important task of project management is to design the communication within the project organization in an effective way (Boland and Tenkasi 1995). Consequently, the structured sharing and communication of relevant information is crucial for the overall success of any project.

2.2 A Theoretical Efficiency Model

Looking at the challenges of FDWH projects (Behrmann et al. 2008), effective and efficient team communication plays an important role for project success. In this part of the paper we briefly sum up a theoretical model for explaining and predicting project efficiency in FDWH projects that is described in detail in Räkers (2009). This paper aims to test this model. The following brief description is needed to explain our further proceedings. One of the major roles of project managers is to ensure that a project reaches its goals (effectiveness). Additionally, from an economic perspective, s/he should try to reach the predefined goal with as few resources as possible (efficiency). As s/he decides how work should be done (types of action), s/he is the leading part in an ongoing process of decision-making. As different types of actions consume different resources in different amounts, the project manager has to decide – based on her or his assessment of the current situation – what type of action is perceived as the most appropriate for reaching a desired goal.

As a FDWH project’s success depends greatly on successful integration of source systems and consequently on the ETL processes (March and Hevner 2007), they especially have to deal with challenges regarding communication problems and common understanding of involved parties. While team members’ time efforts are the driving cost component of a FDWH project, communication issues become crucial for project success. Therefore, we argue that the efficiency of communication inside the project team has to be managed well to be successful. A conceptualization of IS building on philosophy of language adds thoughts of communication costs in differently developed, domain-specific language communities (Holten 2007, Holten and Rosenkranz 2008). By investing differently in terminology progress, language communities reach different levels of shared understanding and communication efficiency. This principle is shown in Nikolopoulos and Holten (2007) and was used in Räkers (2009) to develop a theoretical efficiency model. From this perspective, each project team is a more or less developed language community that progresses in the project with a subjectively observable efficiency. Only if project team members use and understand the same domain-specific terminology as a know-how held in common, ambiguity and uncertainty are both lowered, and misunderstandings resulting from different interpretations should occur less frequently. Another problem the project manager has to care about is the estimation and measurement of the project’s complexity (e.g., Xia et al. 2005). During the course of the project, as the project manager continuously engages into sense-making of the situation, complexity can be assessed more accurately.

To summarize the proposed model, using the cost function derived from the application of the conceptualization of IS as language communities, goals of higher complexity (increased variety) can only be reached by current members of the project team by investing more resources (i.e., person-days) into the development of domain-specific terminologies and creating a mutual understanding of a situation. The efficiency decreases while complexity that can be handled grows in the cost curve belonging to a project team, having its specific terminology T in place to attack problems of different complexity levels C (see Figure 1, left side).

In general, we argue that in a continuous process of decision-making, a project manager has to check the current project status and make forecasts for the project outcome (goal and complexity). The less unambiguous and uncertain a project environment is the better the forecast will be. Building on this, we propose that there are points D in time where project managers of FDWH projects take actions to shift the project team’s cost curve to a different efficiency level by investing into additional terminology. Doing this initially might raise costs in the short-term, but lowers overall costs of
communication in the project in a long-term view. The project-specific curve of efficiency is the path determined by costs curves and decision-triggered terminology-building steps between them (see Figure 1, right side).

![Diagram](image)

**Figure 1. Proposed Efficiency Model for ETL projects in FDWH**

This directly leads to the central propositions to be addressed in this paper:

P1: The curve of efficiency really exists in real project situations and can be observed.

P2: Projects go through phases with different levels of efficiency by creating additional terminology (terminology-building steps which shift the curve).

P3: Projects which explicitly address communication problems are more successful than others that do not.

We believe that project managers inherently observe projects and their team communication to detect points of efficiency change and that FDW projects should become more successful if project managers explicitly keep this possible shift in mind and react accordingly.

### 3 MULTIPLE CASE STUDY

3.1 Method & Case Selection

We decided to conduct a multiple, embedded case study in order to examine the central proposition made in section 2. Case studies are ideally suited when the investigator has limited control over events and boundaries of a phenomenon and when the phenomenon and the context in which it is investigated are unclear or closely related (Yin 2003). Challenges of communication and requirements specification in FDW projects certainly satisfy these criteria (Miles and Huberman 1994, p. 15). Since we wanted to test the propositions previously made, we followed a deductive pathway (Lee 1989).

Two case selection criteria were applied. First, the research decided to focus on FDW projects in the financial services industry. The selected cases are typical for recent projects in this sector. Second, one of the authors worked for more than six years at a major German consultancy which focuses on IT in the financial services industry. He therefore had broad experience with working in FDW projects, first as a consultant and later as a project manager, enabling him to establish contacts and to get in touch with the key informants. Moreover, he participated for some time in two of the projects (source of potential bias). The trigger for all the projects was the supervisory requirements known as Basel II. These regulations have an impact on the whole structure of a financial service provider and especially affect ETL due to the needed integration of various source systems.
Three FDWH projects of various banks in Europe were selected, which offered access to members of all stakeholders involved, that is, internal project members from IT and business units, external IT and business consultants, and project managers. Table 1 gives an overview of the selected cases.

<table>
<thead>
<tr>
<th>Natural languages used in project</th>
<th>Bank A</th>
<th>Bank B</th>
<th>Bank C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope</td>
<td>German, English</td>
<td>German, English</td>
<td>German</td>
</tr>
<tr>
<td>Project duration</td>
<td>Europe</td>
<td>Europe</td>
<td>Germany</td>
</tr>
<tr>
<td>No. of countries involved</td>
<td>&gt; 13</td>
<td>&gt; 2.5 years</td>
<td>&gt; 2 years</td>
</tr>
<tr>
<td>No. of subsidiaries involved</td>
<td>18</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>Balance value</td>
<td>100 bn. EUR</td>
<td>70 bn. EUR</td>
<td>7 bn. EUR</td>
</tr>
<tr>
<td>Project size (person-days)</td>
<td>about 100,000</td>
<td>about 40,000</td>
<td>about 5,000</td>
</tr>
<tr>
<td>Core system change</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

3.2 Data Collection & Data Analysis

The risk that participants displayed retrospective bias was minimized because at the time of the interviews, all selected projects or directly related successor projects were still ongoing. Data was collected from various data sources and with the help of different data collection methods, enabling triangulation (Eisenhardt 1989, p. 539). The interviews were undertaken in 2008, two years after the involved researcher had dropped out of the projects in order to counter his potential bias. Unstructured and semi-structured interviews, project documentation, and self-estimation surveys of project members activities were used to generate the data for a case study diary (Yin 2003, p. 105). The interviews were guided by an interview outline which was not shared with the interviewees and was only used by the researchers as a checklist and guideline. The interviewees were encouraged to provide a narrative of their experiences as freely as possible. The interviews were all recorded and fully transcribed. Follow-up e-mails and draft reports were sent to request clarifications and to offer informants the possibility to provide feedback and comments. Natural controls and treatments that were already in place were utilized in testing (Lee 1989, p. 39). For instance, the control was already in place by virtue of focusing on persons which were participants in more than one of the three projects, interviewing the same person for different projects.

3.3 Case Descriptions

Bank A is an international banking group having 18 major subsidiaries. The project at Bank A has been started in 2001 and was planned to end in December 2007. To fulfil the requirements of Basel II and to enable a group-wide calculation one of the main tasks of the project was to implement a central data warehouse. Primarily, data of transactions, collaterals, customers and rating information of both head office and subsidiaries had to be delivered into the central data warehouse. To achieve this, each subsidiary had to deliver specification-conform flat files to be imported into the central data warehouse. At the beginning, the core project team at headquarters, which consisted of consultants and Bank A experts from business and IT had designed a detailed data model for Basel II, and specifications for data loading interfaces in XML Schema Definition (XSD) format, which were send out to the project members in the subsidiaries (i.e., network banks). At the beginning of the project, an initial training workshop was conducted to explain the XSD format to the subsidiaries.

Bank B is a large European banking group with foreign subsidiaries. In order to deal with Basel II demands, Bank B put up a FDWH project in order to fulfil the requirements of Basel II. The project started in 2005 and officially ended in 2007. In addition, Bank B was in a post-merger situation and was the target of an acquisition during the project, so the ongoing consolidation process in the banking industry is considered as well. In addition, beside the Basel II-related data warehouse project, another large project was established in Bank B at the same time. A new core banking system was developed and switched productive during the second half of 2006. So nearly every department was involved in
two large and important projects at the same time while doing the normal daily work as well. From core banking systems over the group-wide data warehouse to the reporting and supervisory systems, nearly every application and business process was affected. All in all, Bank B was undergoing enormous organizational changes, and the changing environment heavily influenced the project.

Bank C is mid-sized European bank with mainly national activities. While implementing a new core banking system the bank’s management decided to take the opportunity to renew the technical landscape regarding regulatory systems and the underlying data warehouse as well. The project started in 2005 and was finished by the end of 2007. One of the major goals was the implementation of the Basel II requirements, which makes this case comparable to the cases of Bank A and B. Like in the other cases, this project was supported a consultancy and all departments of the bank were involved in the project. Looking at Bank C, this case gives us a different view on a smaller and mostly national bank to compare and analyze whether size and geographical diversity have impact on the projects and the theorized curve of efficiency.

3.4 Case Findings

3.4.1 Bank A

The complexity at Bank A was perceived as high (1) due to new regulatory demands and (2) the existence of various important subsidiaries with different environmental situations. However, since the specification was assumed to be unambiguous by the project managers due to the use of XSD, it was sent to each subsidiary via e-mail without further descriptions. Furthermore, the project managers believed that employees in the subsidiaries would share the same knowledge about the banking domain as employees at head office, that is, that the meaning of each of the used terms and descriptions would be understood without any problems. Since after the distribution of the specification all support requests of the subsidiaries could be satisfied by a contact person at headquarters, the head office team assumed that each subsidiary had understood the specification in the intended way. But first completeness checks of the flat files that were delivered by the subsidiaries showed several problems resulting from a different understanding of the meaning of the terms. It seemed as if employees from head office and subsidiaries spoke different languages, since the interpretations of the published specification varied to a considerable degree, although the required data and data structures were perceived to be unambiguous due to the use of XSD and additional extensive textual descriptions.

“A classical example is that the central business data base table was named ‘counterparty’ due to a translation error […]. This caused a lot of confusion because therefore this term appears relatively often in the data warehouse. […] In fact, there are really 13 data warehouses that have to be loaded identically, but which weren’t loaded identically […]. This was the case in the first implementation phase, that one just gave the documents to everybody and everybody did interpret them. And the result was 13 interpretations. […]” (Mr. JS, project worker)

“At the moment of the rollout, specific concepts were not clear for the people on site, what turned out only afterwards, and specific types of business naturally only existed locally or, again, used other terms, so that it was not clear to them how to convert that adequately into the warehouse.” (Mr. WB, project manager)

After this problem surfaced (following the first release of the data warehouse that had already been implemented), a change in the course of action was observable at Bank A. Additional workshops were conducted once the project management team realized that the sent XSD specifications were not enough for coping with proliferating complexity of the situation. During the workshops, changes to the data model in subsequent releases were introduced and the meaning of terms and concepts was discussed in detail with all involved stakeholders of business units and subsidiaries. After the workshops, once it was realized that the delivered data quality was still not satisfactory, the project
managers decided to engage into local visits of head office project members at problematic subsidiaries. Patterns of terminological discourse are traceable once the existence of separate sub-terminologies and jargons had been understood and accepted, and those sub-terminologies began to be consolidated into a terminology by the project team (generation of language community). The project managers explicitly decided to switch from written specifications to direct face-to-face communication. The joint discussion, clarification and learning of terms (i.e., the creation of a language community) were time-consuming, but perceived as necessary for reducing ambiguity and to be able to deal with the complexity at hand.

“At one point in the project [...] it was recognized that the specification must be improved and refined [...] then one realized that this XSDs are not enough, and then one created the data delivery concept, and drove out to the subsidiaries to explain this. This were points where one noticed that they needed a little more support, but the real turning point, where this issue had full management attention, came actually in April 2007, where one said the data delivery does not work at all [...] and then did decide that we need so-called task forces, [...] very different persons now became involved, who really were on-site [at the subsidiaries] and went through the data with them weekly, looked at the data fields and discussed every single thing. This actually was one of the greatest turning points for me. There, the upper management realized ‘Now we have to do something, and we have to spend money [...].’” (Mr. BK, project worker)

The on-site visits and the discussions with all project members and subsidiary employees finally succeeded in establishing a shared, mutual understanding of terms. Ambiguity was mostly lowered by the subsequent face-to-face discussions between all project members. Afterwards more complex topics were discussable.

“Absolutely, it became apparent that the number of questions which one brought back from the countries and the on-site visits, that by looking at these one saw, that many open questions existed and once we had those, we could clarify those. [...] There are still questions, but these questions are much more difficult to answer because most of these require much more detailed knowledge. Also it didn’t run like it’s on rails afterwards directly, but I think, the amount of questions which had been clarified in this did certainly justify this [the on-site visits].” (Mr. JS, project worker)

Thus, the project went through phases of different communication policies due to project management decisions. Each of these decisions changed the style of communication. For example, based on assessments of the self-estimation survey, these changes doubled the time spent on communication and also doubled efforts for on-site meetings. By conducting expensive on-site meetings, a shared terminology and a common understanding regarding project issues was built. After those changes in policy, project members state a perceived increase of efficiency in the project in later phases compared to earlier phases.

3.4.2 Bank B

In general, the complexity of the project was perceived to be very high, since (1) Bank B was faced with new regulatory demands, (2) Bank B was target of an acquisition, and (3) a new core banking system was implemented in parallel to the Basel II project. Ambiguity and uncertainty were perceived as being high due to the existence of several jargons and terminologies in the project and its surrounding environment. As in Bank A, particularly the creation of a shared understanding between head office and subsidiaries proved to be difficult. On the one hand, the interfaces to the subsidiaries and the core banking system were very complex. On the other hand, the types of businesses in foreign countries were fundamentally different from businesses at head office, leading to an even greater complexity and much more ambiguity for the project, which was not perceived as a problem at the beginning. Similar to Bank A, the project managers at first did not address that different understandings and meanings for the same words existed in Bank B.
“[...] There was this nice module, which was called [offer processing application] [...]. Under this term, one can imagine quite a lot, and there we had a lot of discrepancies between both projects, even concerning the question where this is positioned with regard to IT, technically. And then we had different coordination meetings where one said ‘OK, what really is the content of this system?’ so that later you really could clarify the business activities and could limit responsibilities and interfaces relatively well.” (Mr. TA, project manager)

But as different understandings proved to be the case, the procedure for dealing with those problems explicitly changed as the project progressed (change in the course of action). According to our findings, the project at Bank B was characterized by two distinct procedure models, where the shift from procedure 1 to procedure 2 was consciously made by the project management team. The first phase was specification-driven whereas the second phase was communication-driven. For example, the project manager’s self-estimation efforts for on-site meetings were three times higher during the use of procedure 2 than before using procedure 1. For some of the subsidiaries which showed immense problems, so-called task forces were established. The technical face-to-face coaching on-site was very important to make members at the subsidiaries understand the requirements. Therefore the creation of a language community by intensive face-to-face communication was one of the central factors for project success and the handling of the high complexity.

“This then changed from the initial procedure of ‘We send you a concept, look what you got and send us back the results’ to a much more workshop-oriented procedure. So one did say ‘We have to support you much more, we come over to you and look everything through in workshops together with you’ and in a second step one conscripted the colleagues respectively and said ‘We let our project member sit with you and he will support you for the next days and weeks.’” (Mr. TA, project manager,)

Due to the fact that one key informant was a team member of both projects in Bank A and Bank B, we could directly address the difference between both projects. The initial external complexity and uncertainty due to Basel II are comparable. Bank A has more subsidiaries than Bank B, but Bank B was the target of an acquisition during the project and also implemented a new core banking system.

“Uncertainty in both projects was rather almost equal. At Bank A, it simply lasted five years [...] and had a multiple of effort [of Bank B]. [...] they certainly started in 2003 in earnest, and practically, they now have reported for the first time [...]. They had a much longer project runtime. Fair enough, it is a slightly larger bank with much more subsidiaries, so they have 18 subsidiaries in different countries [...], this is really much bigger in dimension. But they also had to make some real turnarounds and extra loops within the project, concerning the subject of classification: what is market risk, what is credit risk? Where do I represent the credit risk, in which data warehouse, where do I represent the market risk, and so on. They certainly got closer to the target more serpentine [...].” (Mrs. SK, project worker).

So while external complexity, ambiguity and uncertainty that both projects faced due to new regulatory requirements and other external circumstances were comparable, Bank B reached the same goal faster (two years compared to six years), and the quality of the processes and product were also perceived as being clearly superior. Basically, Bank B changed its course of action (investing in language community creation and shifting the curve of efficiency) earlier and faster.

“[Bank B’s project] also had a much shorter duration, thus we have now in 2008 reported with the standardized approach, and the project has started in 2006, that is, we have implemented the basic functionality for Basel II in two years [...] the duration compared to the number of resources that have been burnt, is certainly much lower than at Bank A [...].” (Mrs. SK, project worker)

3.4.3 Bank C

The FDWH project at Bank C had to integrate the topics “regulatory reporting” (Basel II, loss database) and “controlling” (operational and business control). The to-be-developed data warehouse was intended to be an integrated database for both project topics. Consequently, changes in one area
had significant impact on the other area, which had to be taken into account. During the project, three different consultant teams were active at Bank C and had to work closely together to implement the new core banking system, the regulatory reporting system and the central data warehouse. Since the data warehouse team had to build interfaces and integrate data from both other systems, this was deemed one of the most important tasks for project success. External and internal complexity were both perceived to be very high by the project management team (1) due to environmental factors (e.g., new regulatory requirements) and (2) the organizational situation (e.g., implementation of a new core banking system). However, in contrast to the previous projects, no subsidiaries did exist. At the beginning, the presence of three teams was hindering the project due to the existence of several different terminologies in the teams. The teams had quite separate IT-related jargons and also different understandings of banking-related technical terms that were important for the project. Between those sub-language communities, misunderstandings occurred quite frequently during the project. This was caused by the perceived ambiguity concerning the different meanings of terms in the different teams.

“The example ‘validierter Saldo’ [German: validated balance] strikes me on this level. […] this wasn’t a problem at the beginning of the project, […] but something that happened at some time during the project […] for us it was clear, this is an established term and means the actual capital disbursed […] but in the core banking system team, they understood it differently, or in the core banking system the data field with this description yields something different respectively, namely the original capital disbursed […] this led to many bad coordination loops and to costs on our side because we had to adapt the data history.” (Mr. VC, project worker)

As in the previous projects, the general procedure was consciously changed by the project managers in order to create a shared language understanding and to reach the project goal on time. For example, the efforts for communication became three times as high as before.

“[…] In the first phase, one had the attitude, we really don’t want to understand the core banking system, instead, we want to have interfaces filled, delivered data requests catalogues respectively and one hoped that this will match somehow. Then the second phase […] ‘OK, that isn’t really working out, now we sit together and try to understand their basic problem.’ And in the last phase, I had the feeling that one had a look at the core banking system oneself a little bit, did discuss this with some bank employees and the interface developer […].” (Mr. JD, project worker)

This consolidation and the intensive discussions of all involved stakeholders reduced ambiguity and finally led to better data quality. The project was perceived as running more smoothly once ambiguity had been reduced or when it was noticed that a different understanding of a term was existing.

“When you […] talked this through together, then you had no discrepancies. In certain circumstance, it wasn’t clear for single data field, because our software has its requirements and if you don’t name them explicitly, they are hard to infer from the field names or field descriptions. […] you have to address such things explicitly if this was somehow lost in the torrent of data requests. […] Such things are very important, meaning the precise description of data requests or technical contents. In my judgment, this accounts for more than half of project success in such an IT or data warehouse project.” (Mr. VC, project worker)

4 DISCUSSION

To link our findings to our propositions we sum up our findings from the cases.

- In each case the project management team got to a point where a change of the general procedure and original project plan (type of action) was undertaken in order to increase efficiency or at least being effective. The types of action became more communicative and sense-making-related and succeeded in creating a more shared and joint terminology. So within all cases, the existence of a curve of efficiency as theorized before can be observed (proposition P1).
• Project members of all projects state that fundamental problems in communication and understanding existed which led to a change in the project procedure. We recognize these approaches as an increase of terminology creation and as a shift to a different level of efficiency (proposition P2).

• By changing the efficiency, all three projects assured project success. Projects which addressed those communication issues earlier (i.e., Bank B compared to Bank A) were perceived to be more successful than others (proposition P3).

Two important consequences for research and practice arise from our findings. First, the extension of the cost curve to a cost curve with multiple efficiency levels (curve of efficiency) allows the explanation of communication behavior and possible types of action for project managers in complex project situations. Therefore we are very confident that the derived efficiency model, which was applied to FDWH projects in this paper, is stable for other types of complex IT projects as well. This leads us to suggest that replicating our study in other settings and using a diversity of different research methods (e.g., case studies, surveys, experiments, action research) will lead to similar results. Second, project managers in practice should explicitly address communication efficiency in their project management approaches and project plans for integration and ETL processes in complex (FDWH) projects. Variety and ambiguity are major problems in complex IT projects. It has been over 60 years since Ashby (1956) proposed that only variety destroys variety, and over 25 years since Daft and Macintosh (1981) argued that ambiguity and uncertainty are important organizational design factors. But still IT project managers make the mistake of misjudging the impact of those factors and try to implement IT before the variety of a problem situation has been matched and ambiguity has been reduced by creating a language community. Usually, issues and problems of communication, understanding and sense-making are not a part of the “accepted” or “rationalized” costs. But as our findings suggest, those costs should be explicitly addressed and incorporated in the management process of complex IT projects.

Case studies can confirm, challenge, or extend a theory because they can be used to reject propositions (Lee 1989). The theoretical efficiency model itself has been previously derived from literature and exploratory field studies, and has been challenged by us with case study data. Qualitative methods were employed to generate an interpretive understanding of the problem situations and to test the propositions (e.g. interview-based evaluation techniques). Nevertheless, further studies are needed to confirm the efficiency model for complex IT projects beside current FDWH projects and to generalize our findings. The chosen cases deal with comparable project topics, but the findings are not derived from a ceteris paribus (“all other things equal”) configuration; instead, we rely on “natural” situations, where influences of other factors cannot be excluded. Furthermore, we greatly rely on subjective impressions from case study participants and our own interpretation. In order to counter those problems, rigorous case studies that want to test theories must address four requirements (Lee 1989, p. 42), which we explicitly tried to address. As a check for falsifiability, we verbally formulated propositions based on the model of the curve of efficiency, which have subsequently been tested against the data. As a check for logical consistency, the propositions follow from the logic of this model. The cases have been selected in order to provide generalizability and to allow checking for consistency. As a check for empirical validity, the theorized curve of efficiency has been confirmed and corroborated. Since this study was interested in providing first empirical evidence for the curve of efficiency and not in refuting any rival theory, a check for relative predictive power has not been addressed yet.

Since we wanted to provide first evidence for our propositions, this research is mostly dealing with construct validity. According to Yin (2003), construct validity is critical in qualitative research and can be improved by three tactics. The first is the use of multiple sources of evidence, which provides multiple measures for the same phenomenon. We explicitly addressed this issue and examined three projects with a multitude of informants in a multiple, embedded case study. The second tactic is that key informants review the case study, which we have addressed as well: the research project was introduced and discussed with the key informants; draft reports were sent to the key informants for
Finally, Yin (2003) recommends maintaining a chain of evidence, which we did by formulating propositions based on the efficiency model, collecting case study data, discussing it with our informants, and creating a case study diary that allows us to trace back from findings to initial propositions. However, further tests are necessary to corroborate our findings, e.g., in laboratory experiments.

5 CONCLUSION & OUTLOOK

The variety of complex IT projects is often underestimated and therefore has to be managed. We used a theoretical model based on efficiency considerations to explain the general behaviour of integration and ETL processes in FDWH and derived three major propositions. Three projects were evaluated in a multiple, embedded case study. Our findings provide evidence for corroborating our propositions. Project managers dealing with integration and ETL processes in FDWH can benefit from explicitly addressing communication issues in their project plans and general project procedures. By finding ways to remove existing communication barriers and misunderstandings, the efficiency of a project team can be improved. The project manager has to decide how much efforts need to be invested into this improvement to be most efficient for reaching the desired project goal.

From a theoretical perspective, the efficiency model is a building block for knowledge on communication in FDWH projects. We contributed to project management by providing concepts for this efficiency model. If “sound” FDWH project management is among the desired goals, researchers and practitioners may benefit from our insights on how to measure the relationship between communication, complexity and ambiguity, and FDWH project success. Of course, several instances are needed to support a theory. At present, the projects are the only setting for which our findings are valid. Our study needs a replication of the findings in additional cases. To strengthen our findings, we will conduct additional case studies to deal with more special topics of other projects and action research settings for explicitly addressing communication barriers in early phases of projects.

References


A COMMUNICATION EFFICIENCY MODEL FOR ETL PROJECTS IN FINANCIAL DATA WAREHOUSING

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0615.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Cost-benefit analysis, Data Warehousing, Management practices, Theory Building</td>
</tr>
</tbody>
</table>
A COMMUNICATION EFFICIENCY MODEL FOR ETL PROJECTS IN FINANCIAL DATA WAREHOUSING

Räkers, Marc, zeb/information.technology, Schlossstr. 22, 48455 Bad Bentheim, Germany, marc@raekers.com

Abstract

The financial industry relies greatly on information technology (IT) because of its work with immaterial goods. Nowadays, most information collected by a financial institution is usually stored in a central data warehouse system. This central financial data warehouse (FDWH) is permanently under construction. Requirements from all over the bank have to be met by FDWH projects in an ongoing process. These requirements need to be communicated to the FDWH project team to do implementations properly throughout the system landscape. Especially the creation of extraction, transformation and loading (ETL) processes depends on the project team’s communication ability and given communication barriers. Enhancing recent research in FDWH a theoretical efficiency model based on philosophy of language and project management fundamentals is built in this paper. The conceptualization of information systems development projects as language communities is an important presumption for this theoretical model. Suggestions derived from this model lead project managers of FDWH projects to more appropriate decisions keeping efficiency drivers in communication in mind.

Keywords: financial data warehousing, theoretical efficiency model, project management, communication efficiency.
1 INTRODUCTION

Financial data warehouses (FDWH) are an important and relevant issue for most financial services providers today. Especially large banking groups with foreign subsidiaries have to deal with new regulatory demands, e.g., Basel II, IFRS or Sarbanes-Oxley Act (Behrmann and Räkers 2008). Moreover, financial service providers try to employ DWH for controlling their risk management approaches. It is general consensus that DWH support the management perspective on business processes technically, but their implementation is extremely costly (Vassiliadis 2000). Therefore, a clean design and specification of the system is of great importance. Consequently, the involvement of the management users and management’s support are deemed key factors for DWH quality and system success (Wixom and Watson 2001, p. 35 f.). Since DWH projects often fail or significantly exceed budgets, existing research has concentrated on quantitative or qualitative analyses of success factors as well as contemporary best practices for building DWH (Weir, Peng and Kerridge 2003, e.g., Herrmann and Melchert 2004, Hwang, Ku, Yen and Cheng 2004, Watson, Fuller and Ariyachandra 2004, Hwang and Xu 2007).

March’s and Hevner’s (2007) thorough literature research warns how links between data warehousing, strategic decision-making and evaluation are under-researched. For March et al. (2007), integration (i.e., the consolidation of data from disparate sources into one consistent body of data) lies at the heart of decision-making tools and is one of the biggest challenges for DWH designers (March et al. 2007, p. 1036). Effective integration can only be achieved by experts who have access to knowledge about internal and external business contexts (i.e., “what questions are momentarily important?”) and access to knowledge pertaining operational systems that have to be integrated (i.e., “which operational systems provide data for answering those questions?”) (March et al. 2007, p. 1041).

This theoretical argument is consistent with statements which suggest that FDWH projects show specific integration characteristics (Behrmann et al. 2008):

- **Financial data warehouse projects have to deal with several business domains and fields of knowledge.** This leads to different domain knowledge in different departments and subsidiaries and specific fields of knowledge arise. To build up a centralized and integrated system a common understanding between all involved department and subsidiaries has to be reached.

- **Financial data warehouse projects are characterized by a high semantic complexity.** Several business domains can occur within a single financial institution. Due to the different fields of knowledge a high number of context-dependent homonyms (e.g., limit, facility, book value, market price) and synonyms (e.g., debt security, bond, obligation) can be observed. Due to this variety, the definition of terms (e.g., dimensions and measures) is challenging.

- **Specification-based approaches are not sufficient in financial data warehouse projects.** Beside the specification additional methods are required, for instance, face-to-face communication.

- **Financial data warehouse projects require knowledge transfer methods and strong interaction.** For a successful implementation a common understanding between all involved project members has to be ensured. This cannot be reached by merely interchanging a written specification.

In this paper a theoretical model is developed to explain and predict interconnections in FDWH projects that lead to project management decisions which change general procedures within these projects over time. During phases of terminology creation, a foundation for more efficient communication (i.e., better aligned to the current project environment) is built by the project team. This changes the ability of the project team to solve problems either by using formal specifications or using face-to-face communication. We argue that the beginning of a phase of terminology creation is decided by the responsible project manager based on efficiency considerations and the perception of project status and future project progress.

The underlying research approach is in line with Lee’s (1991), approach of integrating positivist and interpretive approaches in organizational research. After having created a subjective understanding of
everyday meanings and common sense within observed organizations, which provides the basis for the interpretive understanding, the researcher creates a positivist understanding in order to explain the empirical reality – the explanation being a scientific theory which can be tested against the subjective meaning as recorded in the interpretive understanding (Lee 1991, pp. 351-354). Therefore, as a first step, the subjective understanding of involved persons in FDWH projects was analyzed by creating an interpretive understanding of the specifics of FDWH projects (Behrmann et al. 2008). In order to create a positivist view on FDWH projects, this paper introduces an efficiency model which explains theoretically why observed behavior and procedures in FDWH projects operate as they do. Further research will test this model in additional FDWH project situations and try to apply the model in a broader scope than financial data warehousing.

The remainder of this paper is structured as follows. In section 2, related work concerning research approach, software project management and language theory is presented. Section 3 contains the communication efficiency model for FDWH projects. Finally, section 4 summarizes the findings and gives an outlook to further research.

2 RELATED WORK

ETL projects take about 70 percent of the efforts of a data warehouse project (Kimball and Caserta 2004, p. 391) and therefore need professional project management. Accordingly, Kimball et al. (2004) suggest specific project structures and organizational rules to manage the inherent complexity of ETL projects. To complete a project successfully, the objectives (goals) have to be reached within time, cost and performance as the most critical project dimensions (Jurison 1999). Projects are usually conducted in four generic phases (e.g., Jin and Levitt 1996, Kunz, Christiansen, Cohen, Jin and Levitt 1998):

1. Conceptual phase: the project scope is determined by the project sponsor and the project manager. Software designers work together with the project manager on possible arrangements of the future software.

2. Planning phase: the project manager builds a project plan containing project tasks. Each task has allocated resources, duration and a defined result. This information is generated by conducting estimations with different techniques as summarized by Boehm (1981, p. 203).

3. Execution phase: in the execution phase the software development and testing of the built software product takes place.

4. Termination phase: During the termination phase the project is documented, closed and delivered to the operations department of the company. The project sponsor approves the results to the project manager.

Furthermore, complexity is identified as one of the major problem-driving factors in information system development (ISD) projects (Jurison 1999, Xia and Lee 2005). With the advent of the fourth generation of ISD methodologies in the 1980ies it became clear that intensive user involvement is needed for coping with proliferating complexity, and ISD has an important social component as well (Hirschheim, Klein and Lyytinen 1995, p. 36). The social system resulting from an ISD project is closely interconnected to the project team itself. Therefore, purely mechanistic and formalized system development methodologies need to be treated more critically (Fitzgerald 1996). For example, to be able to handle complex ISD projects successfully, different kinds of communication between project members need to be combined efficiently. Effective communication has already been researched concerning requirements (Coughlan and Macredie 2002). Although requirements play an important role for FDWH projects and especially ETL processes, this paper focuses on explaining operational project management and not on building an additional methodology for requirements specification.

Adding to this, a conceptualization of IS building on philosophy of language adds thoughts of communication costs in differently developed, domain-specific language communities (Holten 2007, Holten and Rosenkranz 2008). By investing differently in terminology progress, language communities reach different levels of communication efficiency. This principle is shown in
Nikolopoulos and Holten (2007). This and other findings form organizational research (e.g., Boisot and Li 2006) suggest that the building of a shared terminology and language community is beneficial for coordination and communication.

3 A THEORY EXPLAINING MANAGEMENT IMPACT OF COMMUNICATION EFFICIENCY IN FDWH PROJECTS

3.1 Basic Components of an Efficiency Model for FDWH Projects

Further development of the efficiency model proposed by Nikolopoulos and Holten (2007) is based on project management basics, for example, as summarized by Jurison (1999). The relevant management dimensions can be derived directly from Jurison (1999) and have been observed to be of relevance in empirical studies of FDWH projects (Behrmann et al. 2008, Räkers and Rosenkranz 2008):

- **Time**: the time needed to reach a project’s goal. Time is a constraint in projects as there usually is a specified deadline in a project plan. Furthermore, time is an effort-driving component of a project. As overall tasks like project management need resources every day during the running time of the project, a longer project needs more resources for these tasks.

- **Cost**: the amount of efforts needed to reach a project’s goal. The main driver for costs is the assigned project team members. Personnel expenses are the relevant cost driver in FDWH projects, because the hardware and technological components are usually needed anyway, are becoming more and more a reliable and well-tried commodity nowadays (Avison and Fitzgerald 1995, p. 6), and do not influence the organizational view on project management this paper tries to address. The working hours of a team member can be split into travel time, meetings, telecommunication etc.

- **Performance**: the created results of a project. The performance of a project is tracked by project managers throughout the whole project. Performance corresponds directly to the benefit a project has for its sponsor. For the project sponsor, the benefit is the primary goal of the project. From an economic point of view, the (rationalized) goal of a project can be summarized as creating a benefit with least possible costs under the constraint of time.

In ISD projects, complexity has been often identified as one of the main reasons for problems. Complexity is often underestimated when starting a project. As complexity is one of the problem-driving factors mentioned by Jurison (1999, p. 4), it becomes an important part of the efficiency model as well. However, complexity is difficult to pin down. Generally, the term has a multiplicity of meanings (Flood and Carson 1993) and is a fundamentally subjective concept (Backlund 2002). In this paper, we build on Ribbers and Schoo (2002) and argue that the complexity of a (FDWH) project evolves from the variety of a project environment. Complexity arises from the inherent *variety of elements and interconnections* of an ISD project (Rosenkranz and Holten 2007). The project organization (i.e., the project team and its organizational setup) needs to be able to handle this variety. According to the law of requisite variety only “variety destroys variety” (Ashby 1956), the project team essentially needs to build up this variety itself. To be able to handle a certain degree of complexity, the variety of this situation needs to be addressed by a project team with at least the same variety at hand (Bar-Yam 2004).

In the beginning of a project, the project manager usually estimates the complexity of the project environment and builds up a project plan to address time, costs and the expected result (benefit for the project sponsor). After the project has started, the project manager has to track the status of the project. This is done by measuring time, cost and results while at the same time estimating whether the rest of the plan fits to the actual situation. As ambiguity and uncertainty in new situations usually are comparably higher at the beginning of the project, the more the project progresses, the less ambiguous and certain it becomes – maybe by risk elimination techniques (Boehm 2003, p. 9). The estimations become more realistic and more precise. While the desired project goal does not change, the variety...
(i. e., complexity) of both the environment and the required tasks to cope with this variety become clearer with every step that is done towards the project goal.

A project plan consists of project tasks that are linked to each other as a network of activities. The most popular illustration of a project plan is a Gantt chart to visualize interdependencies of project tasks. All of these project tasks also have a desired benefit, a timeframe and costs. The project manager arranges different types of actions to reach his or her major project goal. Each type of action has its own profile regarding benefit, time and costs. During the project, the project manager comes to decisions about changing the project plan by reassembling the actions and changing types of actions to come to a different structure of benefits, time and costs. A constraint coming from the project sponsor is to deliver the expected project results, which means to be effective in the project. The project manager is usually trying to achieve this by also having economic efficiency in mind. Efficiency is defined here as the benefit reached per cost. Efficiency is important to the project manager for two reasons. Being efficient during the first part of the project increases freedom in later parts of the project. Furthermore, higher efficiency might be rewarded by the project sponsor. However efficiency becomes most important for the project manager when s/he realizes that the project goal cannot be reached with the current project plan, because the variety (i. e., complexity) of the project was underestimated earlier. As the expected benefit is fixed by the project sponsor, the efficiency can be increased by decreasing the project costs. It was mentioned above that the costs of a FDWH project are driven by personnel expenses for ETL process creation. Furthermore communication drives personnel expenses, what makes communication costs a very important influence factor for FDWH projects.

3.2 Matching FDWH Project Management with Communication Costs

One way of bringing theories about communication costs and project management together is having a look at the project costs. Assuming that the costs of communication have a major impact on project costs and assuming that communication has a major impact on project efficiency as well, the total costs of communication are chosen as an important measure for the suggested efficiency model. As benefit and time are constraints given by the project sponsor, the remaining influence factor is complexity. A visualization of this context is shown in Figure 1.

![Figure 1. Cost Curves: Total Costs of Communication given Degrees of Complexity (adapted from Nikolopoulos & Holten (2007))]
for a language community to deal with it, the variety of the problem becomes greater than the variety of the language community (i.e., project team). To cope with this increasing variety, the terminology of the language community has to be increased as well. In a project environment, for example, this can be done by decisions about activities (D) of the project management, that is, explicit decisions to engage into language community creation and sense-making. On the one hand, these decisions increase the total costs of communication. On the other hand, they subsequently create a more effective terminology for the language community and shift the project to a different cost curve.

In Figure 1, the upper cost curve is less efficient for project degrees of complexity below C1, but is more efficient for degrees of complexity above C1. We propose that if a project manager adapts the estimations for the degree of complexity in his or her project from a level below C1 to a level above C1, and his or her project team uses terminology T1 right now, s/he will decide (D1) to engage into terminology creation and to switch to terminology T2 by investing in this terminology training (special type of action).

3.3 Enhancing the Cost Curves to an Efficiency Model

To enhance and detail this model findings from FDWH project research (Behrmann et al. 2008) are applied. Proceeding from a terminology T1 to another terminology T2 in FDWH projects means to change the style of communication from a specification-based way to a more personal face-to-face communication. By doing this, a higher degree of complexity can be handled by the project team. This is consistent with usual theories about media richness and information processing (e.g., Daft, Lengel and Trevino 1987) Therefore, the shift (D1) from the cost curve of terminology T1 does not proceed in a straight line, but is more diagonal as shown in Figure 2.

![Figure 2. Efficiency Model for ETL Projects in FDWH](image)

During large and long-running projects, multiple of these decisions are observable. This leads to a step-wise increase of total communication costs concerning the degree of complexity of a project. While the project team deals with more and more complexity over time, this step-wise increase should be observable in real projects. Consequently, as multiple cost curves are relevant for a project manager’s decision (i.e., a set or continuum of cost curves), there are a lot possibilities to be taken into account. Depending on his or her estimation of how a single cost curve appears, s/he (intuitively or deliberately) identifies thresholds (such as C1 or C2 in Figure 2).

The next step in our construction is to fit the project manager’s estimation of project complexity into this model. In case the estimated project complexity is below C1, the project stays with terminology
T1. Having an estimated project complexity between C1 and C2, the project manager should decide to invest into terminology T2. If the estimated project complexity is above C2, the project manager should think about additional terminology building or enhancements, and at last decide to switch to terminology T3. The example in Figure 2 shows a project where the project manager decided (D1) to switch from terminology T1 to T2 in a first step. After more complexity was encountered within the project, new estimations showed additional need for communication and terminology investments which led to decision D2 and terminology T3.

![Figure 3. Curve of Efficiency for a Specific Project](image)

Figure 3 shows the resulting curve of efficiency based on decisions of the project manager for this theoretical and exemplary project. As the project passes a complexity threshold, the project manager gets evidence whether his or her decision was right or wrong. Anyway, s/he does not know for sure at all, because s/he had no chance to observe the real cost curve s/he used before s/he made his or her decision (based on estimation). Obviously, taking into account the previous remarks, there is a lot of subjective appreciation and perception of the project manager in the model. However, this is precisely the reason why this model should be able to explain and predict project observations in real situations (which, in contrast to laboratory situations, are characterized by high variety, ambiguity, and uncertainty) very well. It should become clear why project managers act and decide as they do with regard to communication efficiency.

3.4 Explicit Consideration of Constraints

Adding to this basic efficiency model the constraints for the project manager are addressed in this section to completely incorporate all influences identified in section 2. The basic constraints for this model are the fixed benefit and the time limit to finish the project successfully.

The fixed benefit as a background factor for the project’s complexity might change over time while the scope of the project is adjusted. This is normally done by bringing change requests into a project. Changes of scope (e.g., due to external perturbations) can have a direct impact on the resulting curve of efficiency by changing the available or (subjectively perceived from the project manager’s point of view) possible cost curves. The fixed time limit (deadline) is a more important constraint for a FDWH project. As financial companies today rely on IT to a very high degree (Guo, Tang, Tong and Yang 2006) and the data warehouse is a central system with a lot of interdependencies to core banking systems and distributive systems, a precise timing is needed for the release cycle of a FDWH. Postponing of FDWH releases has high impact on the whole IT infrastructure of a bank since release cycles are highly interrelated.
Figure 4 adds this time constraint to the efficiency model. As there is no time dimension included in the model yet, the time has to be aligned to each possible way a project manager can chose to go. A fixed duration for a time period is defined and every time a period ends a dot is put on a curve of efficiency. By changing from one curve to another ways are split and each way has its own remaining distance until the end of the next time period is reached. As projects have to deal with a general time constraint, only a predefined number of spots should be passed by the project to be successful. The following paragraph contains a short example of how the time constraint can have impact on management decisions.

Given a time limit of three time periods, every end of a time period is represented by a black dot. After taking decision D1, the project manager has to decide whether s/he stays on the cost curve associated with terminology T2 or makes a decision D2 to change to the cost curve of terminology T3. Given an (subjectively) estimated degree of complexity for the project of a little more than C2, the project would reach the fixed time limit in period three before the estimated degree of complexity for the project is reached. That is, the project would be overdue. Although switching to T3 would be more efficient regarding total cost of communication, the time constraint limits possible decisions for staying with the cost curve of terminology T2. To state it more simply, the shift from one cost curve to another does not only come with the price of investing resources such as project workers, but also consumes time in itself.

![Figure 4. Curves of Efficiency and Time Constraint](image)

In addition to the constraints discussed above, projects can have a cost constraint as well. The project sponsor usually might announce a cost limit or the project is done for a fixed price, so the project manager has to meet a cost limit. This cost constraint is added in Figure 5. Enhancing the previous example of Figure 4, the project manager now has to decide whether s/he wants to violate the time or the cost constraint. S/he can keep the cost constraint by switching to the cost curve of terminology T3 and choosing curve of efficiency E1.

3.5 Explaining Special Cases

There might be situations where the predefined complexity cannot be reached by the current cost curve. Looking at the example of Figure 4 in Figure 6, this would be the case if the estimated project degree of complexity becomes too high to be handled by the cost curve of terminology T2. There would be no other way than switching to the cost curve of terminology T3 to be able to reach the project goal defined by the project sponsor. In that case, the time constraint would need to be violated. The curve of efficiency E1 would be the result of the project manager’s decisions.
Another phenomenon having impact on curves of efficiency is reusability. If the investment in terminology building phases does not only help the current project but also other activities of a company, this reusability effect can influence the curve of efficiency. Therefore the project manager has to estimate to which extend the costs can be split up between costs that can be directly allocated to his or her project and costs that can be allocated to other projects or general company cost centers. This might decrease the total cost of communication for the terminology building phases (i.e., directly after decision D1 and D2). The cost curves of terminology T2 and T3 would have to be shifted down in that case.

3.6 Limitations

The efficiency model with its different cost curves is only meaningful and applicable if one of the following situations occurs:
- The actual estimated project degree of complexity is higher than the project degree of complexity estimated before.
The scope of the project was changed and new estimations for the curve of efficiency are needed to adapt the project plan. The curve of efficiency has to be adapted due to possible time constraint violations becoming clear by actual estimations. The curve of efficiency has to be adapted due to possible cost constraint violations becoming clear by actual estimations.

Therefore the efficiency model is only suitable for complex projects like FDWH projects. Simpler projects without change of scope or with limited complexity will not need to be reviewed and adapted as often. So the efficiency model is only suitable to explain and predict management decisions in complex projects depending on communication efficiency. Another limitation is that the whole approach is based on subjective understandings of the involved project members and project managers. As project managers rely on information coming from the project team, the project team evaluates itself. Furthermore the approach is based on the precondition that project managers work in line with the goals of the project sponsor and no other interfering goals exist.

Although the model shows only a few possible cost curves in the examples, there is an infinite number of cost curves due to the huge set of possible terminologies. Each term added to a terminology produces a new terminology. So possible terminologies and their corresponding cost curves for a specific project lead to a floating intersection of cost curves for this project. However, the project manager cannot have all of these possibilities in mind. It depends on his or her subjective awareness of different terminologies or approaches to gain them. Most times the project manager will not be able to define concretely how a new terminology should look like, but makes some kind of best guess what terminology and corresponding cost curve might be reachable. That brings project managers to just define how much costs are invested and to hope that a suitable cost curve is reached. This once again makes clear that project management has to do with experience and a feeling for efficiency. As far as the model is developed, it is only applied to the financial data warehouse context. Further research should focus on generalizing the model to other complex ISD projects and implications for less complex projects.

4 CONCLUSION & OUTLOOK

A theoretical model for communication efficiency in FDWH projects was derived based on a theory of communication costs and project management fundamentals. The model is based on exploratory research of the last years and should be able to explain and predict identified specifics of FDWH projects.

The model suits the conditions for theoretical models given by Lee (1989) based on Popper (1965):

- **Falsifiability**: the theoretical model is presented in a way that allows testing against the theoretical approach by conducting empirical studies focusing on it.
- **Logical consistency**: the inherent mechanisms of the model are explained in a step-by-step way to assure logical consistency. Additionally, special cases and common project situations were stressed.
- **Relative explanatory power**: the theoretical model explains FDWH project observations in a more sophisticated and detailed way than the introduced general communication cost model or the introduced general software project management classification.
- **Survival**: since the first idea for a cost-based approach to FDWH in the beginning of 2007, no better applicable model could be found by our research group.

As all criteria are met for the moment, further research steps can be taken to corroborate the theoretical model for FDWH projects in real settings. Currently, additional research focuses on testing the theory in different FDWH project situations in multiple case studies (Räkers and Rosenkranz 2009) and future research will try to apply the theoretical approach to ISD projects beside financial data warehousing as well.
To sum up the findings and formulate theoretical suggestions coming from the efficiency model for FDWH projects, there are already practical implications for the behavior of project managers. They should keep in mind that communication costs are one of the most relevant drivers for project efficiency, and that the complexity of the project has to be judged and estimated throughout the project to stay on the most efficient curve. Due to the characteristics of communication cost functions, investments in terminology should be done as soon as a need for a higher estimated degree of complexity for a project is identified. Practically, this means to engage team members more into face-to-face communication when new variety enters the project environment or becomes visible.

References


## CONTINUANCE USAGE INTENTION IN MICROBLOGGING SERVICES: THE CASE OF TWITTER

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0164.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Partial Least Squares, Social Networking (e.g. Facebook, second life), Technology Acceptance Model (TAM), Structural Equation Modeling</td>
</tr>
</tbody>
</table>
CONTINUANCE USAGE INTENTION IN MICROBLOGGING SERVICES: THE CASE OF TWITTER

Abstract

The emergence of Web 2.0 has brought with it a plethora of social networking technologies. Among these, microblogging has emerged as a new and popular tool for short, frequent communication via Web postings. The most popular microblogging service, Twitter, has established a large user base, in spite of numerous criticisms. This study aims to examine why this is the case. In particular, the study develops a model of microblogging use continuance based on theories of continuance, habit and critical mass. The model is then tested by means of a Web survey of Twitter users (n=131) and structural equation modelling using the partial least squares technique. The results suggest that continued use intention is strongly determined by perceived usefulness, satisfaction and habit ($R^2=0.454$) which together provide a strong explanation for Twitter user behaviour. Further, critical mass and frequency of prior behaviour, both influenced by social network size, are strong determinants of the habit construct. The paper rounds off with conclusions and implications for future research and practice in this very new area of inquiry.

Keywords: microblogging; microblogs; continuance; critical mass; habit; Twitter.
1 INTRODUCTION

Dramatic changes are afoot in how users interact using the World Wide Web. The concept of ‘Web 2.0’ or ‘social software’ has become used to describe revolution-like developments associated with Internet applications like wikis, weblogs and social networking software. Although the term ‘Web 2.0’ was first introduced around 2004, the development process has been a continual one, spawning ever new types of software. One of the newest developments in this area is that of ‘microblogging’. As the technology is so new there are even several opinions about its name. In addition to ‘microblogging’ terms like ‘microsharing’ and ‘activity streaming’ are also typically used.

Notwithstanding the strong adoption of microblogging by private users and enterprises there is, as yet, very little academic research on the topic. Microblogging in general and the currently most popular microblogging service Twitter (www.twitter.com) in particular are presently among the most discussed Web 2.0 tools in the ‘blogosphere’ and in spite of substantial positive arguments many users complain about a lack of functionality or reliability issues. Considering the fact that Twitter use is nevertheless not declining but sharply rising there must be strong reasons why users continue using the service. Discovering an explanation for this finding could help us to understand the processes in microblogging platforms as well as to provide implications for the successful future implementation of microblogs in enterprise contexts.

The rising adoption of Twitter is a perfect setting for research in IS continuance theory. The purpose of our work is to understand the continued use of the microblogging service by its general users. The research question to be answered is “Why do users of the Twitter microblogging service continue to use the platform?” For this purpose we developed a research model using continuance theory (Bhattacherjee 2001) and related theories of habit (Limayem et al. 2007) and critical mass (Van Slyke et al. 2007). Data was collected via a Web survey of Twitter users. The data was then analysed using structural equation modelling via the partial least squares technique (Chin 1998).

The structure of the paper is as follows. In the next section we provide some background on the nature and growth of microblogs. In the following section we discuss the theoretical development of a research model of continuance in microblogging services. In the third section we describe the study’s methodology before presenting the results of the analysis in section four. Finally we discuss the findings and their implications for research and practice in the conclusions.

2 BACKGROUND

Microblogs are a smaller version of weblogs combined with features for social networking and mobility. Users post short updates without a headline or additional information in their microblog. Other members can add them to their social network (i.e. ‘follow’ them). The messages from a member’s social network appear in a chronologically ordered and combined view on their starting page. Most microblogging services limit the number of characters used in a posting to 140 or similar. The goal is to animate users to post short messages often in their microblogs. Access to microblogging services is also possible using mobile text messages, desktop clients or several third party applications.

The first and currently the most popular microblogging service is Twitter. Having earned an estimated US$20 million of venture capital (Crunchbase 2008) the service has seen an exponential growth in its user base since launching in July 2006 to an estimated 3 million users in September 2008 (Twitter Facts 2008). Twitter shapes the understanding of microblogging by asking the user “what are you doing”. Around Twitter a whole ecosystem of third party applications have been developed and much of the microblogging vocabulary has its origins in this particular service, e.g. the verb “tweet” for the action of posting into a microblog. In the U.S. Twitter appears to have reached critical mass and is becoming common sense. For instance, President Barack Obama used Twitter massively in his
election campaign. As Twitter has become perceived as a standard many so-called Twitter-clones have emerged, in part due to the relative ease in imitating the software. Serious competitors such as Pownce, Plurk and Jaiku tried to differentiate from Twitter by offering wider functionality or slightly different approaches. Google’s acquisition of Jaiku (in 2007) and Facebook’s offer for Twitter worth US$500 million in shares (in 2008) demonstrated the strategic importance of microblogs. Eventually, a first rationalisation of the market led to the closure of Pownce in December 2008. Notwithstanding, new competitors entered the market in 2008 with open source microblogging software like laconi.ca; different installations of this program communicate with each other and contribute towards building a big peer-to-peer microblogging network.

Gartner (2008) added microblogging to its hype cycle in 2008 and predicts that the technology will garner a sharp rise in popularity. This was also a consequence of the fast growing debate on using Twitter-like tools in professional contexts. While Twitter itself can be an interesting channel for companies to develop their brands and improve their customer service, many argue that due to privacy regulations there should be such a tool behind the company firewall. Major companies such as IBM, Oracle and SAP experimented with so-called enterprise microblogs and in the second half of 2008 special enterprise microblogging products like Communote, Presently and Yammer were launched.

3 THEORY AND RESEARCH MODEL

In this section we present prior research and concepts that have informed the development of the research model tested in the study. Three main strands of theory have contributed to theoretical development: continuance theory as the basic theory of our investigation complemented with concepts of habit and critical mass. Let us briefly discuss each of these areas in turn.

3.1 Continuance theory

Continuance theory within the discipline of information systems stems from initial research in marketing. Expectation-confirmation theory (ECT) emerged from the consumer behaviour and services marketing literature and has proven broadly robust in a number of service contexts (Dabholkar et al. 2000, Oliver 1993). The general thrust of ECT is the assessment of post-purchase intentions, as influenced by initial expectations about a product or service, subsequent adoption and use (consumption) and the formation of perceptions about performance as influenced by the confirmation or not of initial expectations, the latter determining the level of satisfaction with a purchase and subsequent repurchase or use discontinuance. Bhattacharjee (2001) was the first to fully formalize the theory into an ex-post framework that could be applied to the domain of information systems, adapting the theory to be applied post-acceptance and to encapsulate perceived usefulness (Davis et al. 1989, Davis 1993) as a replacement construct for expectations. Perceived usefulness has consistently proven to be an important construct in longitudinal adoption to post-adoption behaviour (Davis et al. 1989, Karahanna et al. 1999). Thus Bhattacharjee’s (2001) model relates the constructs of perceived usefulness and satisfaction to the extent of confirmation of a user’s expectations about an IS, whereby expectations that are fulfilled drive greater satisfaction and perceived usefulness. High levels of perceived usefulness are also posited to lead to greater satisfaction with a system. In turn, the outcome variable of continuance intention is determined by the level of satisfaction with an IS and the perceived usefulness of the system.

Bhattacharjee’s model has been successfully applied to individual user contexts involving the Web, such as online banking, and the Internet more broadly (Bhattacharjee 2001, Limayem et al. 2007). Microblogging services are also distributed systems reliant on Internet technology, and the focus of this study is on individual users; Bhattacharjee’s theory (which appears to have broader application and generalization in any case) is adopted as a suitable core for a model examining continuance behaviour in this context.
3.2 Habit

Habit refers to the extent to which behaviour has become automatic as a result of prior learning (Limayem et al. 2007). Previous research has found a strong relationship between habit and continuance behaviour. However, the use of habit in research models is more complex. Research has variously examined habit as a moderator between intention and actual behaviour, as a direct effect on actual behaviour, and as an indirect effect on behaviour that primarily determines intentions. Our focus in this research is on use intentions rather than actual behaviour, and so naturally we focus on the latter of these formulations. Our focus on intentions is in line with a core body of previous IS literature (Legris et al. 2003), whilst the focus on indirect habit effects is a view that is held in a number of previous studies that have examined the effects of habit (and the much used proxy construct of experience) on behavioural intentions (Bagozzi & Warshaw 1990, Legris et al. 2003).

In addition to the effect of habit on intention to continue using an IS, we also posit that habit is significantly influenced by satisfaction. Limayem et al. (2007) in their comprehensive definition, application and analysis of the habit construct in continuance theory find very strong support for the linkage between satisfaction and habit. Further, they find that frequency of prior usage and comprehensiveness of usage both have a positive associations with habit, both assertions that we also make for the purposes of model development.

The creation of habit requires a stable context conducive to its formation through repetition or practice (Orbell et al. 2001); we would hold that such a context exists when focusing on individuals’ behaviour with respect to a single system such as a microblogging service. This position is in line with that of Limayem et al. (2007) in their study of habitual use of the Internet.

3.3 Critical Mass and Perceived Critical Mass

Critical mass in technology adoption has been discussed in the literature for many years (e.g. see Markus 1994; Rogers 1995). Some of the seminal work in this area was that of Rogers (1995) as part of his diffusion of innovation theory. Rogers (1995) defined critical mass in terms of a ‘tipping point’ whereby a certain minimum number of users have adopted an innovation which then feeds into rapid continued adoption of the new technology, at which juncture further adoption is self-sustaining.

Further research into the application of critical mass theory to electronic communication media has demonstrated that universal access to a communication medium is major driver for critical mass; the decision to use a particular interactive media for routine communication relies on others in a communication network having adopted it and an overall social consensus (Markus 1994). Social context thus envelops individuals’ choices and “As more and more individuals in a system adopt an interactive communication innovation, the innovation is perceived as increasingly beneficial to both previous and potential adopters” (Van Slyke et al. 2007, p 274). As a result, positive network externalities ensue as a result of technology use (Lou et al. 2000, Van Slyke et al. 2007).

Van Slyke et al. (2007) examined perceived critical mass within the context of instant messaging, which has similarities to microblogging. They applied diffusion of innovation theory and found direct effects on numerous variables, including social norms, perceived relative advantage, perceived compatibility, ease of use and behavioural intention. As part of continuance theory we focus on continuance intention; however, in our habit model we posit that the effects of critical mass will manifest themselves indirectly through automatic behaviour. Thus, based on Markus (1994), an individual will exhibit greater habitual behaviour and choose a technology for interactive communication if there is a greater perception of critical mass. Such a perception will be driven, at least in part, by the size of an individual’s network – the extent of their personal communication network on the microblogging service. Further, the size of a network is posited to play a role in determining the frequency and comprehensiveness of an individual’s behaviour: the greater the extent of the network, the greater possibilities there are for communication, coordination and cooperation, and the more activity that will take place.
3.4 Research Model and Hypotheses

A combination of all of the above theorisations into a single model results in Figure 1. Here we use Bhattacherjee’s (2001) continuance theory as basic core of our research model (marked white in Figure 1). This leads to our first basic hypothesis:

**H1: Bhattacherjee’s continuance theory can be applied to explain Twitter usage.**

Further, we extend the model with the construct of habit and the relationship from satisfaction (Limayem et al. 2007). In contrast to Limayem et al. (2007) however, we model it as a direct effect on continuance intention:

**H2: Satisfaction has a direct effect on habit.**

**H3: Habit has a direct effect on continuance intention.**

We accept the habit indicators of frequency of past behaviour and comprehensiveness of usage introduced by Limayem et al. (2007) and extend these constructs with perceived critical mass due to the property of Twitter as a communication tool.

**H4: Habit in microblogging usage is driven by: (a) perceived critical mass; (b) frequency of past behaviour; and, (c) comprehensiveness of usage.**

As a basic determinant for these three constructs we introduce social network size:

**H5: The construct of social network size is a driver of: (a) perceived critical mass; (b) frequency of past behaviour; and, (c) comprehensiveness of usage.**

![Research Model](image)

**Figure 1. Research Model.**

4 STUDY DESIGN AND METHOD

In this section we briefly outline the process of data collection, scale development and data analysis. Let us examine each of these in turn.
4.1 Data Collection

Twitter is clearly the best known microblogging service on the Web. This is one reason why we chose this platform as the basis for our investigation. Another point is that Twitter has had several downtime and reliability issues during 2008 and most recently additional, smaller competitors have emerged with broader functionality. Even in the face of this, Twitter’s user base experienced exponential growth. It can be supposed that there are strong reasons why existing users continue to use Twitter and new users choose this service over the various alternatives. Finally, there was a practical reason for selecting Twitter. Twitter has an accessible Web API (application programming interface) which enabled us to retrieve data about users and thus gave us the opportunity to measure real behaviour for some constructs. For this reason survey participants were asked to give a valid Twitter user name.

Data was collected via the survey Web site, QuestionPro. Data collection was promoted mainly in Twitter with a few blog postings outside the platform. Promotion was initiated only from one of the author’s Twitter accounts using this network as a starting point for viral marketing and ‘hashtags’ to reach completely different communities of users. Hashtags are usually short terms describing several topics, such as an event (i.e. #smclondon08 for Social Media Camp London 2008). People attending such an event and writing about it on Twitter use this tag in their postings. In searching for this term one is able to get related messages even from people who are not in one’s network. These streams from Web-related events were used to post advertising for the survey. Overall, we received a total of 131 usable survey responses during the 13 days the survey was open (of the 138 originally collected, 7 were excluded due to incorrect user names or closed accounts which prevented us collecting usage data from the Twitter service). Based on the logfiles of QuestionPro, some 26.0% of the participants were from Germany, 18.3% from the US, 16.8% from the UK, 10.2% from the rest of Europe and 6.1% from the rest of the world (with 22.9% unknown). This shows a wide distribution of the survey. In addition, based on the API data, some 45% of the participants registered their specific username on Twitter in the year of the survey (2008) and 47% in the year before (2007). Only 6% were initial adopters in Twitter’s first year of operation (2006).

4.2 Measurement

Scale development was largely based on a strong foundation of scale items from existing literature. The survey items selected and their provenance is demonstrated in Table 1. All but three constructs were measured using traditional 7-point Likert scales. Our constructs were extended with the new construct of social network size and two other constructs were measured in different ways to the existing literature. First, using Twitter’s API it was possible to get data for both the frequency of past behaviour and social network size using a custom software programme developed by one of the authors. Frequency of past behaviour was defined as the two weeks prior to survey completion. Based on the date the survey was completed by each respondent we were able to measure this accurately based on the number of Twitter updates posted on the service. We used two weeks rather than the four weeks of Limayem et al. (2007) because we considered it more appropriate for the nature of this interactive medium. Social network size was based on the complete network of ‘followers’ and ‘following’ connections. Comprehensiveness of usage relied on a new conceptualisation of the variety of usages of Twitter based on a simple three-part categorisation into communication, coordination and cooperation, with each being activated as one-way and/or two-way (six uses in total).

4.3 Data Analysis

Data analysis was performed using a variance maximization approach to structural equation modelling (SEM) and associated statistics for validity and reliability. More specifically, we used the partial least squares (PLS) technique with reflective indicators in Smart-PLS 2.0 (Ringle et al. 2005). The PLS technique has become increasingly popular in information systems research, marketing and in
management research more generally in the last decade or so, influenced by its flexibility; indeed, PLS does not have the same distributional assumptions of normality for data and is able to handle small- to medium-sized samples (Chin 1998, Compeau & Higgins 1995).

<table>
<thead>
<tr>
<th>Construct</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Usefulness</td>
<td>Limayem et al. (2007)</td>
</tr>
<tr>
<td>Confirmation</td>
<td>Bhattacharjee (2001)</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>Bhattacharjee (2001)</td>
</tr>
<tr>
<td>Continuance Intention</td>
<td>Bhattacharjee (2001)</td>
</tr>
<tr>
<td>Habit</td>
<td>Limayem et al. (2007)</td>
</tr>
<tr>
<td>Perceived Critical Mass</td>
<td>Van Slyke et al. (2007)</td>
</tr>
<tr>
<td>Comprehensiveness of Usage</td>
<td>The general idea came from Limayem et al. (2007). However, this was implemented using 6 kinds of Twitter uses (communication, coordination and cooperation, with each being one-way and/or two-way)</td>
</tr>
<tr>
<td>Frequency of Past Behaviour</td>
<td>The number of Twitter postings during the preceding 14 days before survey participation</td>
</tr>
<tr>
<td>Social Network Size</td>
<td>The user’s number of followers and number of friends (‘following’) in Twitter</td>
</tr>
</tbody>
</table>

Table 1. Sample Measures.

5 RESULTS

In this section we discuss the results of model testing, including scale validity and reliability and the results of our partial least squares path modelling analysis.

5.1 Tests for Validity and Reliability of the Measures

Table 2 demonstrates that the scale items exhibit high levels of convergent validity – the extent to which theoretical scale items are empirically related. The loadings of the measures on their respective constructs in the model range from 0.727 to 0.955, with all being significant at the 0.1% level.

Table 2 also demonstrates that all of the constructs fulfil the recommended levels with reference to composite reliability (CR) and average variance extracted (AVE). All items were higher than the cut-off of 0.50 recommended by Fornell and Larcker (1981), ranging from 0.695 to 0.891. Similarly, the values for composite reliability are very good, ranging from 0.842 to 0.942, well above the reliability values of 0.70 and 0.80 that are typically advised for building strong measurement constructs (Nunnally 1978, Straub & Carlson 1989).

Table 3 examines the extent to which question items measure the construct intended or other related constructs, otherwise known as discriminant validity. Fornell and Larcker’s (1981) standard test for discriminant validity was used, whereby the square root of average variance extracted for each construct is compared with the correlations between it and other constructs; discriminant validity is demonstrated if the square root is higher than the correlations. Table 3 clearly indicates that each construct shares greater variance with its own measurement items that with other constructs with different measurement items, with a good margin of difference.

Table 4 provides an additional test for discriminant validity. Here we utilized the cross-loading method of Chin (1998). The method prescribes a requirement for measurement items to load higher on a construct than the scale items for other constructs and for no cross-loading to occur. Item loadings in the relevant construct columns were all higher than the loadings of items designed to measure other constructs; similarly, when glancing across the rows the item loadings are considerably higher for their corresponding constructs than for other constructs.
Overall, the results of testing for validity and reliability are very positive and provide us with a high
degree of confidence in the scale items used in the study.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Item</th>
<th>Loading</th>
<th>Mean</th>
<th>St. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmation (CONF)</td>
<td>CONF1</td>
<td>0.862</td>
<td>5.71</td>
<td>0.103</td>
<td>20.444</td>
</tr>
<tr>
<td></td>
<td>CONF2</td>
<td>0.904</td>
<td>5.63</td>
<td>0.113</td>
<td>51.131</td>
</tr>
<tr>
<td></td>
<td>CONF3</td>
<td>0.724</td>
<td>5.14</td>
<td>0.123</td>
<td>9.636</td>
</tr>
<tr>
<td>Critical Mass (CM)</td>
<td>CM1</td>
<td>0.885</td>
<td>4.33</td>
<td>0.158</td>
<td>21.662</td>
</tr>
<tr>
<td></td>
<td>CM2</td>
<td>0.618</td>
<td>5.28</td>
<td>0.101</td>
<td>6.992</td>
</tr>
<tr>
<td></td>
<td>CM3</td>
<td>0.880</td>
<td>4.15</td>
<td>0.156</td>
<td>23.852</td>
</tr>
<tr>
<td>Habit (HABIT)</td>
<td>HABIT1</td>
<td>0.916</td>
<td>5.49</td>
<td>0.130</td>
<td>51.533</td>
</tr>
<tr>
<td></td>
<td>HABIT2</td>
<td>0.930</td>
<td>5.50</td>
<td>0.127</td>
<td>70.021</td>
</tr>
<tr>
<td></td>
<td>HABIT3</td>
<td>0.819</td>
<td>4.76</td>
<td>0.135</td>
<td>24.200</td>
</tr>
<tr>
<td>IS Continuance Intention (CONT)</td>
<td>CONT1</td>
<td>0.865</td>
<td>5.10</td>
<td>0.116</td>
<td>24.278</td>
</tr>
<tr>
<td></td>
<td>CONT2</td>
<td>0.842</td>
<td>5.93</td>
<td>0.102</td>
<td>26.120</td>
</tr>
<tr>
<td></td>
<td>CONT3</td>
<td>0.867</td>
<td>5.05</td>
<td>0.124</td>
<td>23.480</td>
</tr>
<tr>
<td>Perceived Usefulness (PU)</td>
<td>PU1</td>
<td>0.915</td>
<td>6.02</td>
<td>0.095</td>
<td>40.056</td>
</tr>
<tr>
<td></td>
<td>PU2</td>
<td>0.890</td>
<td>5.92</td>
<td>0.096</td>
<td>21.853</td>
</tr>
<tr>
<td></td>
<td>PU3</td>
<td>0.923</td>
<td>5.92</td>
<td>0.089</td>
<td>31.792</td>
</tr>
<tr>
<td>Satisfaction (SATIS)</td>
<td>SATIS1</td>
<td>0.918</td>
<td>5.87</td>
<td>0.092</td>
<td>41.291</td>
</tr>
<tr>
<td></td>
<td>SATIS2</td>
<td>0.923</td>
<td>5.81</td>
<td>0.092</td>
<td>53.296</td>
</tr>
<tr>
<td></td>
<td>SATIS3</td>
<td>0.842</td>
<td>5.53</td>
<td>0.103</td>
<td>18.985</td>
</tr>
<tr>
<td></td>
<td>SATIS4</td>
<td>0.846</td>
<td>5.61</td>
<td>0.100</td>
<td>18.678</td>
</tr>
<tr>
<td>Social Network Size (SNS)</td>
<td>SNS1 (Followers)</td>
<td>0.933</td>
<td>217.86</td>
<td>24.677</td>
<td>25.364</td>
</tr>
<tr>
<td></td>
<td>SNS2 (Following)</td>
<td>0.955</td>
<td>255.53</td>
<td>30.013</td>
<td>71.194</td>
</tr>
</tbody>
</table>

Note 1: CR = Composite Reliability; AVE = Average Variance Extracted
Note 2: Usage Comprehensiveness (mean=4.11, standard error=24.68) and Frequency of Past Behaviour
(mean=121.58, standard error=14.33) are single-item constructs

Table 2. Psychometric Table of Measurements.

<table>
<thead>
<tr>
<th>Construct</th>
<th>CONF</th>
<th>CM</th>
<th>HABIT</th>
<th>CONT</th>
<th>PU</th>
<th>SATIS</th>
<th>SNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmation (CONF)</td>
<td>0.834</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical Mass (CM)</td>
<td>0.428</td>
<td>0.804</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habit (HABIT)</td>
<td>0.624</td>
<td>0.562</td>
<td>0.890</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IS Continuance Intention (CONT)</td>
<td>0.575</td>
<td>0.420</td>
<td>0.556</td>
<td>0.858</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Usefulness (PU)</td>
<td>0.667</td>
<td>0.505</td>
<td>0.710</td>
<td>0.610</td>
<td>0.909</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction (SATIS)</td>
<td>0.554</td>
<td>0.347</td>
<td>0.536</td>
<td>0.569</td>
<td>0.594</td>
<td>0.883</td>
<td></td>
</tr>
<tr>
<td>Social Network Size (SNS)</td>
<td>0.219</td>
<td>0.286</td>
<td>0.282</td>
<td>0.021</td>
<td>0.288</td>
<td>0.251</td>
<td>0.944</td>
</tr>
</tbody>
</table>

Table 3. Correlations Between Constructs (Diagonal Elements are Square Roots of the Average Variance Extracted).
Table 4. Loadings and Cross-Loadings for Reflective Measures.

<table>
<thead>
<tr>
<th></th>
<th>CONF</th>
<th>CM</th>
<th>HABIT</th>
<th>CONT</th>
<th>PU</th>
<th>SATIS</th>
<th>SNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONF1</td>
<td>0.862</td>
<td>0.269</td>
<td>0.509</td>
<td>0.500</td>
<td>0.520</td>
<td>0.462</td>
<td>0.151</td>
</tr>
<tr>
<td>CONF2</td>
<td>0.904</td>
<td>0.418</td>
<td>0.631</td>
<td>0.455</td>
<td>0.671</td>
<td>0.495</td>
<td>0.184</td>
</tr>
<tr>
<td>CONF3</td>
<td>0.724</td>
<td>0.380</td>
<td>0.396</td>
<td>0.498</td>
<td>0.454</td>
<td>0.427</td>
<td>0.220</td>
</tr>
<tr>
<td>CM1</td>
<td>0.288</td>
<td>0.885</td>
<td>0.456</td>
<td>0.267</td>
<td>0.381</td>
<td>0.208</td>
<td>0.180</td>
</tr>
<tr>
<td>CM2</td>
<td>0.419</td>
<td>0.618</td>
<td>0.454</td>
<td>0.475</td>
<td>0.439</td>
<td>0.335</td>
<td>0.183</td>
</tr>
<tr>
<td>CM3</td>
<td>0.319</td>
<td>0.880</td>
<td>0.434</td>
<td>0.266</td>
<td>0.388</td>
<td>0.286</td>
<td>0.313</td>
</tr>
<tr>
<td>HABIT1</td>
<td>0.617</td>
<td>0.533</td>
<td>0.916</td>
<td>0.482</td>
<td>0.659</td>
<td>0.498</td>
<td>0.319</td>
</tr>
<tr>
<td>HABIT2</td>
<td>0.572</td>
<td>0.471</td>
<td>0.930</td>
<td>0.499</td>
<td>0.670</td>
<td>0.506</td>
<td>0.241</td>
</tr>
<tr>
<td>HABIT3</td>
<td>0.470</td>
<td>0.496</td>
<td>0.819</td>
<td>0.504</td>
<td>0.562</td>
<td>0.423</td>
<td>0.187</td>
</tr>
<tr>
<td>CONT1</td>
<td>0.435</td>
<td>0.420</td>
<td>0.443</td>
<td>0.865</td>
<td>0.488</td>
<td>0.510</td>
<td>0.100</td>
</tr>
<tr>
<td>CONT2</td>
<td>0.579</td>
<td>0.345</td>
<td>0.544</td>
<td>0.842</td>
<td>0.640</td>
<td>0.533</td>
<td>0.174</td>
</tr>
<tr>
<td>CONT3</td>
<td>0.436</td>
<td>0.308</td>
<td>0.418</td>
<td>0.867</td>
<td>0.396</td>
<td>0.397</td>
<td>0.065</td>
</tr>
<tr>
<td>PU1</td>
<td>0.660</td>
<td>0.470</td>
<td>0.673</td>
<td>0.556</td>
<td>0.915</td>
<td>0.583</td>
<td>0.279</td>
</tr>
<tr>
<td>PU2</td>
<td>0.516</td>
<td>0.453</td>
<td>0.630</td>
<td>0.531</td>
<td>0.890</td>
<td>0.516</td>
<td>0.256</td>
</tr>
<tr>
<td>PU3</td>
<td>0.634</td>
<td>0.456</td>
<td>0.633</td>
<td>0.577</td>
<td>0.923</td>
<td>0.518</td>
<td>0.249</td>
</tr>
<tr>
<td>SATIS1</td>
<td>0.549</td>
<td>0.403</td>
<td>0.565</td>
<td>0.555</td>
<td>0.583</td>
<td>0.918</td>
<td>0.246</td>
</tr>
<tr>
<td>SATIS2</td>
<td>0.517</td>
<td>0.318</td>
<td>0.530</td>
<td>0.503</td>
<td>0.581</td>
<td>0.923</td>
<td>0.267</td>
</tr>
<tr>
<td>SATIS3</td>
<td>0.386</td>
<td>0.197</td>
<td>0.341</td>
<td>0.449</td>
<td>0.412</td>
<td>0.842</td>
<td>0.131</td>
</tr>
<tr>
<td>SATIS4</td>
<td>0.481</td>
<td>0.277</td>
<td>0.420</td>
<td>0.494</td>
<td>0.496</td>
<td>0.846</td>
<td>0.224</td>
</tr>
<tr>
<td>SNS1</td>
<td>0.168</td>
<td>0.270</td>
<td>0.261</td>
<td>0.014</td>
<td>0.269</td>
<td>0.247</td>
<td>0.933</td>
</tr>
<tr>
<td>SNS2</td>
<td>0.239</td>
<td>0.270</td>
<td>0.271</td>
<td>0.025</td>
<td>0.275</td>
<td>0.230</td>
<td>0.955</td>
</tr>
</tbody>
</table>

5.2 Test of the Research Model

The results of PLS path modelling are shown in Figure 2. For this analysis we utilised the software package, Smart-PLS (Ringle et al. 2005). The shaded items are those that have been added to Bhattacherjee’s (2001) basic continuance model. A power analysis in G*Power 3.0 (Faul et al. 2007) shows that the sample size (n=131) has good power for explaining medium population effects ($f^2=0.15$; $\alpha=0.05$; $1-\beta=0.88$), and is thus suitable for the testing of the model under these conditions.

All relationships in Bhattacherjee’s (2001) original model are strongly supported by the data (H1), with links from confirmation to perceived usefulness ($R^2=0.445$) and satisfaction at the 0.1% and 1% levels of significance respectively. Similarly, perceived usefulness is strong driver of satisfaction ($R^2=0.397$; $p<0.001$). Both perceived usefulness and satisfaction are strong drivers of Twitter users’ continuance intentions at the 1% levels of significance. Together with habit, which provides a smaller but nevertheless significance influence (H3; $p<0.05$), some 45.4% of variance in Twitter use continuance intentions is explained by the model.

In line with the findings of Limayem et al. (2007), we find that satisfaction is a major determinant of the manifestation of automatic behaviour (H2; $p<0.001$). Along with strong linkages from the frequency of past behaviour (H4b; $p<0.05$) and perceived critical mass (H4a; $p<0.001$), these variables account for a considerable 47.4% of the variance of habit. We consider this to be a very strong explanation in statistical terms. The linkage between comprehensiveness of usage and habit is not supported. This finding corresponds with additional comments given by the survey participants suggesting that there is a wide range of possible usages but that one particular use can be enough to create value for a user. The comments included: “I use twitter as a learning tool”, “We use Twitter to feed in blog posts from our blog network” and “I use Twitter mainly for entertainment purposes & social networking.” Together these give an impression of different usage possibilities and a possible
explanation why comprehensiveness of usage may not be important in creating habit in the microblogging context.

Further, we find that social network size exerts a strong influence over perceptions of critical mass for Twitter (H5a; p<0.001), the frequency of previous usage of the service based on the last 14 days (H5b; p<0.01), as well as comprehensiveness of usage in our model (H5c; p<0.01), although the latter does not link further to habit.

6 DISCUSSION AND CONCLUSIONS

This paper has tested an extended research model of continuance behaviour in the context of the most popular microblogging service, Twitter. To our knowledge this is the first academic research paper that has examined use continuance behaviour with respect to this very new social networking tool. The results suggest that the continuance behaviour of Twitter users is strongly determined by their perceptions of value of using the service, satisfaction manifested in their usage of the tool and prior development of automatic behaviour or habit surrounding the use of the service (H1 and H3). Further, we find that social network size influences the development of past use behaviour and of perceptions of critical mass, as users see the value of interacting with a significant group of nodes in their network (H5). In turn, habit is driven by satisfaction (H2), past usage behaviour and the aforesaid perceptions of the significant size of the network via perceived critical mass (H4). However, one link is not supported – between comprehensiveness of usage and habit – which leads to the partial falsification of hypothesis 4 (H4c).

Although the findings are both original and significant, there are a number of limitations of note in the research. In using the Twitter API for retrieving data about the participants there is a clear advantage in collecting direct measurements. However, the obverse is that there could be the possibility of ‘missing’ use data. We used each user’s number of Twitter postings during the 14 days before survey participation as a measurement for past behaviour. Notwithstanding, there are Twitter usages such as
plain reading of other’s postings or building a social network which are hidden from the API. Although posting updates on Twitter is seen as the primary usage it would be interesting to see further studies comparing individual perceived frequency of past behaviour with measured posting behaviour. In this way the importance of passive usage could be discovered.

Linked with this limitation is an issue that emerged during the study. We used six categories (communication, coordination and cooperation, each one-way and/or two-way) with real-world examples to capture comprehensiveness of Twitter usage. Nevertheless, a number of comments from respondents suggested introducing additional usage classifications such as work/personal, seminars/events or mobile/desktop. Clearly the development of a comprehensive, valid and reliable typology of Twitter usage is a matter for future research.

Another limitation of the study is the sample size. Although the sample has good explanatory power for medium-sized effects ($f^2=0.15; \alpha=0.05; 1-\beta=0.88$), a larger sample could help to reveal smaller population effects, which may for example uncover missing relationships. A further limitation of our research model is the absence of demographic and other factors, such as age, gender, experience with Twitter, use intensity and so on. Since we may have a heterogeneous sample of a global nature; capturing and analyzing information on the sample may help to understand the behaviour of different groups of global consumers.

From a practical perspective, our research provides useful guidelines not only in terms of understanding public microblogging users, but in terms of other nascent applications of microblogging. Enterprise microblogging or “Twitter for the enterprise” is currently a very hot topic of discussion among Web 2.0 developers. Our research model provides practitioners with some advice regarding not only how to launch a microblogging service successfully but also how to make sure that users continue to use it. In addition to the core influences of perceived usefulness and satisfaction the strong relationship between habit and continuance intention could be an important starting point. Practitioners need to design high quality microblogging services that provide clear benefit and that provide a satisfying experience, but that also become permeated within users’ daily routine activity to such an extent that usage becomes habitual. Incentives to build a core of users may be valuable in creating a self-sustaining body of users and in building perceptions of universal access. Clearly organisations will need to expend considerable effort in trying to create habit. It is not only important to find good initial use cases and incentives to create past behaviour but also to provide perceived critical mass. The corollary of this is interesting; this implies that there is actually no need to start a microblogging service within a whole organisation as long as there is perceived critical mass among the current set of users. A good starting point might be a single department or a single subsidiary. In our research, no relationship was detected between comprehensiveness of usage and habit, implying that there is no need for a wide range of uses of microblogging to build use continuance. One or two strong valuable use cases could be enough to create habit among users.

In conclusion, we perceive that microblogging is clearly becoming an established and much used interactive social medium. This study has shed light on a number of explanations for this phenomenon. We hope that this research paves the way for others to begin investigating aspects of user behaviour surrounding this very new and interesting social networking tool.

References


BEYOND USER ACCEPTANCE: The determinants of THE INTENTION TO PRODUCE USER CREATED CONTENTS ON THE INTERNET

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0455.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Virtual community, Social Networking (e.g. Facebook, second life), User participation, Technology Acceptance Model (TAM)</td>
</tr>
</tbody>
</table>
BEYOND USER ACCEPTANCE: DETERMINANTS OF INTENTION TO PRODUCE USER CREATED CONTENTS ON THE INTERNET

Lee, Habin, Brunel University, Kingston Lane, Uxbridge, Middlesex, UB8 3PH, UK, habin.lee@brunel.ac.uk
Koh, Woonjung, Hanyang University, Seoul, South Korea, nayawoonjung@nate.com
Kim, Jong Woo, Hanyang University, Seoul, South Korea, kjw@hanyang.ac.kr
Ghoneim, Ahmad, Brunel University, Kingston Lane, Uxbridge, Middlesex, UB8 3PH, UK, ahmad.ghoneim@brunel.ac.uk

Abstract

The advance in User Created Contents (UCCs) web sites like YouTube changed the role of Internet users from contents receivers to contents creators; a role which requires more pro-active user behaviour. However, the literature on user behaviour in information technology lacks theories that explain the pro-active user behaviour of producing and sharing UCCs with others on the Internet. This paper aims to reveal the major attributes of Internet users that have a positive impact on the intention to produce UCCs on the Internet. Extant related theories are reviewed to extract major factors of Internet users that lead to the production of UCCs. A questionnaire survey is administrated to 400 sampled respondents in South Korea to test the relationships among the identified factors. The results show that playfulness, self-expressiveness/sharing intention, innovativeness, computing skills and reward have a positive impact on the intention to produce UCCs. In particular, innovativeness turned out to have the biggest impact, while social participation is not a significant factor. Mediator variables such as age, gender and types of UCC also turned out to be playing a role in the causal relationships among the factors and the intention to produce UCCs. A model pertaining to the intention to produce UCCs online is developed and tested. The academic and practical implications of the study are also discussed in details.

Keywords: virtual community, social networking, user participation, technology acceptance.
1 INTRODUCTION

The American weekly newspapers ‘Times’ selected YouTube as the invention of the year in 2006 (Time, 2006). Furthermore, the weekly newspapers ‘The Economist’ envisaged that Web 2.0 and user created contents (UCCs) was the mega-trend of year 2007. There is no widely accepted definition of User Created Contents (UCC). According to Wunsch-Vincent and Vickery (2006) UCCs are “content made publicly available over the Internet reflecting a certain amount of creative effort and which is created outside of professional routines and practices”. A more simplified definition is provided by Chin (2006) who defined UCC as content created by users themselves. Internet users are in the centre of this new trend by creating original and creative text, audio, image and video contents which require much more efforts than traditional Internet activities such as simply posting a reply to a web page. They create and upload their own images, audios, and videos using their mobile phone onto Internet web sites such as YouTube, FaceBook, and Phandora TV. One of the enablers of this trend is the evolution of web 2.0 that aims to facilitate bi-directional interactions between web sites and their users overcoming the one-sided information provision by web sites to their users in web 1.0. As a result, the role of users in web pages development is becoming more important (or proactive) than before as they are the major source of web contents.

Despite the new trend on the behaviour of Internet users, there is no study that aimed to investigate and explore the reasons behind such online user behaviour. Existing research on user behaviour towards Information Technology (IT) adoption have been concentrating on user acceptance (Fishbein and Ajzen 1975; Davis, 1989; Mathieson, 1991; Thompson et al., 1991; Davis et al., 1992; Taylor and Todd 1995; Moor and Benbasat 1996; Venkatesh et al., 2003). For example, Venkatesh et al., (2003) summarise the major factors for user acceptance proposed by previous studies to be: “perceived ease of use”, “perceived usefulness”, “attitude”, “subjective norm”, “extrinsic & intrinsic motivation”, “perceived behavioural control”, “job-fit”, “complexity”, “social factors”, “visibility”, “compatibility”, and “voluntariness of use”. However, user acceptance is passive user behaviour and those identified factors fall short in explaining the proactive behaviour of producing, uploading and sharing digital contents with others on the Internet; a behavior that goes well beyond technology acceptance. Most of the UCC web sites have two different types of users: content creators and content viewers. While extant factors can be employed to explain the behaviour of content viewers, it is necessary to identify the factors which explain the behaviour of content creators. Furthermore, most existing studies focused on IT in both the organizational and academic contexts in which new IT (excel, word processor, task supporting systems) are necessary or very helpful to execute users’ given tasks. As a result, the acceptance of the ITs is mandatory or implicitly mandatory rather than voluntary to users. However, UCC web sites are purely voluntary and the difference in perception is under researched which strove the authors to attempt to fill this gap in the literature.

This paper aims to identify major factors that make Internet users create and share contents on the Internet in a process that requires investing intensive efforts in terms of time and resources. The study undertaken in this paper identifies seven major factors that would explain the behaviour of content creators via a thorough literature review and proposes a conceptual research model. The proposed conceptual model is tested on 400 sampled respondents in South Korea. The results show that playfulness, self-expressiveness/sharing intention, innovativeness, computing skills and reward have positive impact on the intention to produce UCCs. In particular, innovativeness turned out to have the biggest impact in comparison to social participation which does not have a significant impact on UCCs production. The academic and practical implications of the study are also discussed.

The organisation of the paper is as follows. Section 2 proposes the research model of the paper. Section 3 explains the methodology used to test the research model and section 4 presents the results. Lastly, section 5 discusses the implications of the findings and concludes the paper.
2 RESEARCH MODEL

2.1 Factors affecting the intention to produce UCC

Even though existing studies did not differentiate between the intention to use and the intention to produce, it is possible to identify the factors that initiate and affect the proactive behaviour of online content production and dissemination.

Since the 1990s, researchers such as Ahn et al., (2007), Moon and Kim, (2001), and Webster and Martocchio, (1992) focused on playfulness as a factor that explains why people adopt information systems. Barnet (1990) defines playfulness using five generic attributes namely; perceived volunteerism, social volunteerism, physical volunteerism, explicit pleasure and humorous sense. He also found that differences in individuals’ playfulness positively affect the use of computers. The content creators may find it playful during the process of creating exciting contents and this leads to the repetition of the contents’ creating behaviour. As a result, the first hypothesis is derived as follows:

H1: Playfulness has a positive effect on the intention to produce UCCs.

Artists express their ideas, ideology and feeling by producing artistic materials. Similarly, the purpose of producing UCCs can be compared with that of artists. Boneva et al., (2001) differentiates expression based communication from instrumental communication in computer mediated communication research. While instrumental communication is used in general behavioural communication, expressiveness is used in communications to improve mental familiarity and sharing. In this paper, expressiveness is defined as individual consciousness of how much s/he considers UCC as a tool to express himself/herself. Pederson et al., (2003) revealed that expressiveness is one of the major factors for the acceptance of multimedia messaging services (MMS). Otherwise stated, users of MMS use the services as a tool to express themselves to others in order to share their emotions. Using the same rationale, it can be inferred that the producers use UCCs as the tool to express themselves to others in public via the Internet medium. Consequently, the second hypothesis is derived below:

H2: Expressiveness has a positive impact on the intention to produce UCCs.

Most of UCCs are shared within online communities. The users produce UCCs and upload them on the Internet to share with others rather than keeping them locally on their PCs in the form of private diary entries. Butler (2001) and Hare (1976) define the Internet community as a process in which information and knowledge are shared among the community members. Butler (2001) insists that the sharing activities are an important factor to explain the dynamics of the Internet community. From this perspective, one of the reasons of producing UCCs on the Internet can be considered to be motivated by the intention to share UCCs with other community members. The intention to share is defined in this paper as “the intention to provide others with their knowledge or accept others’ knowledge through UCCs on the Internet community”. Based on the preceding the third hypothesis is deduced:

H3: Intention to share has a positive impact on the intention to produce UCCs.

The innovation diffusion theory provides a useful analysis framework to explain how and why new technologies are accepted and used by users (Atkin and LaRose, 1994; Leung and Wei, 1998; Lin and Jeffres, 1998; Li and Yang, 2000; Reagan, 1987). According to innovation diffusion theory (Rogers, 1995), early adopters or pioneers tend to be proactive in accepting new technology or ideology and in taking risks. Hence, innovative people tend to have stronger interests in new ideas or information and consequently have a high probability to buy new products before others (Lin and Jeffres, 1998; Ha 2005; Ban 2007). The adoption of technological innovation is determined by individuals’ innovativeness or volunteerism to try new products or technologies. Innovative people have skills to understand complex technical knowledge, and tend to adopt innovation when it is not widely adopted by other social members. This segment represents those who are young, adventurous, active in...
external activities and pro-active to adopt new ideas in societies (Rogers, 1995). As a result, it is natural to derive the following hypothesis.

**H4:** Innovativeness has a positive impact on the intention to produce UCCs.

Rogers (1995) also relates innovativeness with social participation by noting that people who adopt innovative ideas more than others tend to have a stronger desire to participate in social affairs. In this paper, social participation is defined as the level of interests or actual participation of an individual in social affairs. Human beings naturally tend to be part of society and have a basic desire to be involved in social affairs that may affect their lives. Within the online community literature, social participation (sometimes referred to as ‘social enhancement’ or ‘affection social presence’) is one of the major motivators that triggers people to participate in and contribute to online communities (Dholakia et al., 2004; Shen and Khalifa, 2007). As the number of users of UCC web sites increases, many of the UCCs contain politic or environmental messages. For example, edited video clips of the President of United States were one of the most popular contents of YouTube after the Iraq War in 2003. As a result, it can be deducted that people create and upload digital contents on web sites as a way to participate in social affairs. Therefore, the fifth hypothesis is derived as follows:

**H5:** Social participation has a positive impact on the intention to produce UCCs.

On the other hand, producing UCCs requires users to possess a certain level of computing skills. A computing skill is defined as a competence that enables people to achieve given tasks using computer based information systems or related technologies (Lee et al., 1995). Nelson (1990) asserts that the acceptance of information technology is affected by the level of individual knowledge or skills as well as the technology itself. In the literature, it is reported that computing skills affect perceived usefulness and playfulness (Igbaria et al., 1996; Webster and Martocchio, 1992; Kouparis, 2002). UCCs require certain level of computing skills to edit or create digital contents. It can be inferred that the higher the level of computing skills a user has, the stronger s/he is attracted to create the digital contents. Based on the above, the sixth hypothesis is derived as follows:

**H6:** Computing skills have a positive impact on the intention to produce UCCs.

Finally, Park et al., (2006) report that people share UCCs in order to be famous or make money though this has been investigated only in South Korea. The result is in line with existing studies that identify monetary reward as one of the major motivator of online communications. Kang et al., (2006) measures the economic incentives of UCCs through qualitative metrics. In their study, the interviewees confirmed their satisfaction with the economic reward from the UCCs they produced. This leads to the final hypothesis of this paper:

**H7:** Reward has a positive impact on the intention to produce UCCs.

Figure 1 summarises the proposed conceptual research model based on the derived hypotheses:

![The Conceptual research model on Users' intention to produce online UCCs](image)
2.2 Control Variables

According to the global leader in internet media and market research Nielsen//NetRatings by Wunsch-Vincent and Vickery (2006), men are 20% more likely to visit YouTube than women. Furthermore, of those who actively produce UCCs 86% are males. In the meantime, while Blogging is very popular in countries such as China, India, and Iran, a recent Microsoft survey shows that of all Asian Internet users, women are very active bloggers accounting for 55% of bloggers in Asia with men mostly overwhelming blogging in India (Stern, 2006). There is a clear indication that there are differences in the level of involvement of males and females in UCCs. This level of involvement differs depending on the geographical location as per the aforementioned statistics. As such, in this study gender was introduced as a control variable to measure its effect on Internet users’ intention to produce UCCs in South Korea.

The social drivers of UCCs production such as willingness to share, contribute and to create online communities are altering the media consumption habits of internet users particular within the younger age group of 12 to 17 years old. According to the Nielsen//NetRatings’ (2006) report, UCC has already begun to have an impact on traditional media industries. The statistics show that users aged between
15-24 years of age reduced their consumption of offline media in the UK. This reduction is driven by Internet use (Ofcom, 2006). This implies that out of the various possible age groups to view or produce UCCs 15-24 years old are probably the most potentially effective segment. This claim is also supported by findings in the report that state that that YouTube visitors aged between 12-17 years of age (those who mainly view but might not necessarily have produced UCC before) are the highest among all the various age groups. For those internet users who publish UCCs namely amateur video contents, 75% are under the age of 25. In the light of the above, it is clear that the age group to target falls around the 12-15 years old, hence the authors defined a control variable that would accommodate such a segment namely; 10s and 20s years of age to test the effect of the two age groups on the intention to produce UCCs.

Reflecting on the previous subsection 2.1.2, it is apparent that within such an age segment lay two user categories namely; those who have previously created UCCs and those who are mainly viewers of UCCs with possibly no previous experience in UCCs production. The research distinguishes between these two categories of internet users and introduces a control variable (self / non-self produced UCC) to test the effect of the level of involvement in UCCs on the intention to produce UCCs.

3 METHODOLOGY

In order to test the conceptual research model presented in section 2, a survey questionnaire method was applied. The operationalisation of the variables was conducted via a literature review and are summarised in Table 1. A pilot survey had been performed to test the validity of the questionnaires before they were administrated to the respondents. The sample for data collection was targeted at teenagers and people in their twenties in South Korea as they are the most active users of UCCs (Yim, 2007). Four hundred respondents who had an experience of producing or plan to produce UCC in the Seoul and Kyong-gi province have been sampled randomly for a questionnaire survey. Out of a total of 400 completed questionnaires 389 were used for the analysis after filtering out 11 non-sensible questionnaire responses. Male respondents represent 45.2% of the respondents while 54.8% of respondents were females... The respondents’ ages were distributed as follows: 49.6% teenagers, 46.3% in their twenties, 3.1% in their thirties, and finally 1% in their forties. Most of the respondents had more than five years of Internet experience and about 60% of the respondents were visiting UCC related web sites at least 3 times a week. As a result, most of the respondents had enough knowledge to respond to the questions.

4 DATA ANALYSIS AND RESULT

4.1 Validity and reliability test

To test the reliability of the survey constructs, Cronbach’s alpha value was used. As for the validity test, firstly, the exploratory factor analysis was used to judge unidimensionality of sub variables. Secondly, the principal component analysis using varimax rotation has been employed to extract factors to reduce the number of factors with minimum loss of data. First, the factors whose eigen value is bigger than 1 from principal component analysis were extracted. Then, varimax rotation was performed on the extracted factors. The analysis of variables resulted in 6 factors which were used as measurements of users’ intention to produce UCCs. Among the 7 variables, self-expressiveness and sharing intention have been represented and grouped into one factor based on the fact that users want to show off themselves by sharing their contents to provide others with a pleasure. The detailed data of the factor analysis is not shown here for the conciseness of the paper but the variables were well tied to each other according to each factor and the factor loading values were bigger than 0.6.
A reliability test using Cronbach’s alpha was performed after excluding the invalid variables from the factor analysis. As shown in Table 2, the most of the alpha values are higher than 0.8, hence the collected data can be considered as reliable for the variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>PF</th>
<th>SE</th>
<th>SI</th>
<th>IN</th>
<th>SP</th>
<th>CS</th>
<th>CO</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cronbach’s alpha</td>
<td>0.850</td>
<td>0.531</td>
<td>0.858</td>
<td>0.842</td>
<td>0.926</td>
<td>0.850</td>
<td>0.775</td>
<td>0.898</td>
</tr>
</tbody>
</table>


4.2 The research model test

Multiple regression analysis and ANOVA were employed to test the hypotheses in the conceptual research model presented in Figure 1. Six among seven factors were used as independent variables after merging ‘self-expressiveness’ with ‘sharing intention’ while the intention to produce UCCs was defined as the dependent variable. The result of the regression analysis and the ANOVA results are summarised in Table 3 and 4 respectively.

<table>
<thead>
<tr>
<th>R</th>
<th>R²</th>
<th>Adjusted R²</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.684</td>
<td>0.467</td>
<td>0.459</td>
<td>0.76701</td>
</tr>
</tbody>
</table>

*Table 3  Regression analysis validation*

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum squares</th>
<th>d.f.</th>
<th>Mean square</th>
<th>F</th>
<th>significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression</td>
<td>197.288</td>
<td>6</td>
<td>32.881</td>
<td>55.892</td>
<td>0.000**</td>
</tr>
<tr>
<td>Residual</td>
<td>224.733</td>
<td>382</td>
<td>0.588</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td>422.021</td>
<td>388</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 4  ANOVA results*

From the above tables, the $R^2$ value of the regression equation is 0.467 and the significance level of the ANOVA is 0.05. As a result, the regression equation can be considered valid to explain the dependent variable namely; the ‘intention to produce UCCs’.

(a) According to Table 5 and (b)

Figure 2 (a), five factors apart from social participation turned out to be significant in explaining the intention to produce UCCs without a multi-collinearity issue.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Beta</th>
<th>t</th>
<th>Significance</th>
<th>Collinearity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tolerance</td>
</tr>
<tr>
<td>Playfulness</td>
<td>0.152</td>
<td>2.981</td>
<td>0.003</td>
<td>0.536</td>
</tr>
<tr>
<td>Self-expression/sharing intention</td>
<td>0.118</td>
<td>2.264</td>
<td>0.024</td>
<td>0.515</td>
</tr>
<tr>
<td>innovativeness</td>
<td>0.398</td>
<td>7.998</td>
<td>0.000</td>
<td>0.562</td>
</tr>
<tr>
<td>Social participation</td>
<td>-0.058</td>
<td>-1.234</td>
<td>0.218</td>
<td>0.624</td>
</tr>
<tr>
<td>Computing skill</td>
<td>0.121</td>
<td>2.831</td>
<td>0.005</td>
<td>0.762</td>
</tr>
<tr>
<td>Reward</td>
<td>0.127</td>
<td>2.733</td>
<td>0.07</td>
<td>0.649</td>
</tr>
</tbody>
</table>

*Table 5  Estimated regression equation*
Further analysis was made to investigate the effect of control variables such as gender, age and UCC type on the intention to produce UCCs with the results shown in Figure 2 (b) and Figure 3. According to the results, innovativeness turned out to be the most important factor that affects the intention to produce UCCs regardless the type of contents, gender or age. On the other hand, social participation does not affect the intention to produce UCCs. Figure 2 (b) shows the relationships between the six factors and the intention to produce UCC according to the type of contents. Playfulness, self-expressiveness, computing skills and reward turned out to be significant factors for producing self-produced UCCs, while they do not have significant impacts for non-self produced UCCs. In particular, computing skills and reward were the most significant factors for self produced UCCs which can be explained by the fact that a high level of computing skills is required for self produced UCCs. In addition, the desire to be compensated for their idea and creativeness on the UCCs is one of the major motivator of the self producers. Figure 3 (a) shows that males and females are affected by different factors in producing UCCs. Males are affected by playfulness and reward, while females by self-expressiveness / sharing intention and computing skills. On the other hand, according to Figure 3 (b), the major motivation to produce UCCs for teenagers is playfulness and computing skills while the motivation for twenties are self-expressiveness / sharing intention and reward. Finally, Table 6 summarises the hypotheses test results as discussed above.
It is worth to note that the proposed research model does not include some factors used to explain the intention to accept new ITs in previous studies like social influence, perceived ease of use, perceived usefulness as those factors are strongly related to the passive behavioural intention. Hence they are not appropriate to explain the pro-active behavioural intention to produce UCCs without any external pressure. Lu et al., (2005) report a positive impact of social influence on the acceptance of new IT like mobile services which were provided to users in non-working environments. They argue that users feel that the use of new IT can improve their (social) image which demonstrates an example of how social influence can affect user acceptance. However, the behaviour of producing UCCs is performed in virtual society with anonymity, and the impact of the social image does not play a significant role in the context. As a result, the role of social influences which represents the influences of other people who belong to a user’s social network was excluded from the research model.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Description</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Playfulness has positive impact to the intention to produce UCCs</td>
<td>Accept</td>
</tr>
<tr>
<td>H2</td>
<td>Self-expressiveness has positive impact to the intention to produce UCCs</td>
<td>Accept</td>
</tr>
<tr>
<td>H3</td>
<td>Innovativeness has positive impact to the intention to produce UCCs</td>
<td>Accept</td>
</tr>
<tr>
<td>H4</td>
<td>Social participation has positive impact to the intention to produce UCCs</td>
<td>Reject</td>
</tr>
<tr>
<td>H5</td>
<td>Computing skill has positive impact to the intention to produce UCCs</td>
<td>Accept</td>
</tr>
<tr>
<td>H6</td>
<td>Reward has positive impact to the intention to produce UCCs</td>
<td>Accept</td>
</tr>
</tbody>
</table>

Table 6  The conceptual research model hypotheses test results.

As shown in Table 6, all hypotheses apart from H4 were accepted. In particular, innovativeness turned out to be the most significant factor for producing UCCs. Even though there are studies that highlight the important effect of personal innovativeness on user acceptance (Lu et al., 2005; Agarwal and Prasad, 1998a; Agarwal and Prasad, 1998b), the result in this paper revealed that the innovativeness has the strongest impact among others on the acceptance of purely voluntary and computing skill demanding IT processes such as UCCs. In their studies, innovativeness was found to be a moderator rather than a direct determinant of the intention to use new ITs. Such a difference can attribute to the fact that the role of users in ITs in general is still passive compared to being pro-active as demonstrated in UCC web sites.

The acceptance of H5 needs to be interpreted in relation to existing studies that report the irrelevance of computing skills to IT acceptance. Computing skills have not been considered as one of the major factors in the acceptance of IT in the literature. The main reason being the high level of user knowledge of IT as a result of the widespread of computers and Internet technology. As a result Igbaria and Nachman (1990)’s proposition that Information Systems (IS) satisfaction is affected by computer use was considered not longer valid in modern society. However, the results in this study showed that computing skills significantly affect the intention to produce UCCs. This is because UCCs production is a more complex task than using other back office systems as such production requires more advanced computing skills to create and edit videos and pictures using digital devices and software.

On the other hand, the motivation for social participation turned out to be an insignificant factor for the production of UCCs. This is due to the nature of the respondents who are mostly teenagers or in their twenties who normally do not have strong interests in politics and social affairs (Kimberlee, 2002).

The role of mediation variables such as age, gender and UCC type turned out to be significant in explaining the relationships between the factors and the intention to produce UCCs. In particular, the identified factors can explain the intention to produce purely self-produced UCCs, but not non-self
produced UCCs. This implies that the identified factors have more to do with proactive user behaviour because self-produced as opposed to non-self-produced UCCs require more effort. With regards to gender differences, the results are in line with existing studies that reported gender differences in user acceptance of new IT (Venkatesh et al., 2003). In this study, females are more sensitive towards computing skills to produce UCCs; which again is in line with existing studies that report that in general, females have less computer experience and have a negative attitude towards accepting new IT (Schumacher and Morahan-Martin, 2001). Furthermore, playfulness turned out to have a significant impact only on males’ intention to produce UCCs. In other words, males produce UCCs as a tool to play, while females do not perceive it in this perspective. This result concurs with existing studies by Phillis et al., (2002) that report gender difference on electronic games. With regards to age, it was found that teenagers tend to be more sensitive to computing skills than those who are in their twenties as teenagers are usually not well trained on IT as much as twenties. Also, those who are in their twenties consider compensation is a more important factor than playfulness, which is not the case for teenagers. Finally, innovativeness turned out to be significant while social participation is insignificant. Both facts are regardless of differences in age, gender or UCCs’ type.

The paper delivers a number of contributions to both theory and practice. Firstly, to the authors’ knowledge, this is the first effort to identify the major factors that affect Internet users’ intention to produce UCCs. A process that requires more proactive user behaviour than that of just accepting new information technology. In particular, innovativeness turned out to be the most significant factor to explain this proactive user behaviour. Secondly, this paper joins existing studies that highlighted the role of mediator variables namely: age and gender in user acceptance literature from proactive user behaviour’s perspective. Thirdly, from a practical perspective, managers of UCC web sites would benefit from the research model when setting up their strategies as to focus on innovativeness and playfulness to enable a sustainable production of UCCs by their users.

The goal of the paper was to identify major factors that affect the intention to produce UCCs and as a result it lacked the scope to focus on the components that explain the internal users’ process in producing UCCs on the Internet. As an example and a potential future research, structural equation modelling (Kline, 2005) would be a good approach to reveal the procedural structure among the factors identified in this paper, along with any potential external factors in future studies. The implications of the study are self-evident. Firstly, the contextual variable that determines how a new IT is introduced in a society or an organisation affects the acceptance of the technology. As shown in this study, the UCCs are typical example of voluntarism, and the innovativeness plays the most significant role in the user acceptance in this context.

References


Proceedings ECIS 2009


Li, S.S. and Yang, S.C. (2000). Internet shopping and its adopters: examining the factors affecting the adoption of Internet shopping. Paper Presented at the 35th Anniversary Conference by the School of Journalism and Communication at the Chinese University of Hong Kong, Hong Kong.


INTENTION TO LEARN IN MMOG: EXAMINING THE ROLES OF PEER INTRINSIC AND EXTRINSIC MOTIVATIONS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0696.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Collaboration, Learning, Virtual world, Technology Acceptance Model (TAM)</td>
</tr>
</tbody>
</table>
INTENTION TO LEARN IN MMOG: EXAMINING THE ROLES OF PEER INTRINSIC AND EXTRINSIC MOTIVATIONS

Kong, Joseph Siu-Lung, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong, slkong@student.cityu.edu.hk

Kwok, Ron Chi-Wai, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong, isron@cityu.edu.hk

Abstract

Massively Multiplayer Online Game (MMOG) is a unique categorization of electronic game which allows thousands of players to play simultaneously through the Internet in the same virtual environment. A number of researchers have started to introduce the use of MMOG as a new generation of educational platform, allowing players to interact and to learn together through collaborative game-play. However, the answers for the occurrence of collaborative learning behaviour and the motivational drivers for learning collaboratively in a MMOG are still under-researched.

Motivated by such concerns, this study tests a theoretical model to explain individual’s intention to learn by peer motivations. The model employs motivational theories to propose two external motivational factors, namely peer intrinsic motivation and peer extrinsic motivation, and investigates effects of the two new constructs on MMOG players’ intention to learn individually and intention to learn collaboratively by building on the cognitive learning theory.

Based on a study with 94 valid current MMOG player responses, PLS analysis shows that peer intrinsic motivation to play has a significant positive influence on the intention to learn collaboratively, while peer extrinsic motivation to play has a significant positive influence on the intention to learn individually. The results of our findings indicate potential implications to researchers, educators and game developers.

Keywords: Collaboration, Learning, Virtual world, Technology Acceptance Model (TAM)
1 INTRODUCTION

This study aims to bring together Massively Multiplayer Online Game (MMOG) and technology-based collaborative learning streams of research to examine peer motivational factors influencing intention to learn that is under-researched so far. As a hedonic information system, MMOG is a unique categorization of electronic game. By definition, “MMOGs are highly graphical 2- or 3-D videogames played online, allowing individuals, through their self-created digital characters or ‘avatars,’ to interact not only with the gaming software (the designed environment of the game and the computer-controlled characters within it) but with other players’ avatars as well. These virtual worlds are persistent social and material worlds, loosely structured by open-ended (fantasy) narratives, where players are largely free to do as they please – slay ogres, siege castles, barter goods in town, or shake the fruit out of trees” (Steinkuehler 2004). A growing number of researchers have started to introduce the use of MMOG as a new generation of educational platform (e.g. Gee 2004, Childress and Braswell 2006, Mikropoulos 2006), allowing players to interact with each other, and so as to learn together through collaborative game-play. However, the answers for the occurrence of collaborative learning behaviour in MMOG, and the motivational drivers for learning collaboratively in a MMOG environment are still under-researched.

To address this under-researched area, the objectives of this study are to examine the occurrence of collaborative learning behaviour and motivational drivers in MMOG by extending motivation to play to the peer level, bridging the characteristics of MMOG with peer motivations, and investigating the effect of peer motivations towards players’ behavioural intention to learn in the context of MMOG.

In the next section, we present the literature review. Then we discuss the research model, hypotheses and methodology, followed by the discussion of the results, implications, limitations and future research directions. The conclusion of this study is presented in the final section.

2 REVIEW OF LITERATURE

The first stream of reviewed research is on MMOG game-based learning which believes MMOG is a new type of computer-mediated communication application and the unique characteristics of MMOG may trigger learning behaviours of game players in the gaming context. The second stream of reviewed research is on the explanation of relationships between motivational factors and intentions to play/use/learn in a gaming context by Technology Adoption Model (TAM).

2.1 MMOG game-based learning

MMOG is a new type of Computer-Mediated Communication (CMC) application, with recreational and playful context that are designed to support dynamic social structures (Pena and Hancock 2006). MMOG consists of several unique characteristics that may trigger learning behaviour in the gaming context:

- **Avatar / Virtual Identity**: The existence of an avatar is the player’s representation which enhances the player’s sense of identity and presence, and helps the player to successfully perform the learning tasks (Mikropoulos 2006). The use of avatar in MMOG aligns with Gee’s (2004) example of virtual identity in gaming environment that the self created avatar is necessary for players to commit and take on a new identity they value and in which they become heavily invested in deep learning.
- **Co-presence**: Co-presence is the sense of being there in other places and being together with other people (Biocca et al. 2001). It is the fundamental prerequisite of collaborative learning and it is claimed to be one of the crucial social components of computer-mediated communication (Spears and Lea 1992).
• **Collective Identity:** Players can join *guilds* (named teams) to socialize and play together. Collaborative-competition is introduced in MMOG and encourages collaboration among in-team players to compete with out-team players. According to the social identity theory (Tajfel and Turner 1979), when individual is faced with such situation, collective identity would be introduced and individuals tend to act in the benefit of the team and discriminate the out-teams.

• **Transparency:** Transparency in MMOG is to allow players to retrieve values of attributes of avatars in the environment. Such characteristic allows informational influence to happen between players’ actions in the virtual environment, which in return generate an internalization process when a player perceives information as a means to enhance his/her knowledge above that of reference groups (Kelman 1961).

2.2 Technology Acceptance Model (TAM): Motivation to Use/Play and Intention to Use/Learn

A game is not necessary to be designed for educational purpose in order to make learning happen from the player’s point of view. Squire (2005) argues that commercial computer and video games have mostly been overlooked by educational technologists. Such argument aligns with the constructivist belief of learning, which views knowledge not as information transmitted from a teacher to a learner, but rather as the result of the active engagement of the learner in the comprehension and conception of the information (Bradsford 1979). Looking at the constructivist belief of learning from the Information Systems (IS) perspective, computer games and their subset MMOG fall into the category of hedonic information system (Van der Heijden 2004), in which learning in a MMOG can be a process through motivating a player with an intention to learn and play the game individually and collaboratively.

According to motivational theories (Deci and Ryan 1985), motivation to play can be determined by two fundamental types of motivation, namely intrinsic motivation and extrinsic motivation. From the IS perspective, the area of research on users’ acceptance of information systems is mainly rooted in the Technology Acceptance Model (TAM) (Davis 1989), which explains user’s intention to use an information systems is affected by perceived usefulness, perceived ease of use, and perceived enjoyment. Based on motivational theories, Davis et al. (1992) further clarifies perceived usefulness as an example of extrinsic motivation, whereas perceived enjoyment is an example of intrinsic motivation.

Moreover, MMOG is more than single player participation due to its “massively multiplayer” nature, in which players’ participation can be motivated by his/her peers in the same team. According to cognitive learning theory, learning consists of two components — personal and social (Entwistle and Entwistle 1992). Player-to-player interactions in MMOG allow learning to occur not only individually but also collaboratively. In addition to individual engagement in the learning process, computer-mediated learning can be initiated and motivated at peer level through collaborative learning interaction (Alavi 1994).

With the recognition of the importance of peer motivations, learning intentions, collaborative interaction in game-based learning reported in the TAM and cognitive learning research, this study proposes two new constructs, *peer intrinsic motivation* and *peer extrinsic motivation*, and studies their effects on intention to learn individually and intention to learn collaboratively in the context of MMOG game-based virtual environment.

3 METHODOLOGY

3.1 Research Model

This study proposes two new constructs, *peer intrinsic motivation* and *peer extrinsic motivation*, and studies their effects on intention to learn individually and intention to learn collaboratively (See
Figure 1 for the research model, and Table 1 for the definitions of constructs). This study is set in the context of a MMOG game-based virtual environment, World of Warcraft (WoW). The reason for us to pick WoW because it is currently the best selling MMOG in the world with over 11.5 million subscribers worldwide (on 23rd December 2008, Blizzard Press Release: http://www.blizzard.com/us/press/081121.html).

Figure 1. Research Model

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer Intrinsic Motivation to Play (PIM)</td>
<td>… is the perception that an individual will want to perform an activity (i.e. playing a particular game), driven by the desire to have his/her peers engaging in enjoyable, self-determined, and competence-enhancing activity</td>
</tr>
<tr>
<td>Peer Extrinsic Motivation to Play (PEM)</td>
<td>… is the perception that an individual will want to perform an activity (i.e. playing a particular game), driven by the desire to have his/her peers obtaining tangible or intangible external rewards</td>
</tr>
<tr>
<td>Intention to Learn Individually (ILI)</td>
<td>… is the intention for an individual to learn to perform an activity (i.e. playing a particular game) by him/her own efforts continuously for a period of time</td>
</tr>
<tr>
<td>Intention to Learn Collaboratively (ILC)</td>
<td>… is the intention for an individual to learn to perform an activity (i.e. playing a particular game) together with his/her peers (i.e. helping each other to learn) continuously for a period of time</td>
</tr>
</tbody>
</table>

Table 1. Definitions of Constructs

The concepts of peer intrinsic and peer extrinsic motivations for a team are new and form the core of this study. A thorough Google Scholar and online article database search did not show any prior work on these constructs.

In distinguishing intrinsic and extrinsic motivations for an individual, it is very important to note that motivation that comes from peers is counted as external factors, not on the game itself (that is being mainly studied in the current game research), but on knowing about team members’ enjoyment of the game in general (peer intrinsic motivation), and on knowing that team members achieving certain game rewards (peer extrinsic motivation).

3.1.1 Social Comparison and Intention to Learn Individually

This hypothesis is formulated based on Tesser and Campbell’s (1990) argument that the closer the relationship an individual is to someone, the more likely that envious comparison processes will occur. In the context of MMOG, players can be from the same team or belong to the same guild, fighting for the same goal in the same quest (a collective task). Transparency characteristic of MMOG allows players to read his/her peers’ values of attributes and observe their performance. Applying the concept of envy into the virtual environment of MMOG, feeling of envy will arise when one player perceives another player possesses more superior ability than his/her own.
Envy was seen to be helpful in motivating people to increase their performance in battle or civil competitions such as athletic contents. The effect of upward social comparisons on negative psychological reactions could be explained by the relevance of the superior target person as a role model (Lockwood and Kunda 1997). Consistent with this reason, the social learning literature has found that individuals tend to imitate others whom they perceived as being more successful and/or more highly rewarded (Bandura 1986).

In the MMOG environment, the feeling like “I wish I had what you have” is a result of social comparison between the player and his/her peers. Besides the negative feeling of envy, players may also appreciate the successful players of their team because their external rewards will be shared by the peers including the ones who envy others. In other words, in addition to envy that players focus on competition, players may also focus on helping behaviour and target for sharing of collaborative victory. To our knowledge, the amount of research is limited for this kind of peer extrinsic motivation, generated with the feeling of envy from upward social comparison process and collaborative victory, will motivate a player’s intention to learn on his/her own in order to differentiate him/her from his/her peers in a MMOG context. Therefore in this study, we stipulate that a MMOG player, who focuses on peer extrinsic motivation in the game-play process, will have a high tendency to learn on his/her own in the MMOG environment. Therefore, we have the following hypothesis:

Hypothesis 1: Peer extrinsic motivation to play will have a significant positive influence on intention to learn individually in a MMOG context

3.1.2 Knowledge Sharing and Intention to Learn Collaboratively

Kalling and Styhre (2003) comment on the relative lack of attention paid to the role of motivational factors that influence knowledge sharing behaviours. As reviewed by Bock et al. (2005) on the benefit of knowledge sharing, knowledge sharing can benefit individual, group and organization. As MMOG is a kind of hedonic information systems, some MMOG players may focus on the fun-aspect rather than productive-aspect of the game. In general, we may consider the benefits of knowledge sharing at three levels: individual fun, group fun and organizational fun, where group and organizational fun may belongs to our definition of peer intrinsic motivation. Individual fun is possibly an interesting area for investigation, however, this study focuses mainly on peer motivations and suggests further research on this construct.

According to studies of social capital and knowledge sharing, individuals will be motivated to contribute their knowledge in the community. Wasko and Faraj (2005) proposed two constructs that would affect individual motivation for knowledge contribution: (1) reputation and (2) enjoy helping. In order to share knowledge, individuals must think that their contribution to others will be worth the effort and that some new value will be created, with expectations of receiving some of the value for themselves (Nahapiet and Ghoshal 1998). These personal benefits or “private rewards” are more likely to accrue to individuals who actively participate and help others (Von Hippel and Von Krogh 2003).

In addition to enhancing their reputations, individuals may also receive intrinsic benefits from contributing knowledge. Self-evaluation based on competence and social acceptance is an important source of intrinsic motivation that drives engagement in activities for the sake of the activity itself, rather than for external rewards (Bandura 1986). Players may perceive that helping others with challenging problems is interesting, and because it feels good to help other people by solving their problems (Kollock 1999). Prior research in electronic networks suggest that individuals are motivated intrinsically to contribute knowledge to others because engaging in intellectual pursuits and solving problems is challenging or fun, and because they enjoy helping others (Wasko and Faraj 2000). Therefore, instead of focusing on comparing the extrinsic rewards of peers, a MMOG player who focuses on peer intrinsic motivation is likely to share knowledge and learn collaboratively through helping each other to learn in the MMOG environment, and therefore we have the following hypothesis 2:
Hypothesis 2: Peer intrinsic motivation to play will have a significant positive influence on intention to learn collaboratively in a MMOG context

3.2 Development of the Questionnaire

A survey design was used to investigate the aforementioned hypotheses. The literatures on behavioural intention and motivation are reviewed. Survey questions were constructed with reference to prior empirical work and modified to fit the context of this research. All questions were phrased from the perspective of current WoW players, and anchored on a seven-point scale from Strongly Disagree (1) to Strongly Agree (7).

Questions on peer intrinsic motivation and peer extrinsic motivation were drafted based on Work Preference Inventory (Amabile et al. 1994). These questions assessed intrinsic and extrinsic motivation in terms of work preference. Eight questions (four for peer intrinsic motivation; four for peer extrinsic motivation) were selected from the inventory and modified according to the definition of peer intrinsic motivation and peer extrinsic motivation from the perspective of current WoW players.

Questions on intention to learn individually and intention to learn collaboratively were created with reference to construct “behavioural intention” from Venkatesh et al. (2003). With reference to their original source, these questions are modified to assess one’s behavioural intention to learn within the next month. Eight questions (four for intention to learn individually; four for intention to collaborate) were created based on the definition of intention to learn individually and intention to learn collaboratively.

A total number of 16 items were generated for the validation process of the constructs. Conceptual construct validation was carried out following Moore and Benbasat’s (1991) card sorting procedure. Table 2 presents the measurement of items for the questionnaire.

<table>
<thead>
<tr>
<th>Peer Intrinsic Motivation (PIM)</th>
<th>1 (Strongly Disagree) to 7 (Strongly Agree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIM1</td>
<td>I am strongly motivated by knowing my teamates find the game enjoyable.</td>
</tr>
<tr>
<td>PIM2</td>
<td>I am strongly motivated by knowing my teamates find the game exciting.</td>
</tr>
<tr>
<td>PIM3</td>
<td>I am strongly motivated by knowing my teamates find the game fun.</td>
</tr>
<tr>
<td>PIM4</td>
<td>I am strongly motivated by knowing my teamates find the game interesting.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Peer Extrinsic Motivation (PEM)</th>
<th>1 (Strongly Disagree) to 7 (Strongly Agree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEM1</td>
<td>I am strongly motivated by knowing my teammates get experience points.</td>
</tr>
<tr>
<td>PEM2</td>
<td>I am strongly motivated by knowing my teamates earn reputation points.</td>
</tr>
<tr>
<td>PEM3</td>
<td>I am strongly motivated by knowing my teammates get in-game rare weapons.</td>
</tr>
<tr>
<td>PEM4</td>
<td>I am strongly motivated by knowing my teamates earn gold coins.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intention to Learn Individually (ILI)</th>
<th>1 (Least Likely) to 7 (Most Likely)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILI1</td>
<td>I will continue to learn the skills on my own to perform better in the game World of Warcraft.</td>
</tr>
<tr>
<td>ILI2</td>
<td>I intend to learn on my own continuously how to defeat different kinds of creatures in the game World of Warcraft.</td>
</tr>
<tr>
<td>ILI3</td>
<td>I believe I will continuously practice my skills on my own to better control my avatar in the game World of Warcraft.</td>
</tr>
<tr>
<td>ILI4</td>
<td>I plan to continuously acquire more knowledge related to the game World of Warcraft on my own.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intention to Learn Collaboratively (ILC)</th>
<th>1 (Least Likely) to 7 (Most Likely)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILC1</td>
<td>my teammates will continue to help me learn how to play better in the game World of Warcraft.</td>
</tr>
<tr>
<td>ILC2</td>
<td>I intend to continuously help my teammates learn how to defeat different kinds of creatures in the game World of Warcraft.</td>
</tr>
<tr>
<td>ILC3</td>
<td>I believe I will continuously help my teammates practice their skills to better control their avatars in the game World of Warcraft.</td>
</tr>
<tr>
<td>ILC4</td>
<td>my teammates plan to continuously help me acquire more knowledge which is related to the game World of Warcraft.</td>
</tr>
</tbody>
</table>

Table 2. Measurement of Items in the Self-Reported Survey
3.3 Data Collection

This study utilizes an online survey website to collect data from current players of WoW. Players of WoW are invited to fill in the survey through online WoW communities, like game forums and corresponding network on Facebook. To guarantee the quality of respondents, raw data are filtered based on the following criterions:

3.3.1 World of Warcraft Experience

One “yes-or-no” question at the beginning of the questionnaire required the respondent to answer “Have you ever played World of Warcraft?” All respondents who indicated without experience in WoW will be filtered.

According to the design of WoW, players create their avatar by selecting race and class. Limited by the design, several classes are exclusive to certain races. Based on these rules of combinations, two questions are included in the questionnaire, which ask the respondent to select the race and the class of their strongest avatar. The combination of race and class will then be checked with the above rules. Those who fail this check, or indicate “others” in the selection, will be dropped for analysis.

3.3.2 Definition of Current Participants

Dependent variables of this research are intention to learn individually and intention to learn collaboratively (see Table 1 for their definitions). One question is added in the questionnaire and the subject is required to disclose “When was the last time you played the game World of Warcraft?”. Only those responses which indicate “within one day” and “longer than one day and within one week” will be used for analysis.

The survey received a total number of 94 valid responses from current active players of the WoW after corresponding data filtering procedure. Table 3 summarizes the demographic information of respondents:

<table>
<thead>
<tr>
<th>Gender</th>
<th>Male</th>
<th>71</th>
<th>75.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>23</td>
<td>24.5%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WoW Experience</th>
<th>Less than 1 year</th>
<th>19 20.2%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>More than 1 year but less than 2 years</td>
<td>25 26.6%</td>
</tr>
<tr>
<td></td>
<td>More than 2 years but less than 3 years</td>
<td>30 31.9%</td>
</tr>
<tr>
<td></td>
<td>More than 3 years</td>
<td>20 21.3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Last Play *</th>
<th>Within one day</th>
<th>81 86.2%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Longer than one day and within one week</td>
<td>13 13.8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall Weekly Participation</th>
<th>Mean 23.3 hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>20.0 hours</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other MMOG Experience</th>
<th>Yes 57 60.6%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No 37 39.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age Group</th>
<th>15 or below</th>
<th>1 1.1%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>16 - 20</td>
<td>35 37.2%</td>
</tr>
<tr>
<td></td>
<td>21 - 25</td>
<td>33 35.1%</td>
</tr>
<tr>
<td></td>
<td>26 - 30</td>
<td>16 17%</td>
</tr>
<tr>
<td></td>
<td>31 - 35</td>
<td>8 8.5%</td>
</tr>
<tr>
<td></td>
<td>36 or above</td>
<td>1 1.1%</td>
</tr>
</tbody>
</table>

* This study only consider players who participated in WoW within one week as “current participants”

Table 3. Demographic Information of Respondents
3.4 Data Analysis

The research model and hypotheses were estimated using 200 iterations of the bootstrapping technique in PLS Graph 3.00.

3.4.1 Measurement Model

The Average Variance Extracted (AVE) of PLS analysis attempts to measure the amount of variance that a latent variable component captures from its indicators relative to the amount due to measurement error, and so as to assess the convergent validity of the constructs. AVE values should be greater than the generally recognized .50 cut-off. The Internal Composite Reliability (ICR) of PLS analysis attempts to assess inter-item reliability, and so as to ensure internal consistency of indicators. Acceptable values of an ICR for perceptual measures should exceed .70 and should be interpreted like a Cronbach’s Coefficient Alpha. Table 4 summarizes the measurement model results with all ICR and AVE values satisfied the recommended values:

<table>
<thead>
<tr>
<th>Measures</th>
<th>Items</th>
<th>Composite Reliability</th>
<th>Average Variance Extracted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention to Learn Individually (ILI)</td>
<td>4</td>
<td>0.964</td>
<td>0.871</td>
</tr>
<tr>
<td>Intention to Learn Collaboratively (ILC)</td>
<td>4</td>
<td>0.968</td>
<td>0.884</td>
</tr>
<tr>
<td>Peer Intrinsic Motivation (PIM)</td>
<td>4</td>
<td>0.984</td>
<td>0.939</td>
</tr>
<tr>
<td>Peer Extrinsic Motivation (PEM)</td>
<td>4</td>
<td>0.950</td>
<td>0.827</td>
</tr>
</tbody>
</table>

Table 4 – Result of Confirmatory Factor Analysis

Discriminant validity indicates the extent to which a given construct is different from other constructs. To evaluate discriminant validity in PLS analysis, the square root of AVE for each construct should be higher than the levels of correlations involving the construct (Chin 1998). One criterion for adequate discriminant validity is that the construct should share more variance with its measures than with other constructs in the model (Barclay et al. 1995). Table 5 demonstrates discriminant validity with the square root of AVE for each construct is greater than the levels of correlations involving the construct.

<table>
<thead>
<tr>
<th></th>
<th>ILI</th>
<th>ILC</th>
<th>PIM</th>
<th>PEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILI</td>
<td>0.933</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ILC</td>
<td>0.621</td>
<td>0.940</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIM</td>
<td>0.308</td>
<td>0.467</td>
<td>0.969</td>
<td></td>
</tr>
<tr>
<td>PEM</td>
<td>0.339</td>
<td>0.376</td>
<td>0.589</td>
<td>0.909</td>
</tr>
</tbody>
</table>

Note: ILI: Intention to Learn Individually  ILC: Intention to Learn Collaboratively  PIM: Peer Intrinsic Motivation  PEM: Peer Extrinsic Motivation

* The shaded numbers in the diagonal row are square roots of the average variance extracted.

Table 5 – Correlation between Constructs

The explanatory power of the structural model is evaluated by looking at the $R^2$ value of the two dependent constructs. Model testing is divided into two separated PLS analysis: (1) test of hypotheses, and (2) test of the full model. The first analysis tests the proposed hypotheses of this study, while the second analysis tries to validate the proposed research model by considering the other possible paths as full model.
3.4.2 Test of Hypotheses

**Hypothesis 1:** Peer Extrinsic Motivation to Play will have a significant positive influence on Intention to Learn Individually in a MMOG context

As shown in Figure 2, the $R^2$ value for Intention to Learn Individually is 12%. The proposed path between Peer Extrinsic Motivation to Play and Intention to Learn Individually is significant ($\beta = 0.34$, $p < 0.01$). Statistically, this result showed that Peer Extrinsic Motivation to Play had a significant positive influence on Intention to Learn Individually in a MMOG context.

**Hypothesis 2:** Peer Intrinsic Motivation to Play will have a significant positive influence on Intention to Learn Collaboratively in a MMOG context

As shown in Figure 2, the $R^2$ value for Intention to Learn Collaboratively is 22%. The proposed path between Peer Intrinsic Motivation to Play and Intention to Learn Collaboratively is significant ($\beta = 0.47$, $p < 0.01$). Statistically, this result showed that Peer Intrinsic Motivation to Play had a significant positive influence on Intention to Learn Collaboratively in a MMOG context.

3.4.3 Test of the Full Model

**Paths to Intention to Learn Individually**

The $R^2$ value for Intention to Learn Individually is 13%. As shown in Figure 3, for the links to Intention to Learn Individually, only the path between Peer Extrinsic Motivation to Play and Intention to Learn Individually is significant ($\beta = 0.24$, $p < 0.1$). The path between Peer Intrinsic Motivation to Play and Intention to Learn Individually is not significant. With the consideration of the full model, this analysis result shows that by including the path between Peer Intrinsic Motivation to Play and Intention to Learn Individually into the PLS analysis, it will not affect the existing significant path discovered in the first analysis.

**Paths to Intention to Learn Collaboratively**

The $R^2$ value for Intention to Learn Collaboratively is 23%. As shown in Figure 3, for the links to Intention to Learn Collaboratively, only the path between Peer Intrinsic Motivation to Play and
Intention to Learn Collaboratively is significant ($\beta = 0.38$, $p < 0.01$). The path between Peer Extrinsic Motivation to Play and Intention to Learn Collaboratively is not significant. With the consideration of the full model, this analysis result shows by including the path between Peer Extrinsic Motivation to Play and Intention to Learn Collaboratively into the PLS analysis, it will not affect the existing significant path as discovered in the first analysis.

Table 6 summarizes the comparison of analysis results:

<table>
<thead>
<tr>
<th></th>
<th>$R^2$</th>
<th>Test of Hypotheses</th>
<th>Test of the Full Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILI</td>
<td>0.116</td>
<td></td>
<td>0.133</td>
</tr>
<tr>
<td>ILC</td>
<td>0.218</td>
<td></td>
<td>0.233</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\beta$ (Sig. Level)</th>
<th>Test of Hypotheses</th>
<th>Test of the Full Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIM $\to$ ILI</td>
<td>n/a</td>
<td></td>
<td>0.167 (not sig.)</td>
</tr>
<tr>
<td>PIM $\to$ ILC</td>
<td>0.467 ($p &lt; 0.01$)</td>
<td></td>
<td>0.376 ($p &lt; 0.01$)</td>
</tr>
<tr>
<td>PEM $\to$ ILI</td>
<td>0.340 ($p &lt; 0.01$)</td>
<td></td>
<td>0.240 ($p &lt; 0.1$)</td>
</tr>
<tr>
<td>PEM $\to$ ILC</td>
<td>n/a</td>
<td></td>
<td>0.154 (not sig.)</td>
</tr>
</tbody>
</table>

Note: ILI: Intention to Learn Individually  ILC: Intention to Learn Collaboratively  PIM: Peer Intrinsic Motivation to Play  PEM: Peer Extrinsic Motivation to Play

Table 6. Comparison of Analysis Results

4 DISCUSSION AND IMPLICATIONS

4.1.1 Academic

For academic researchers, the study contributes to a better understanding of the motivation to play in a MMOG context. Hedonic information systems (i.e. MMOG) differ from utilitarian information systems in terms of reasons for use, and this study is one of the very few that has attempted to investigate the motivational drivers of user acceptance in a hedonic information system context. More specifically, this study introduces two new constructs, namely peer intrinsic motivation and peer extrinsic motivation to explain players’ behavioural intention to learn individually and collaboratively in the MMOG environment. Using the card sorting method and various validations in the data analysis, this study introduces items which are possible to assess the degree of peer motivations and intention to learn under the MMOG context.

Furthermore, unique characteristics of MMOG were put into the explanations of hypotheses in this study, and the results showed that a player’s perception on peer intrinsic motivations is associated with his/her intention to learn collaboratively, while the perception on peer extrinsic motivations is associated with his/her intention to learn individually. This study is one of the early attempts to associate characteristics of MMOG with peer motivational factors and learning intentions in a MMOG context, which may help to act as a theoretical foundation for further studies.

4.1.2 Educators

For education practitioners, the results from this study suggest that in a MMOG environment, players’ peer intrinsic motivation is associated with their intention to learn collaboratively while players’ peer extrinsic motivation is associated with their intention to learn individually. Although the scope of this piece of research does not take into consideration of the context of learning material that MMOG will deliver to the player, the result from this paper is suggesting that educators who want to adopt MMOG
as a platform will need to pay special attention on how they want to manipulate the peer motivational factors and learning intentions of the players from a psychological perspective.

4.1.3  Game Developers

Finally, according to cognitive learning theory, knowledge acquisition consists of two components — personal and social (Entwistle and Entwistle 1992). In the MMOG environment, a player’s intention to learn can also be categorized into intention to learn individually and intention to learn collaboratively. From the player’s point of view, the design of MMOG forces them to collaborate, and the struggle between the feelings of “envy” and “collaborative victory” may apply to the player from the psychological point of view. The design of the game interfaces should give reasonable expectation to each player on the number of experience points needed for promoting towards the next level, and what kind of spell that his/her avatar can cast (with preview demonstration on the effect).

4.2  Limitation of Study and Future Research Directions

The data of the presented study was collected using a survey method from MMOG players of WoW only. We suggest future research may conduct a broader survey and compare players of different MMOG. Also, in addition to a snap-shot of self-reflected data collection, a longitudinal study of individual online gaming behaviours is required for in-depth understanding of team dynamics in MMOG environment.

5  CONCLUSION

This study examines players’ intention to learn in a hedonic information system - MMOG. More specifically, this study investigates the roles of players’ Peer Intrinsic Motivation and Peer Extrinsic Motivation on both Intention to Learn Individually and Intention to Learn Collaboratively in MMOG environment. The results indicate that Peer Intrinsic Motivation to Play has a significant positive influence on the Intention to Learn Collaboratively, while Peer Extrinsic Motivation to Play has a significant positive influence on the Intention to Learn Individually. The results from this exploratory study also expand our understanding of the use of MMOG for collaborative learning purposes as well as its potential causes of threats like addictiveness rooted by players’ continuous “play-to-learn” and “learn-to-play” modes of engagement in MMOG.

References


CONSULTANTS EXPERIENCE OF REQUIREMENTS ELICITATION CONVERSATIONS – AN EMPIRICAL MODEL

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0345.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Information systems analysis and design, IS modeling and design, IS professionals, Research methodology</td>
</tr>
</tbody>
</table>
CONSULTANTS’ EXPERIENCE OF REQUIREMENTS ELICITATION CONVERSATIONS – AN EMPIRICAL MODEL

Davey, Bill, RMIT University, 239 Bourke St Melbourne, 3001 Australia, Bill.Davey@rmit.edu.au

Cope, Chris, Latrobe University PO Box 199 Bendigo 3552 Australia, c.cope@latrobe.edu.au

Abstract

Research continues to show that poor requirements elicitation leads to costly information systems failure. We also know that the most information can be obtained by a consultant conducting conversations with clients. Research was conducted into consultants’ experience of these conversations. Twenty professionals involved in these conversations were interviewed and Phenomenographic analysis of all of the quotes within transcripts from these interviews was undertaken. The analysis resulted in a list of five distinctly different ways in which these conversations can be experienced by consultants. These are given the labels: domination, manipulation, problem resolution, negotiation and partnership. The labels are attached to a range of attributes where differences were found in the data. These well defined categories of description allow us to structure professional development and to clarify future research efforts.
1 INTRODUCTION

In this paper we report research into the conversations between information systems (IS) consultants and clients that are part of requirements elicitation (RE). The research investigated the consultants’ experiences of the conversations. We intend the term “consultant” to be synonymous with “business analyst”, “analyst programmer” or “senior analyst”, among others. We use the term “conversation” where others might use “interview” to avoid confusion between the discussions between consultants and clients as part of RE, and the research interviews.

An early stage of developing an information system is requirements elicitation (RE). There is evidence to suggest that accurate requirements elicitation is difficult and that failure to elicit accurate requirements can cause considerable problems. Davis, Fuller, Tremblay, & Berndt (2006) found “accurately capturing system requirements is the major factor in the failure of 90% of large software projects” echoing earlier work by Lindquist (2005) who concluded ”poor requirements management can be attributed to 71 percent of software projects that fail; greater than bad technology, missed deadlines, and change management issues.” The cost of this failure is enormous. One study found that failed or abandoned systems cost $100 Billion in the USA alone in 2000. (Browne & Rogich, 2001) Not only does a failed system cost money, but fixing mistakes made at the requirements elicitation stage accounts for 75% of all error removal costs (Urquhart, 1999).

What is the nature of the communication between humans that occurs during requirements elicitation? We are indebted to Davis, Dieste, Hickey, Juristo, & Moreno (2006) for a fairly comprehensive review of the research into requirements elicitation techniques. Some findings were:

- Interviewing is cited as the most popular requirements elicitation method
- Structured interviews gather more information than unstructured interviews
- Unstructured interviews gather more information than sorting and ranking techniques

As structured interviewing appears to be the most common RE technique, an important step in avoiding IS development failure would seem to be conducting effective, structured conversations between consultants and clients.

Significant research has been put into improving RE conversation performance. One area of research starts with theories of cognition. Cognitive science tells us that communication between people is hampered by the limitations of human cognition and by problems that arise when communication needs to be conducted by language (Bolloju, 2004, Kim & Peterson, 2001, Siau & Tan, 2005, Gaines, 2003, Hudlicka, 1996)

Several researchers have taken these theories and tried to find techniques for overcoming the cognitive difficulties. The most recent of these was research (Pitts & Browne 2007, Browne & Rogich 2001) which found that using procedural prompting strategies designed to overcome cognitive difficulties produced more meaningful requirements than other prompting techniques. Cognitive science has been used to investigate stopping behaviour. That is, the elicitation of requirements must eventually be called to a halt. Some work has been done by Pitts & Browne (2004) in determining what factors lead to a best choice of when to stop.

The process of requirements elicitation is often seen as a process of mutual education of consultant and client. Some researchers have taken current educational theory and applied it to improving elicitation. Efforts include applying collaborative elaboration – from education theory (Majchrzak, Beath, Lim, & Chin, 2005), mental imagery (Zmud, Anthony, & Stair, 1993), collaborative requirements negotiation (EasyWinWin) and group communication theory (Grünbacher & Briggs, 2001).

Despite this knowledge of possible improvements in RE conversations, they continue to fail. If we accept that the consultants are those who must perform better then we must conduct studies of the way consultants relate to RE conversations. An understanding of this relationship will enable us to
influence these conversations. One powerful way of approaching variation in experience is the method of phenomenography. This method has been used to examine the way in which students experience computing concepts (Bruce et al., 2004; C. Cope, 2002; C. Cope, 2006; Stoodley, Christie, & Bruce, 2004), the way in which teachers experience teaching computing, (C. Cope & Ward, 2001; Kinnunen, McCartney, Murphy, & Thomas, 2007; Levy & Ben-Ari, 2007; Lund & Baker, 1999) in experience of professional management skills (Chen & Partington, 2006) in experience of nursing (Falk, Wahn, & Lidell, 2007; Frank, Asp, & Dahlberg, 2009; Johansson, Swahn, & Strömberg, 2007; Jormfeldt, Svedberg, Fridlund, & Arvidsson, 2007) physiotherapy (Larsson & Gard, 2006) patients families (Schröder, Larsson, & Ahlström, 2007; Widang, Fridlund, & Martensson, 2007) and doctors and sick listing (Swartling, Peterson, & Wahlström, 2007). A consistent finding of phenomenographic research is that a phenomenon can be experienced in a limited number of qualitatively distinct ways. These distinctly different ways of experiencing a phenomenon are described as an outcome space of a number of categories of description. These categories give us a viewpoint of the relationship between a person and a phenomenon that has been shown to be useful in making decisions about affecting that relationship.

2 RESEARCH APPROACH

The aim of this research is to uncover the relationship between consultants and RE conversations. To do this we ask the question “What variation exists in the way that consultants experience RE conversations?” An understanding of this variation will help us understand how to influence conversations.

2.1 Participants

In qualitative research based on interviews sample populations are normally drawn until no more variation exists in data being obtained. In phenomenography this is often taken to be about 20 to 25 interviews. Twenty consultants with professional experience in RE conversations were interviewed. All the initial interviewees were involved in dealing with clients in various stages of implementing systems after the signing of a contract between the consultancy company and a client company. To broaden the scope further targeted people were approached. This included analysts offering information system services within a company, those involved in a range of industries including manufacturing, financial services, government services and retail services. The stages of interaction between consultant and client addressed specifically avoided sales conversations and those situations where a product was being implemented with no concern for the specific requirements of the client. Äkerlind (2005 p103) advises demographic variation so as to maximise the chances of variation of experience of the phenomenon as possible. The selected population showed variation in the attributes: place educated, country of residence, age, years of experience (0 to 30 years), current job, scope of projects (company internal to consultancy), company type, industry (NGO, government, retail, banking, manufacturing, insurance) and gender. It must be emphasized that the method does not cross match outcomes from interviews with these demographic variables.

2.2 Interviews

The phenomenographic approach seeks to control the extent to which the interviewer imposes a view on the interview (Bowden & Green, 2005). Interviews were semi structured and started with an introduction that was consistent and defined RE conversations. A closed question “describe your last RE conversation” was asked and followed by open questions. Analysis of these interviews examines transcripts for meaning rather than textual content, so open questions are designed to explicitly reveal the meaning within initial statements of the interviewees, using only words they had introduced. Interviews ran for between 45 and 90 minutes. Each interview was then transcribed.
2.3 Data Analysis

The data gathering and analysis was conducted using a phenomenographic approach as described by Marton (Marton, F. 1981, Marton, F., & Booth, S. 1997) and detailed in Bowden & Green (2005). The result of this analysis should be a description of the variation of experience of the phenomenon expressed as a set of related categories of description of the phenomenon. Each transcript was entered into a specially written database program. The transcripts were examined for every statement that contained some meaning and these quotes were separated from the transcript with a link maintained to the original text. The links allowed quick checks of each quote in context to ensure that meaning was not mistaken due to fragmenting the transcript. Quotes were organised according to any aspects of the phenomenon they addressed. Variations in meaning and structure were allowed to emerge from the quotes through an iterative process comparing quotes, possible dimensions of the phenomenon showing variation and the original transcripts. Multiple iterations of this process were conducted as recommended by Bowden and Green (2005). In each iteration candidate categories are postulated and then checked against the data to see if each candidate category sensibly reflects meaning found in the data. A sensible set of categories is also one where some structure exists that is supported by quotes. A useful category would describe a range of aspects where the quotes indicate that a view of one aspect is related to a view of another aspect, or where there is a simple logical connection between aspects. It is important when constructing categories to ensure that they come from meaning within the data rather than the researcher finding quotes to support a previously determined structure. Candidate categories were compared with the data as a whole to ensure that significant ideas in the transcripts were represented in the categories. Categories were continually checked against whole transcripts to make sure no important sections were not represented in the outcome and that the categories truly reflect things said by the participants without leading from the researcher. When this check was done it was found that each respondent had several views of the phenomenon. The categories reflect variations in views rather than variation between respondents.

Links between the categories were analysed and presented in the results. An indication that a candidate category be preferred was the existence of several quotes directly supporting the category and quotes explicitly relating dimensions within a category.

3 RESULTS

As a result of analysis of the data five distinct categories of description of experience of RE conversations emerged. These categories are a way in which the researcher seeks to create a logical structure from the variation found in the data. Many concepts were described in the transcripts, each exemplified by multiple statements showing a range of meanings. For convenience of reporting these have been summarised through three dimensions:

- **aspects of the requirements** – what is their nature, how are they seen in terms of the aim of the conversation, where they come from.

- **aspects of the people** - how the consultant sees themselves, how the clients are seen and the relationships between them such as power and personal relationship.

- **aspects of the dynamics** of the conversation are experienced in a range of ways. Concepts such as behaviours that are noticed, strategies that are employed and how progress is experienced. This aspect has been called ‘The act of the conversation’

Each category reported here includes a single transcript quote, chosen from several available, that helps show the distinction between the name of a category and the meaning within the data. The categories are reported in order. The categories range from a collection of many of the simple ideas to a collection of the complex. There is no intention to imply any other order. It should be noted that
categories arise from all the quotes, not from particular people. A phenomenon can be experienced by the same person in different ways at different times. This was clear in the data.

3.1 Category 1: Domination

In this category a requirements elicitation conversation is experienced as a conflict in which the consultant understands the client's requirements, predetermines a set of requirements, and the conversation is needed to have the client sign off on those requirements. Increasing power is a significant aim of the conversation. The consultant has an understanding, possibly from previous projects, of the business and what an envisaged information system could do for the business. The conversation is required since the client is required to sign off the requirements, even though they probably will not understand them or their significance.

- 'Because [person in the consulting company] had come to a certain conclusion about the ‘right’ way forwards with this, which was effectively one based around conflict. It was based around ‘let’s set up an us against them trial, and see who wins’

The requirements are seen as an absolute, determined by the logic and experience of the consultant. Quote 104 “So we’ll tend to chair it, which does give us the power to direct the conversation in ways we want” is typical of data showing that the consultant goes into a requirements conversation with a clearly defined idea of what the outcome should be.

The people are seen as clients, who are a potential nuisance that need to supply a sign off on a document so that the project can proceed, and consultants with a clear idea of the whole project and a need to get started. Examples that illustrate this include the quote “this person [the client] had very, very limited knowledge on information technology, how it works, what sort of benefit it can bring to an organization” This typical quote shows that the consultant sees the client as not being capable of expressing any opinion of the requirements of the system. This is indicative of the external horizon where the consultant sees their company as having exclusive knowledge of what is ‘good for the client.’

The dynamics are seen as presenting to or teaching the client what it is that the client needs. Data often showed that the consultant had a metaphor with other professions where the knowledge of the professional was the source of progress in the conversation. For example Quote 61 “you know, your doctor can prescribe you a medicine, and you’re comfortable to take it without knowing exactly how it works. And so, ah, yeah it’s usually around getting, getting the client comfortable that they’ve, that you’ve got a solution that they’re happy with.”

3.2 Category 2: Manipulation

In this category a requirements elicitation conversation is experienced as a manipulation of the client so that the client agrees to the requirements that the consultant wants to fulfil. A conversation is experienced as convincing the client that our requirements are what they need. The client need not understand the requirements, but must be convinced that they are desirable.

- So that they [the clients] can get a feel of, they [the clients] obviously don’t understand but the fact that they are asking a question like that, they don’t understand what our [processes of the consulting company] processes are, and so we explain a bit about our processes so that hopefully they [the clients] both, get the answer to their questions and feel happy, but also have a better understanding of what we [the consulting company] do

The requirements are seen as an absolute, determined by the consultant, that the client must be taught enough about to make them happy to sign off. Quote 112 “I [the consultant] think another important
way of explaining something like that is trying to distinguish between the important technical bits that someone is trying to understand and the technical bits that if you don’t understand you still get enough out of it. I guess that relates back to breaking things down into pieces and treating the bits that aren’t that important as a black box” shows that the outcome of a conversation can be experienced as a client that understands the requirements to some extent. The object is to make the client feel happy and considered, rather than bullied into signing off.

**The people** are clients, seen as having power but no worth and consultants with technological expertise who are responsible for getting client “buy in” on the requirements. Quote 14 “what I’ll [the consultant] normally do is I’ll prepare presentations, and what I’ll generally do is I’ll target, um, I’ll try and target it at their [the clients] level.” This type of data shows that consultants can see the elicitation conversation as one way. It also demonstrates a view of the conversation as a “presentation” of the predefined requirements. Clients are seen as students of the requirements, needing to be talked to at a level that enables them to understand.

**The dynamics** are seen as using feedback from the questions asked by the client to both create an understanding of the requirements and convince them that they are valuable. Quote 6 “The problem that we had was that he kept trying to sort of ah, cut us off” indicates that the consultant sees it frustrating if the client attempts to make input into the conversation. The conversation actions are seen as directed by the consultant driving towards an acceptance of the requirements.

**How does this differ from domination**

This category sees the conversations as two way exchange. The client expresses lack of knowledge and the consultant gives that knowledge of the requirements in a teaching like situation. Much of the data talks about how the consultant learned from their own teachers as how to explain and react when people demonstrate lack of understanding. Although, as in Domination, the requirements are brought to the conversation, in this category it is seen as important that the client understand them rather than take on faith that they are good for the business.

### 3.3 Category 3: Problem Resolution

In this category a requirements elicitation conversation is experienced as the exchange of information. The conversation is aimed at both finding the things the client needs to implement their part of requirements and the things the consultant needs to know about the business that make this implementation different from previous examples. Conversations are needed to provide each party with the information needed to put known solutions into a new situation

- “since everybody in the meeting have, know each other really well, so we shoot straight into the problems that we were facing last couple of weeks. And pretty we are pretty much just brainstorming ways of solving those problems”

**The requirements** are seen as a problem that must be solved. Requirements are mutable but easily satisfied by finding an answer. The main problem is understanding what the client really means. This is often done by starting with a solution (consultant provides a set of requirements) and finding ways in which the client is confused by or is happy with that solution. Quote 72 “There are some audiences that are, don’t like taking your suggestion or solution. Ah, and so for those audiences you might just state your understanding of the problem and not state the suggested solution. You still should think about what that suggested solution is so that it still makes sense as you discuss the problem, and perhaps where you lead the discussion.” Much of the data actually uses the word “problem” when describing why a RE conversation was initiated. In early interviews the subject was asked ‘are you sure this was a requirements interview?’ which they affirmed. It seems common, especially where agile methods are being used, that requirements are seen to arise from inspecting initial models. While agile methods do not presume this approach to requirements, some views of the process see no distinction between a problem resolution session and requirements gathering.
The people are seen as having power through their understanding of how the business operates but an ignorance of technology. The client often has a confused idea of what they want. This is often expressed as a problem with a prototype, or a problem with the expression of the requirements.

The dynamics are seen as determining the problem and ensuring that the solution presented solves the problem. Quote 305 “What the purpose of it is, how they imagine using it and I don’t know, what data or whatever they imagine supplying to it and sometimes they don’t understand all the pieces so we need to talk about what those pieces actually are and trying to join that up to something we can actually develop on the software side if it’s a software issue” Data of this type indicates that requirements arise when the client starts to try to answer questions such as ‘can you give me the data that the system needs for input?’ or ‘what should the report look like?’ Again a clear picture of what the system should be has been brought by the consultant, but the RE interviews are seen as allowing the client to make changes based on things they find wrong with the initially proposed solution.

How does this differ from Manipulation: If Manipulation involves a set of requirements that must be communicated then Problem Resolution sees RE as a process of refining requirements as clients encounter problems.

3.4 Category 4: Negotiation

In this category a requirements elicitation conversation is experienced as a type of negotiation. Outcomes are seen as dynamic and power is used but not created. The data shows sets that can be described as “Negotiation.” This category is defined as “a negotiation where the client attempts to specify as much as possible for the needs of the company and the consultant tries to narrow this to things that can be done with the skill set available and within the price for the project.” The consultant is focussing on finding a way of dealing with each request of the client in a way that fits within the needs of the consultant. The conversation consists of finding requirements in the context of “some things we can effectively charge the client for, and some things we can’t effectively charge the client for.”

- “they know that most of the things that I [the consultant] suggest to them [the client] are to the best of their interest, as long as they are not conflicting with [the consulting companies’] interest.”

The requirements are seen as an outcome to be negotiated within the general conditions of a contract or memorandum of understanding. This shows in data that clearly indicates a view of the conversation as a bargaining based on expressed wants against cost. For example Quote 101 “trying to detail where we’ve put on some limitations or where we think something is not realistically possible or and in some cases and this is something we have been dealing with recently actually um giving some different options and different if its something we are going to charge for separately then you know different options and different levels of cost for each option”

The people are seen as clients with business needs and consultants with a set amount of resources allocated to the project. Quote 330 indicates the existence of a distinct category of experience “they know that most of the things that I suggest to them are to the best of their interest, as long as they are not conflicting with [the consultant’s company] interest.” This quote shows that the consultant sees, within a conversation, two ‘sides’ with separate interests. The conversation, then, includes finding things for the project that will have a good business outcome for the client without compromising the interests of the consultant.

The dynamics are seen as conducting a give and take type of negotiation until a “fair” outcome is reached. Quote 269 paints this starkly “just because the client asks for it doesn’t mean a) they should want it and, b) that we want to do it. Um, because the, the starting point of most people in the industry is ‘I should please the client’. And I guess, one level of thought of what that means - if the client asks for something I should say ‘we can do it, and we will do it’. Um, and so that’s the largest most common mistake, that staff make. Um, some of that’s completely selfish, there is nothing about um providing the, what the client actually needs. It’s about some things, some things we can effectively
charge the client for, and some things we can’t effectively charge the client for. Our staff aren’t doing us a good service if they’re doing a lot of things for clients that we can’t charge the clients for.”

This piece of data contains a host of rich detail:

“just because the client asks for it” indicates that the client is seen to have expectations and requirements that the consultant is not aware of in advance. “just because the client asks for it” also indicates that the consultant see the conversation as a two way communication where the client can and will ask for requirements. “doesn’t mean a) they should want it” shows a view of the conversation as a place where the client can ask for things and the consultant use their expertise to comment on the business sense of the proposal. “doesn’t mean a) they should want it and, b) that we want to do it” shows that the consultant is aware of a ‘position’ that he/she is to hold during the conversation that ‘represents’ the interests of the consulting company (or group), as opposed to purely technical considerations.

In explaining this in detail with “some of that’s completely selfish, there is nothing about um providing the, what the client actually needs. It’s about some things, some things we can effectively charge the client for, and some things we can’t effectively charge the client for. Our staff aren’t doing us a good service if they’re doing a lot of things for clients that we can’t charge the clients for.” The consultant is suggesting that the principal focus of the conversation is negotiating a set of requirements within the commercial boundaries of the contract. This is very distinct from a view where the needs of the business system are central. This can be seen as an example of the difference between a logical and physical design problem.

How does this differ from Problem Resolution: In problem Resolution the consultant is still the originator of the requirements. In Negotiation the two parties are committed to a project through a contract, memorandum of understanding or other instrument, but the project is not tightly defined except in business terms. The requirements are specified through the RE process. The consultant knows what they are capable of and the client knows what the business needs. The conversation then becomes a negotiation where the client attempts to specify as much as possible for the needs of the company and the consultant tries to narrow this to things that can be done with the skill set available and within the price for the project.

3.5 Category 5: Partnership

In this category a requirements elicitation conversation is experienced as partnership. Mutually respected parties attempt to determine requirements that will create a greater whole. A conversation is experienced as a step towards finding the real needs of the business and the possible ways in which the consultant can help the business meet those needs. For example the technology and previous implementations are not seen as a constraint. Many of the quotes supporting this category talk about the project as being a small step in an ongoing relationship between the parties.

• “that means they [the client] understand it so deeply, that they can understand where to apply it in a way that you haven’t told them.”
• “what the purpose of meeting them [the client] right now is and um what we are going to be looking to do with them in the future and so what sort of things we should be talking to about when we are meeting them”

The requirements are seen as a short term goal of several discussions. The project is seen as the next step in a developing an ongoing relationship to benefit both parties. Quote 85 “Um try to learn a bit of background about the clients business and um what their specialties are, what their ah focus is and um what their objectives in business”

The people are seen as equals with developing relationship and knowledge of each others needs and abilities. Quote 270 “if something seems a bit crazy, don’t say, don’t just silently say ‘yeah sure, we’ll do that’, say ‘I’m having trouble understanding why you want X, can you give me some more
information about what’s motivating X’. Or if can, if you think you understand what’s motivating it, suggest an alternative solution that achieves what they want and isn’t crazy. Um, and repeat that process until either you understand why they want the thing that didn’t make any sense to you, or, they, together you’re asking for something different that does seem to make sense.”

The dynamics are seen as developing and deepening the relationship between the individuals and their organizations. Quote 167 “Well, if we tell them what they need, um, so in other words, if we thought something was a stupid thing for them to need, we don’t do that, we do the thing we thought they should do. If they’re right that they did need that, they don’t get ‘it’. Um, if we’re right, and they did need it, um, then at the very least someone’s ego has been dented at the client, that we said ‘no, that’s not what you need, this is what you need’. Whereas if we ask them questions until they ask us for something different, they asked us for ‘something’. And so it’s still, you know it, you know that, I mean they’re not dumb. They don’t, they don’t think oh yeah that way was my idea, this way it was still my idea.”

How does this differ from negotiation: Negotiation sees requirements in the context of a single concrete project. Partnership envisages a continuing relationship between the two companies or groups. Conversations seen as partnership do not drive towards an outcome, nor are they seen as failures if a set of requirements does not arise as a clear document. RE is seen as a genuine attempt to find business solutions rather than software specifications. It was found that partnership is an experience of RE across all the types of companies involved in the data collection.

4 DISCUSSION

The aim of this research was to understand variation in the way a consultant can experience an RE conversation. The researcher found that five categories of description of ways of experiencing RE conversations could be constructed. These have been given the labels domination, manipulation, problem resolution, negotiation and partnership. The labels have meaning only to the extent that they reflect a general meaning from the variation in attributes. The categories can be given an ‘order’ based on complexity of meaning in the many quotes that relate to each category. Let us consider each category and look at what correspondence there might be between the category, commercial reality and the literature.

Domination, in terms of the variation in description is the simplest of the categories. That is, the view of the client, the view of the requirements and the impact that the conversation might have on them, is very limited. In a commercial context, however, it could be seen as the quickest way of concluding the RE phase of a project. If clients are seen as having no possibility of contributing to the project, then wasting time on interviewing them is silly. Requirements must be signed off at the start of a project so that project completion can be proven later. Urquhart (2001) found evidence of similar domination in one case where the client saw the consultant as “hiding behind a veil of expertise.”

Manipulation also sees requirements as something simple that are a product of the experienced and professional consultant and hence the process of carrying on a conversation is only complicated by the problem of convincing someone of what seems an obvious given. Again, in a commercial context, it could be motivated by the belief that people not committed to a project can make trouble later if they feel no involvement.

Problem Resolution may seem a different type of approach, but still starts with a consultant supplied project specification. It recognises that new clients can have different ways of doing things that we did not anticipate, that circumstances can change and that the requirements of a perfect solution may need to adapt to circumstances.

Negotiation is a label that should not be confused with the generally accepted meaning. Negotiation has been defined as “an interpersonal decision-making process by which two or more people agree how to allocate scarce resources” (Thompson, 2000). Adair & Brett (2005) found empirical evidence
of a four stage model of negotiation relational positioning, problem identification, solution generation and finding agreement. Our data does not reflect this process oriented view of negotiation, but does reflect the responses one would expect when two people are concentrating on the issue of allocating scarce resources.

Our negotiations are about two parties coming to the table with a set of wants on one side and a set of skills on the other. The conversation tries to find ways of meeting as much of the wants as possible within a set price. It seems to arise from the commercial imperative. The client has signed off to pay for a generally described service, which has been identified in enough detail to form the basis of a contract. The consultant comes to the requirements phase with a general project brief including a contract price and a general description that must be fulfilled. The consultant knows that the system to be specified in the requirements must be built within the price agreed. The consultant also knows that the client must be satisfied with the final system in order for there to be continued business. These constraints on the conversation are consistent with a focus on the conversation as a bargaining or negotiation where these two positions must be traded off in a result5ant requirements document.

Partnership does not recognise the boundaries of a project as being impermeable. RE conversations are seen as being in a much larger picture where new projects arising may be as important as finishing the current project.

5 IMPLICATIONS FOR THE PROFESSIONAL DEVELOPMENT AND FURTHER RESEARCH.

As indicated in the literature section above, much of the research into conducting RE conversations has been directed at the conversations, that is directly focussed on the phenomenon. An examination of the ways that a consultant can experience conversations allows us to gain insight into how consultants can be influenced to improve conversations. This has great implications for the IS development profession as consultants may be experiencing conversations in a way not likely to lead to accurate requirements. Future research will be needed to describe different approaches to conducting interviews. It should now be possible to relate different approaches to different quality of RE outcomes and to relate different approaches to different experiences of RE conversations. If we can analyse the experiences that are likely to lead to higher quality outcomes this would inform our professional development efforts.

6 CONCLUSION

The data showed five distinct ways in which consultants can experience RE conversations: As Domination, Manipulation, Problem Resolution, Negotiation or Partnership. Each of these categories is a label standing for a complex set of variations in the way RE conversations are experienced.

On one level we have a hierarchy of categories of description. Ranging from Domination where the consultant is not even interested in the client understanding the requirements through to Partnership where the requirements are seen as part if an evolving relationship between the consultant and the client. In this case the hierarchy might be useful to a consultant, but different commercial environments make a particular view more relevant than others. Many of the participants in the research mentioned the problem of scope creep. In this phenomenon a price is decided for a project, but in what is supposed to be a process of refining requirements, new requests are introduced that gradually move the project well beyond what was originally quoted. Things a client can see as minor changes and improvements can eventually grow to be far more costly than the whole original project. Hence a Domination view of what requirements elicitation could be the most effective way of avoiding scope creep and could be seen as the most valuable and mature way of looking at requirements elicitation. Each category has its own value. The outcome space describes five different ways of seeing an RE conversation.
References


Lindquist, C. (2005). Required: Fixing the Requirements Mess ; The requirements process, literally, deciding what should be included in software, is destroying projects in ways that arent evident until its too late. Some CIOs are stepping in to rewrite the rules. CIO, 19, 53-60.


Urquhart, C. (1999). Themes in early requirements gathering The case of the analyst, the client and the student assistance scheme. Information Technology & People, 12(1), 44.


Service Scenarios - A Socio-Technical Approach to Business Service Modeling

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0413.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>IS modeling and design, Service innovation, Case Study, IT Services</td>
</tr>
</tbody>
</table>
SERVICE SCENARIOS - A SOCIO-TECHNICAL APPROACH TO BUSINESS SERVICE MODELING

Lars Bækgaard, Aarhus School of Business, Aarhus University, Haslegårdsvej 10, DK-8210 Aarhus V, Denmark, lab@asb.dk

Abstract

The paper presents and discusses the notion of service scenarios. A service is work done by a service executor in interaction with a service consumer. A service scenario is a model of a service system and the roles that are played by the actors that participate and interact during the execution of a service. The model represents the roles and the interactions between the participants. Service scenarios can be used to model specific services and the roles played by human beings and IT systems in the execution of services. The use of service scenarios is demonstrated by means of a case study in a public library. The case study indicates that service systems should be understood as socio-technical systems in which service executors and service consumers co-create value in mutual interaction with each other and with a set of shared resources.


1 INTRODUCTION

The notion of a service is a business concept. Service-orientation represents an organizing principle in which everything that is offered by a business is thought of as a service. The inherent perspective of service-orientation is that a process is viewed as a set of services that is offered to consumers that interact with the services. Each service may itself be the consumer of other services. Thus, a process is viewed as a network of interacting services and consumers. Businesses engage in such networks of services.

Services have been studied in areas like, say, financial services (Dandapani 2004; Homann, Rill et al. 2004; Kumar and van Hilegersberg 2004; Mallat, Rossi et al. 2004; Pan and Vina 2004; Tas and Sunder 2004), health care services (Tan, Wen et al. 2005), and public services (Goldkuhl 2006; Janssen, Gottmaker et al. 2008). Across such diverse areas the service concepts share the common characteristic that a service is work done by one party for another party. Shared services are support processes from which many parties can benefit (Ulbrich 2006).

The purpose of the paper is to present the notion of service scenarios that can be used to model services as socio-technical systems in which service executors and service consumers co-create value. It is necessary to model services in order to design and deliver them in an effective, efficient, and flexible manner. As pointed out by Quartel et al. (2007) the service concept “[…] has so far not been used to its full potential due to the lack of a comprehensive conceptual framework.

Service scenarios support conceptual modeling of existing and envisioned business services and they facilitate analysis of and creative discussions about business services. They can be used to represent the internal structure of a service provider, a service executor, a service consumer, and their shared resources. And they can be used to represent the potential relations (including interactions) between provider, executor, consumer, and shared resources. A service scenario can be viewed as a model that represents a service as a socio-technical work system. The executor of a service can be composed of a combination of human beings and IT systems.

The use of service scenarios is illustrated by means of a case study at a library in which they have been used to model information search services. The case study indicates that service scenarios can be useful tools that can create insights into services and their strengths and weaknesses. Consequently,
service scenarios can support service design. Furthermore, the case study indicates that service scenarios highlight important aspects of services. For example, service scenarios can be used to highlight service systems as socio-technical systems in which service consumers and service executors collaborate to co-create value.

The paper’s contribution is a modeling approach that can be used to model the architecture of services and their execution. Existing service modeling approaches like COSMO (Quartel, Steen et al. 2007) and process modeling approaches like role-activity diagrams (Aburub, Odeh et al. 2007) and BPMN (White 2004) can be used to model service processes. Service scenarios can be used to supplement such approaches in terms of models of structural and architectural aspects of services that highlight the human and material operants that participates in service executions.

Section 2 contains a discussion of business services and their characteristics. Section 3 presents and discusses the notion of service scenarios. Section 4 reports from a case study in which service scenarios are used to model information search services in a public library. Section 5 discusses our findings and some of their implications. Section 6 concludes the paper and suggests directions for future research.

2 BUSINESS SERVICES

The value propositions of businesses and their information systems can be viewed as services that offer support to consumers’ creation of value (Cherbakov, Galambos et al. 2005; Maglio and Spohrer 2008). Such services can be viewed as systems in which actors act upon resources in interaction with service customers (Maglio, Srinivasan et al. 2006). In general, it is too simplistic to claim that a service creates value for its customer. Value cannot be created without active participation of the consumer. Service-orientation is inherently a consumer-oriented approach (Rust and Miu 2006) in which value is co-created (Maglio and Spohrer 2008; Payne, Storbacka et al. 2008; Vargo and Lusch 2008a; Vargo and Lusch 2008b). Services are about doing something together rather than merely producing output (Vargo and Lusch 2008a). Service science is the study of service systems, which are dynamic value co-creation configurations of resources, i.e., people, technology, organizations, and shared information. This implies that the client does much more than receive the result of a service in a passive manner. The client’s actions contribute to the value-creation in a significant and active manner (Maglio and Spohrer 2008).

Service execution implies many types of interaction between service consumers, service executors and shared resources. Service execution is carried out in networks of interacting services in which actors dynamically switch roles and act as executors in some situations and consumers in other situations (Quartel, Steen et al. 2007). Every service system is both a provider and client of service that is connected to other services in value chains, value networks, or value-creating systems (Maglio and Spohrer 2008). A business can be viewed as a network of federated service—a network of interacting components (Cherbakov, Galambos et al. 2005).

Maglio et al. (2006) characterize a service in terms its provider, client, and target. The provider is comprised of an individual or organization and technology operated by or owned thereby. The client is an individual or organization. The target is a part of reality that must be transformed or operated on by the provider for the sake of the client. The target may be comprised of a combination of people, businesses, products, and information. This implies that the effect of a service should be explicitly understood in terms of the changes made to a target that is shared by the provider and the client.

Vargo and Lusch (2008a) focus on the competences of the involved operants, i.e., the actors that participate when a service is executed. This implies that the human and technological actors that act as operants must possess the knowledge and skills that are necessary for the execution of a service.

Mathiassen and Sørensen (2008) characterize services as socio-technical systems in which “… configurations of people and IT artefacts interact to support work, communication, and decision-
making. A service is executed by a combination of human beings and IT systems (Sheth, Verna et al. 2006; Vidgen and Wang 2006).

OASIS’ reference model for service-oriented architecture defines a service as the performance of work by a service provider for a service customer (OASIS 2006). A service provider may be the consumer of other services. A service enables access to capabilities using a prescribed interface that comprises the specifics of how to access the underlying capabilities. There are no constraints on what constitutes the underlying capability or how the service provider implements the access mechanism. Thus, the service could carry out its described functionality through one or more automated and/or manual processes that could invoke other available services.

The implementation of a service is typically hidden from the service consumer except for information and behaviour that are visible through the service interface and information that is required by service consumers to decide whether a given service is relevant. This is known as service information hiding (Barros and ter Hofstede 1998). The effect of executing a service is a realization of one or more effects in terms of delivered information and/or changes to a shared state.

COSMO is a modeling framework that views a service as common interactions, the results of these, and their mutual dependencies (Quartel, Steen et al. 2007). Services are seen as units of composition and decomposition. This implies that services are executed in interaction with other services in networks in which the interactions are constrained by mutual dependencies. A service has a structure, a potential behaviour, and available information that influence its potential interaction.

Kaner and Karni (2007) characterize service systems in terms of the following nine aspects. (1) The customers that benefit from or are affected by the system. (2) The goals and meaning of the system. (3) Inputs to the system; (4) Outputs from and effects of the system. (5) Processes that are performed by the system. (6) Human enablers, i.e., human resources that own and/or operate the system. (7) Enablers, i.e., physical and technological resources which aid in operating the system. (8) Information and knowledge resources that support the system. (9) The system’s environment.

3 SERVICE SCENARIOS

The following service characteristics can be derived from the preceding discussion of service concepts. (1) Service is delivered by means of provider-client interaction. (2) Humans and technology may act as operants. (3) Value is co-created when provider and client interacts. (4) Services are executed in networks of interacting services. (5) The provider and the client interacts with a shared service target. (6) Service execution is constrained and supported by the interaction structures of the involved services. (7) Service execution is constrained and supported by the operants skills, potential behaviour, and available information. (8) Service implementations are hidden for clients by means of service information hiding.

A service scenario is a model of a service system and the roles that are played by the actors that participate when a service is executed. Service scenarios can be used to model specific services and the roles played by human beings and IT systems in the execution of services. Service scenarios are based on characteristics (1)-(6) with the following modifications. The provider-client relation has been extended to a provider-executor-client relation. And the label “target” has been changed to “shared resources”. Service scenarios do not focus on characteristics (7) and (8). Figure 1 illustrates the elements and their potential relations (including interactions) in our conceptual service scenario model.
Figure 1. Conceptual service scenario model

Service scenarios distinguishes between service providers that offers services and service executors that executes services in order to emphasize that a given service can be executed in several different ways by different executors.

A provider possesses the right to offer the service and make it available to potential consumers. A provider and a consumer may negotiate conditions for access to services. This may include service-level agreements about price, payment, availability etc. The provider hires, rents, builds, buys and organizes executors. If the executor is a person the provider may hire him or her. If the executor is an IT system it may be bought, rented or constructed by the provider.

An executor may be a set of persons and/or IT systems that performs the work necessary to execute a service in interaction with a consumer that may be a set of persons and/or IT systems that uses the service.

Providers, executors and consumers may share and interact with a set of resources. The executor and the consumer perform a series of actions when a service is executed. The executor may interact with executors of other services relative to which it acts as a consumer. The effect of executing a service is a realization of a set of effects in terms of delivered resources and/or a change to a set of shared resources.

Example. A bank (provider) may offer a service like “withdraw money from account” to bank customers (consumers) that have accounts in the bank. The execution of the service may include actions like “check pin code” and “dispose money” and it may be performed in different ways with different executors. (1) The service may be executed by software (executor) that interacts with the customer through a web interface. (2) The service may be executed by a bank employee (executor) that interacts directly with customers (consumers) in one of the bank’s departments. (3) The service may be executed by a combination of software (executor) and an ATM machine (executor) that interacts with bank customers through an interface through which the customers can insert bank cards and enter, say, pin codes and amounts. Bank customers share access to accounts (shared resources) with the bank and with the executors of withdrawal services. When a “withdrawal” service is executed the executor changes the balance of the involved account.

4 CASE STUDY

We have performed a case study at a Danish public library in which an information search service offered to library users was changed (Bækgaard, Jørgensen et al. 2007). The case study was carried out as a part of a change project at the library. The purpose of the change project was to identify and implement possible improvements of the library’s mediation of library user’s search for relevant information.

During the project a number of analysis and design activities were carried out. User simulations were used to establish understandings of the current activities. Models of current activities were used to capture aspects of these understandings. Formulation of future stories and brainstorms were used to create visions about changed activities and new ways of using IT systems. Modeling of future situations were used to capture aspects of the visions.

During the modeling process a variety of modeling approaches were used. Service scenarios were found useful because they highlight the socio-technical aspects of services and because they highlight the types of interactions that occur when services are executed. Service scenarios were used in the case
study to model interaction between consumer, executor and shared resources. The provider was not modeled in the case study.

4.1 Current service

The purpose of the service is to satisfy a library user’s information needs. The service is based on communicative interaction between a library user and a librarian that engage in a dialogue about the library user’s information needs, potential search terms, and the relevance of search results. The library user expresses needs for information and the librarian uses his understanding of these needs to search for information resources via library databases and Internet-based search engines.

The diagram in Figure 2 represents the current information search service in terms of a service scenario. The rectangles represent the roles that can be played by human actors, IT systems, and other resources that participate when the service is executed. Elements can be decomposed. For example, the executor is decomposed into a librarian and a set of search systems (database, Internet search engine, etc.). A star means that an element can occur as multiple instances simultaneously. For example, more than one search result can occur at a given point in time. The arrows represent potential interactions between the elements.

![Service Diagram](image)

**Figure 2. Service 1 – Current service**

The current service is executed by a combination of a librarian and a set of search systems. A library user formulates information needs and the librarian asks questions in order to improve and verify his understanding of the information needs. There is an important element of cognitive activities in which the user and the librarian tries to understand the problem at hand and in which they consider possibilities and reflect upon formulations and search results.

The librarian uses his understanding to formulate parameters and queries to one or more search systems. The search results are available for both the librarian and the library user as shared resources. The library user and the librarian analyze and evaluate the search results. The librarian uses “cut & paste” to copy relevant resources from an answer to the resource collection—a text document in which the selected resources from search answers are stored. Near the end of service execution the librarian formats the resource object and enhances it with clarifying comments.

The effects of the service can be characterized as follows. The librarian adds information resources (URLs etc.) that are relevant for the library user to a resource object that contains the selected resources. For example, if an Internet search engine returns one or more relevant URLs these are added to a digital text document using cut-and-paste, they are written on a piece of paper or the screen shots on which they appear are printed. This implies that a resource object is comprised of unrelated digital text documents and pieces of paper.

4.2 Future service

The current execution of the current service has a number of disadvantages. The librarian creates the resource object by means of cut-and-paste, handwriting, and printing operations. Consequently, the librarian creates the semantic integration of the search systems and the resource object. The IT systems
do not support the integration in any way. Also, the resource object itself is heterogeneous and it is not internally integrated because it is composed of hand-written notes, printed screen-shots, and digital text documents. Finally, it is very difficult to reuse past search results and share the information represented by these among librarians and library users.

The diagram in Figure 3 represents the future service in terms of a service scenario. The future service is based on the same interaction between library user and a librarian as the current service (Figure 2). However, the execution of the future service differs from the current service in the following ways.

![Diagram](image.png)

**Figure 3. Service 2 – Future service**

First, a piece of software called resource manager is added to the executor. The librarian uses the resource manager to select and modify resources from search results. This resource object is a structured and integrated digital document that contains structured resource items. When the librarian and the library user identify a relevant information resource (for example a URL) the librarian can use the resource manager to add selected resources to the resource object. Rather than using cut-and-paste or paper-and-pencil to maintain the resource object the librarian marks the relevant part of a search result and tells the resource manager to add the selection to the resource object that is now structured and fully digitalized.

Second, the resource manager stores the final version of the resource object in a resource database that can itself be used as a search system. Consequently, the librarian can use the resource database as a memory of past service executions that may be searched and reused in future executions of the service.

Apart from simplifying the recording of relevant resources the resource manager integrates the otherwise non-integrated search systems. The service has two effects. It produces a resource object with information resources and it records this object in a resource database that can itself be accessed as a search system.

From the librarian’s point of view the service is improved because of the better integration of search systems and the resource object. From the library user’s point of view the service is improved because of the integrated, digital resource object that is delivered as the effect of the service.

### 4.3 Future service without librarian

It is possible to give library users remote access to the resource manager and thereby to offer a version of the service that is executed solely by IT systems. This implies that a library user interacts directly with search systems and with the resource manager. Consequently, the library users must perform the cognitive activities that are performed by a librarian in the previously discussed versions of the service.

The diagram in Figure 4 represents the future service without librarian in terms of a service scenario. The service scenario in Figure 4 differs from the service scenario in Figure 3 in one important way.

---

**Proceedings ECIS 2009**
The library user interacts directly with search systems and with resource manager without the mediating help of a librarian.

![Diagram](image)

Figure 4. Service 3 – Future service without librarian

From the librarian’s point of view the service frees him or her from serving all library users. This may give the librarian more time to other work activities. However, a potential disadvantage is that the resource database may be “polluted” by resource objects of low quality. From the library user’s point of view the service can be executed any time without the library user having to be present at the library. However, a potential disadvantage is that the library user cannot benefit from the knowledge and experience of librarians.

5 DISCUSSION

We have used the library case study to demonstrate how service scenarios can support modeling of important socio-technical aspects of business services. In this section we discuss some of the implications of the case study.

5.1 Implication #1 – Co-creation

*Service scenarios can be viewed as media for experiments, reflections, and decisions about different levels of co-creation.*

The future library service (Figure 3) cannot be understood solely in terms of value created by the librarian for the library user. The library user plays several active roles and participates in and contributes to the value-creating process. The library user interacts with the librarian by means of oral communication in which the user expresses information needs and the librarian suggests interpretations of the needs and asks clarifying questions. An isolated librarian cannot fully determine the value of an information search service. Only the library user can determine this value. And the librarian’s ability to use search systems in relevant manners is fully dependent on input, guidance, and evaluation from the library user.

The librarian uses a software system called a resource manager to create and maintain a shared resource called resource object that is visible to the library user. The resource object contains the (partial) result of the librarian’s search for information. During this search the librarian uses search systems like databases and Internet search engines. The results of these searches are visible to both the librarian and the user in terms of a shared resource called search result. The library user participates actively in the evaluation of search results and the resource object.

The level of co-creation is different in the two services that involve librarians in comparison to the service without the librarian. In the latter service the library user has to do all the cognitive work without any help from librarians.
5.2 Implication #2 – Interaction

*Service scenarios can be viewed as media for experiments, reflections, and decisions about external and internal interactions in service executions.*

The essence of a service is work performed in socio-technical networks in which human beings, technology, and other material objects interact.

COSMO is a framework that supports modeling of executors (COSMO uses the term provider) and consumers (COSMO uses the term user) in a manner that highlights interactions between the two parties (Quartel, Steen et al. 2007). Service scenarios and COSMO actor models are similar in the sense that they represent interactions between consumers and executors.

Service scenarios highlight the three interacting parts of a service system (executors, consumers, shared resources) as a socio-technical work system. Furthermore, service scenarios support modeling of internal interaction in executors. In the library case this corresponds to the internal interaction between librarians and search systems.

COSMO do not support modeling of interaction with shared resources and it does not support modeling of internal interaction in executors. Without the structural picture offered by service scenarios it is hard to use a COSMO model to communicate an overview of the totality of elements and interactions in a service system. In particular, it hard to identify the shared resources and their roles in interactions from a COSMO model.

5.3 Implication #3 – Shared resources

*Service scenarios can be viewed as media for experiments, reflections, and decisions about the roles and use of shared resources during service execution.*

All the three studied library services share the characteristic that the library user and the service share two important resources, i.e., the resource objects and the search results. A search result is an intermediate resource that is used by the librarian and the library user to evaluate potential information resources and their relevance for the library user. The resource objects contain the information resources that the library user finds relevant with respect to his information needs.

COSMO supports modeling of interaction between executor and consumer but it does not support modeling of shared resources. Use cases share this characteristic. Use cases are somewhat service-oriented in the sense that a use case is a part of a system that is supposed to create value for its user (Jacobson, Booch et al. 1999; Kruchten 2000). However, use cases are based on insufficient assumptions that does not support rich modeling of a consumer’s interaction with resources, human actors, and technology.

5.4 Implication #4 – Mediated interaction

*Service scenarios can be viewed as media for experiments, reflections, and decisions about the use of IT systems to mediate interaction.*

Interaction can be mediated by IT tools. In the library case the new services are based on a software component called a resource manager that mediates interaction between human beings and the resource objects that contains the information resources that are created during service execution. In the future service (Figure 3) the librarian interacts with the resource object via the resource manager. In the future service without librarian (Figure 4) the library user uses the resource manager.

Service scenarios support modeling of different levels of direct or mediated interaction and hence they support decisions about how the elements of a service system should interact. For example, they can be used to answer questions like the following. Should communication between librarian and library
user be face-to-face as implied by the scenarios in Figure 3 and Figure 4? Should it be possible for librarians and library users to interact via an IT system?

**Implication #5 – Human beings versus IT systems**

Service scenarios can be viewed as media for experiments, reflections, and decisions about the roles played by human beings and technology in service execution.

The library case study represents two types of services that are commonly used by businesses. The service without librarian (Figure 4) is executed solely by IT systems whereas the other service versions are executed by a combination of human beings and IT systems.

Service scenarios can be used to evaluate advantages and disadvantages of including human beings in service execution. The service scenario in Figure 3 indicates that the communication between librarian and library user has a value-creating potential in the sense that librarians can use their experience and broad access to search databases to improve the quality of the information search. The service that is modeled in Figure 4 does not have this potential but it offers a more flexible service access to library users that can operate without depending on the availability of librarians.

The future service without librarian can be called a self-service in the sense that no human beings participate in the execution. The library user is the only human being that participates when the service is executed. This does not imply that the library user does all the work. The involved IT systems perform work as well and there is a significant amount of co-creation occurring during the interaction between the library user and the IT systems.

**Implication #6 – Service networks**

Service scenarios can be viewed as media for experiments, reflections, and decisions about the roles of services in broader service networks.

The future service (Figure 3) and the future service without librarian (Figure 4) contain two interacting service executors. In both cases the resource database acts as a service for the storing of information resources. Also, the search systems in all three services can be viewed as services even though they have not been modeled as such. In the future service from the library case (Figure 3) the librarian acts as an executor intending to serve the library user in some situations. And the librarian acts as a consumer that uses a service executed by a resource database.

**Implication #7 – Processes and information**

Service scenarios should be supplemented by tools for modeling of information structures and processes.

Service scenarios do not support modeling of information structures and processes. COSMO uses information models to represent the information that is used during service execution. A process model represents the dependencies and temporal organization of the actions. COSMO process models may be used to model the processes executors and consumers. COSMO use information models to express preconditions for the actions of consumers and executors.

6 **CONCLUSION**

We have presented and discussed the notion of service scenarios. A service scenario represents a business service and the participating actors and their shared resources as a socio-technical activity system. Three types of actors participate. A provider offers a service to a set of consumers. An executor performs the work that constitutes a service. A consumer benefits from a service and co-creates its value. When a service is executed a consumer interacts with an executor in the sense that they communicate and exchange resources. The consumer provides input that is necessary for the executor in order to perform the service. The effect of a service can be characterized by the delivered resources and changes made to shared resources.
The use of service scenarios have been demonstrated by means of a case study in which information search services have been modeled by means of service scenarios. The case study demonstrates that service scenarios highlights some of the essential characteristics of business services as socio-technical systems characterized by co-creation, interaction, shared resources, mediation, human participation, and networks.

Future work includes more case studies in which the notion of service scenarios is evaluated and improved. In particular, studies of more complex services than the studied library services are needed in order to evaluate the advantages and disadvantages of service scenarios. We are currently studying medical e-consultations in which patients can communicate with and seek advice from doctors through dedicated web sites. And we are planning a study of consultancy services in farming.

Also, future work includes analysis of process modeling languages and approaches and evaluation of their relevance for modeling the processes of consumers and executors. This includes analysis and evaluation of languages like BPMN (White 2004), EPC (Dehnert 2002; Lübke, Lüecke et al. 2006), action-oriented modeling (Ågerfalk and Eriksson 2004), action diagrams (Goldkuhl 1996), event-activity diagrams (Bækgaard 2004; Bækgaard 2007), action-oriented development (Rittgen 2006), role-activity diagrams (Ödeh and Kamm 2003), resources-events-agents diagrams (Poels, Maes et al. 2007).

7 ACKNOWLEDGEMENTS

We would like to thank the anonymous referees for their valuable and constructive comments and we would like to thank The Danish Strategic Research Council (NABIIT) for support to the eGov+ project.

8 REFERENCES


OASIS (2006). Reference Model for Service Oriented Architecture 1.0, OASIS.


AN EMPIRICAL EXPLORATION OF REQUIREMENTS ENGINEERING FOR HYBRID PRODUCTS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0450.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Empirical study, Research in progress, Practice, Information systems analysis and design</td>
</tr>
</tbody>
</table>
AN EMPIRICAL EXPLORATION OF REQUIREMENTS ENGINEERING FOR HYBRID PRODUCTS

Berkovich, Marina, Technische Universität München, Information Systems (I17), Boltzmannstraße 3, 85748 Garching b. München, DE, berkovic@in.tum.de

Leimeister, Jan Marco, Universität Kassel, Chair for Information Systems, Nora-Platiel-Str. 4, 34127 Kassel, Germany, leimeister@uni-kassel.de

Krcmar, Helmut, Technische Universität München, Chair for Information Systems, Boltzmannstr. 3, 85748 Garching b. München, Germany, krcmar@in.tum.de

Abstract

In this paper we report on an empirical study on requirements engineering of hybrid products. Hybrid products (often also referred to as product service systems) – a combination of product, software and service elements – are an emerging trend on the market. Companies intend to offer holistic solutions for customer problems and not single products. The development of hybrid products differs from the development of “classic” products because of the high-level of technological integration of the elements that hybrid products consist of, the interdisciplinarity and the different lifecycles of their single components. We have conducted fifteen expert interviews to explore current practices in requirements engineering in three industries developing hybrid products: automotive, IT-consulting and system integrators, and medical technology. Our results show that most components of hybrid products are developed independently from each other. Based on our empirical insights we have identified requirements and challenges for the design of an integrated requirements engineering process for hybrid products.

Keywords: requirements engineering, hybrid product, software, hardware, service, empirical study, product-service systems.

1 INTRODUCTION

The success of a manufacturing company is directly related to the success of its products on the market. And the success of a product on the market depends on its capability to meet the needs and wishes of customers and environment (Beiter et al. 2006, Lindemann 2006, Nuseibeh and Easterbrook 2000). These points are essential for the development process. The development process is also affected by factors such as shortened development and production time, as well as cost pressure (Lindemann 2006, Spath et al. 2001). Further, the quality pressure – to produce better products than competitors – is growing. Thus, products are becoming more and more complex, and customer requirements are becoming more specific and leading to the development of individualized products and services built to individual customer needs (Abramovic and Schulte 2007). Customers have no interest in products or services per se, but they want their problems to be solved. The companies offer not a product, but a solution that is “a customized, integrated combination of products, services and information that solves a customer’s problem” (Sawhney et al. 2006). The distinction between products and services is disappearing and a so called hybrid product emerging (Berkovich et al. 2009, Knackstedt et al. 2008, Leimeister et al. 2009). In contrast to classical hybrid products consisting of a physical product and a service part, we also consider a combination of software and service, and software as a hybrid product (Böhmann and Krcmar 2007). Only by coordinating the development process of each element of hybrid products is it possible to achieve a customer’s benefit that is greater than if the product and service were to be bought separately (Schmitz 2008).
The development of hybrid products differs from the development of “classic” products because of the high number of sub-elements, the high level of technological integration and the degree of customer-integration (Leimeister and Glauner 2008, Pahl et al. 2006, Zellner 2008). The literature discusses products – physical products or software – and services separately. In practice, the development of products and services is also reported to take place separately (Leimeister and Glauner 2008, Müller and Blessing 2007).

One of the most important aspects in a development process is the successful handling of the requirements (Brooks 1987, Pohl 2007). Defining and handling the requirements is a very complex and difficult process. Deficiencies in this process often lead to costly project failures (Boehm et al. 2001). Keil (1998) finds that two of the three major risk factors in software development have to do with requirements engineering (RE). RE is especially important for hybrid products which consist of elements with different life-cycles; for example, hardware and software are developed by different disciplines, namely product, software and service engineering. Another aspect is hybrid products aiming at solving specific customer problems. Thus, the RE of hybrid products have to show and consider the different views of the involved disciplines on the RE, and also provide the customer-orientation and coordinated development. For the successful development of hybrid products, well-founded knowledge about the interdependencies of different characteristics of these products, their value-enhancement and the resulting output quality are necessary (Leimeister and Glauner 2008). Thus, it is a very gainful step to develop integrated RE for hybrid products.

This paper presents and discusses practice requirements for the RE process for hybrid products. In the study we first analyzed the literature on product, software and service engineering. Then we conducted a series of expert interviews on RE for hybrid products in practice. We collected empirical data from fifteen interviews in three industries – automotive, consulting and systems integrators, and medical technology. These interviews were the basis for our definition of the requirements and challenges for the design of the RE process for hybrid products from practice.

2 RELATED WORK

To understand the RE for hybrid products, it is important to comprehend what hybrid products are. It is necessary to understand the RE in general and to know how the RE is performed in the three disciplines of product engineering, software engineering and service engineering.

2.1 Hybrid Products and Software

Hybrid products consist either of physical products or software or a combination of these elements and service-components (Becker and Krčmar 2008, Beverungen et al. 2008). There are two main reasons why software is an essential component of hybrid products: First, implementation and consulting activities are very often combined into service bundles (Bonnemeier et al. 2007). Second, the importance of software is growing in general. For example, in the field of embedded systems, producers find that software development often takes more than 50% of the total development costs (Baskerville 1999). In the automotive domain it accounts for 20% to 40% of the production costs (Bender 2005). Today, software is not only a part of the product, but an indispensable part of it: Functionality realized by software cannot be replaced by hardware. Also, the value creation of modern systems relies on software (Liggesmeyer and Rombach 2005). On that account, the strengthened integration of the software development in the overall development process is necessary.

2.2 Requirements Engineering in general

RE is a discipline for handling the requirements, and is an essential part of the development process. Software engineering has already defined a profound understanding of RE as consisting of requirements elicitation, analysis, negotiation and validation. Change management and traceability of requirements are also parts of RE. Several studies illustrate that faulty implementations of the RE
process have an impact on the overall development process (Finkelstein and Dowell 1996, Hall et al. 2002). Failures arising during the RE process resulting in substantial costs for their elimination, especially in the latest phases of the development process (Boehm and Basili 2001, Pohl 2007). The process of RE is one of the most difficult phases of a development process (Damian and Chisan 2006). RE has to involve different stakeholders – users, customers, managers, domain experts, standards, laws, etc. – in order to get the requirements (Cheng and Atlee 2007).

The following two sections discuss each of the following phases of RE: requirements elicitation, requirements analysis, requirements documentation, requirements validation, and change management and traceability.

2.3 RE in Product Engineering

This section describes some aspects of the RE in the domain of product development. Our primary source of information was the top-10 textbooks, according to the amazon.de/amazon.com sales ranking (accessed on 09/17/2008). With the focus on the RE process, we selected only the books describing process models for product development (Ehrlenspiel 2002, Lindemann 2006, Pahl et al. 2006). These books contained useful references to other sources, which we then analyzed (Ahrens 2000, Jung 2006, Kruse 1996, Ulrich and Eppinger 2003, VDI-Richtlinien2221). To cover the latest developments in this area, we analyzed the journal articles of the years 1997 to 2008 and conference proceedings from 2002 to 2007. The Chair of Product Development of the TU München (http://www.pe.mw.tum.de) provided us with a list of relevant journals and conferences. The results of the analysis are presented below.

Requirements elicitation: The main source of requirements is the customer; other sources are legal documents, standards, existing systems, etc. In product engineering (PE), checklists are a widely used technique to elicit requirements (Ehrlenspiel 2002, Pahl et al. 2006). Most analyzed conference-papers and journal articles present minor improvements of existing methods for elicitation. The importance of the customer for a successful product development is emphasized by nearly all approaches, but is restricted to the first phases of the RE process (Jung 2006).

The main task of the requirements analysis is to derive technical requirements from customer requirements. The technical requirements have to be detailed, solution-independent and in a quantified form (Ahrens 2000). Most analyzed conference-papers and journal articles present variations of the QFD method or put their focus on the management of functional requirements.

Requirements analysis. This has the task of resolving conflicts between requirements. Lindemann (2006) and Kruse (1996) purport that it is important to prioritize requirements to resolve conflicts and to define the focus of the development.

Requirements documentation. This requirement is very important for the development process, as misunderstandings and misinterpretations are often caused by inadequately documented requirements. Most approaches agree on the basic rules on how requirements are to be formulated: solution-independent, unambiguous, positive formulation, etc. (Lindemann 2006, Ulrich and Eppinger 2003).

Requirements validation. The main goal of requirements validation is to ensure that the requirements express the product as the customer needs it. The analyzed approaches describe only the subtasks of checking the requirements for completeness and consistency.

The importance of RE for successful development processes has been recognized in product engineering, but it is often seen as merely the initial step of the development process where the customer places the order. Due to the nature of hybrid products, this view of RE is not sufficient. Hybrid products need an even stronger customer orientation, often leading to customer integration (Leimeister and Glauner 2008, Zellner 2008). Services and software are parts of a hybrid product, but the RE in product development mentions neither the integration of services, nor the parallel development of them. Interdisciplinary products and the integration of software development into the product development are handled only marginally.
2.4 RE in Software Engineering

RE is a discipline within software engineering “about defining precisely the problem that the software is to solve” (Cheng and Atlee 2007). We have selected the approach of Sommerville and Kotonya to show the main aspects of RE. This approach defines the phases of RE and takes into consideration the parallel and iterative execution of individual RE phases (see Figure 1).

![RE process according to Sommerville/Kotonya (1998)](image)

The activities of traceability of requirements and change management are also parts of RE. The following sections discuss the phases of RE and the role of the stakeholders in the RE process.

Requirements elicitation. Requirements elicitation is the first phase in the RE process. This activity is defined as “understanding of the goals, objectives, and motives for building a proposed software system” (Cheng and Atlee 2007). It aims at identifying the problem that needs to be solved (Nuseibeh and Easterbrook 2000). For a successful elicitation, it is necessary to consider all different stakeholders (Sommerville and Kotonya 1998). The most popular techniques are: interviews that can be defined as a discussion about the system in the development with different stakeholders, workshops based on a group discussion, scenarios showing interactions in the developing system, and observations and brainstorming (Aurum and Wohlin 2005, Pohl 2007, Somerville and Kotonya 1998). Prototyping can also be used to elicit requirements (Arnold et al. 2003).

Requirements analysis and requirements negotiation. In this phase the requirements are to be concretized. Afterwards the stakeholders negotiate on these requirements. The goal of the negotiation is to solve the conflicts between the requirements, to specify a set of consistent and complete requirements, and thereby to satisfy all stakeholders (Aurum and Wohlin 2005, Somerville and Kotonya 1998).

Requirements documentation. Documenting the requirements is a basis for other RE activities (Pohl 2007). The requirements and the changes of the requirements are to be documented carefully. The common art of documentation is the natural language documentation. Pohl (2007) suggests model based forms of requirements documentation, for example with UML.

Requirements validation. The process of validation means to ensure that “models and documentation accurately express the stakeholders’ needs” (Cheng and Atlee 2007). Boehm (1984) purports that validation should answer the question, “Am I building the right product?” The techniques applied for validation are reviews based on reading and analyzing the requirements, checklists, prototyping and walkthroughs (Pohl 2007, Somerville and Kotonya 1998).

Change management and traceability. Change management has to verify and evaluate changes of requirements and then to examine its impact for the system in development (Somerville and Kotonya 1998). Thus, change management has to guarantee that the proposed change is documented and analyzed, the costs for the implementation are checked, and the change can be realized. Another aspect
is traceability of requirements describing “the life of a requirement, in both a forwards and backwards direction” (Gotel and Finkelstein 1994).

We conclude that RE of software engineering is well elaborated, but not directly applicable to hybrid products. For example, hardware is seen as system context and therefore as unchangeable. The integration of services into the product is not considered.

2.5 RE in Service Engineering

Gill (2004) and Burr and Stephan (2006) define service engineering as follows: service engineering is the systematic development and organization of services by deploying engineering methods, practices and by using tools of the engineering design field. The focus of this paper lies in the use of appropriate methods and instruments.

In the literature there are some different views on service engineering:

- Burr and Stephan (2006) basically see service engineering as the process of developing services. The first step is to collect the ideas in a systematic and effective way. Afterwards the ideas are evaluated and the most auspicious ones selected and documented. Thereafter the requirements are elicited for which the needs and wants of all potential users have to be regarded. The last step is to design the service by defining the following information: technical goals, market goals, etc.

- Schmitz (2000) suggests that service engineering is handled in large parts by marketing divisions. According to him, RE for service engineering is a task that is carried out by qualitative marketing research. To elicit the requirements, the customer should imagine the service and should then express a hypothetical evaluation.

- Ramaswamy (1996) defines services as transaction processes. He separates the quality of a service into quality of service design and quality of service delivery. The service design consists of various steps. The step “Defining Design Attributes” realizes the RE in which the key-customers are identified and their expectations and requirements elicited and prioritized. It appears that RE for services is intended, but no precise methods are provided.

- A process model eliciting customer requirements during a “concept phase” is suggested by Schneider et al. (2006), but no methods are proposed for this activity.

Even though there are several process models for service engineering, they contain no applicable methods. Most approaches agree that it is necessary to elicit the requirements for the services and that the customer is the main source of these requirements. All in all, the role of RE in service engineering remains vague.

3 RESEARCH METHOD

In the literature review described above, we found that in the context of RE, hybrid products were not treated as a single unit. We identified several problems regarding the RE of hybrid products in different disciplines. To solve these problems, we conducted a series of semi-standardized expert interviews. The goal of these interviews was to capture current practices in RE in industry and to collect the requirements of an integrated RE for hybrid products. In the course of the study we interviewed fifteen experts in companies within the trade automobile, medical technology, consulting and system-integrators industry, who were involved in the design of hybrid products. These companies produced products consisting of parts that were typical for hybrid products. We perceived experts to be those employed in the field of RE or those who handled the requirements because of their position, such as project managers or business managers.

We selected three premium manufacturers from the automotive industry and six companies from the top-25 IT-consulting and system-integrators in Germany (http://www.luenendonk.de/it_beratung.php) of the consulting and system-integrator industry. From the medical technology industry, we selected four companies in which we had contacts. Table 1 shows the characterization scheme according to (IfM 2002).
The interview partner selection relied on theoretical sampling, which is suitable for studying problems that are not understood or not fully understood (Glaser and Strauss 1967).

The interviews were carried out between June 2008 and October 2008. Three interviews were carried out by telephone, and twelve were face-to-face. Table 2 gives a summary of the positions held by the interviewees.

<table>
<thead>
<tr>
<th>Size</th>
<th>Sales volume 2007 (in million €)</th>
<th>Employees 2007</th>
<th>Number of companies</th>
<th>Number of interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>big</td>
<td>50+</td>
<td>500+</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>medium</td>
<td>1 – 50</td>
<td>10 – 499</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>small</td>
<td>up to 1</td>
<td>Less than 10</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1  Company size & industry (according to (IJM 2002))

<table>
<thead>
<tr>
<th>Interviewee</th>
<th>Position of the interviewee</th>
<th>Branch of trade</th>
<th>Type of the interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>Member of the Project management group</td>
<td>Automotive</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Beta</td>
<td>Responsible for consulting</td>
<td>Medical technology</td>
<td>Telephone</td>
</tr>
<tr>
<td>Gamma</td>
<td>Responsible for requirements engineering of IT</td>
<td>Automotive</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Delta</td>
<td>Responsible for functional requirements engineering</td>
<td>Automotive</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Epsilon</td>
<td>Member of department of Business Engineering, main focus at RE</td>
<td>Automotive</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Zeta</td>
<td>Responsible for Requirements &amp; Usability Engineering</td>
<td>Automotive</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Eta</td>
<td>Responsible at the interface point to requirements engineering</td>
<td>Consulting/System-integrator</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Theta</td>
<td>Managing Director</td>
<td>Consulting/System-integrator</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Iota</td>
<td>Senior Manager Business Area Supply Chain Execution; focus at RE</td>
<td>Consulting/System-integrator</td>
<td>Telephone</td>
</tr>
<tr>
<td>Kappa</td>
<td>IT Management Consulting with the interface to RE</td>
<td>Consulting/System-integrator</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Lambda</td>
<td>Executive Partner</td>
<td>Consulting/System-integrator</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>My</td>
<td>Principal Health Care, interface to requirements engineering</td>
<td>Consulting/System-integrator</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Ny</td>
<td>Project manager</td>
<td>Medical technology</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Xi</td>
<td>Project manager</td>
<td>Medical technology</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Omikron</td>
<td>Senior Manager Business Area Supply Chain Exec., interface to RE</td>
<td>Medical technology</td>
<td>Telephone</td>
</tr>
</tbody>
</table>

Table 2  Companies

The interview guideline was structured according to the phases of RE as defined in software engineering (Aurum and Wohlin 2005, Somerville and Kotonya 1998). In our analysis we paid attention to the activities of each phase and to the methods applied there. Further, we were interested in the cooperation and interdisciplinary work of the involved disciplines and the role of the customer for the RE process. Another aspect which furthers staying in connection with the RE and with the communication with the customer is prototyping. Davis (1992) posits that using a prototype for the understanding of a problem or a solution helps to improve communication between the customer and the developer. The companies were also asked about the role of the RE in the service development and the integration of services in the development process.
The developed interview guideline was split into the following subject areas: general information about the company and the role of the interviewee, general information about the companies’ RE, all phases of RE (elicitation, documentation, analysis, negotiation, validation and verification), change management, prototyping, and integration of services (with a special focus on the composition of service- and product development).

The interviews were recorded when it was technically possible and only if the interview partners were in agreement. The records were then transcribed and analysed using Mayring’s (2007) techniques.

4 FINDINGS

This section summarizes the main results of the interviews. We take a look at each phase of RE as defined in the previous sections. The RE process in general, prototyping and the integration of services are described in greater detail.

Requirements Engineering process: Thirteen (out of 15) of the interviewed partners have an iterative process for RE that relies basically on models suggested in the literature. Only 2 of the interviewed partners do not have an iterative RE process since they deal with maintenance projects and use a change-request process. A minority of companies have a continuous requirements management: they not only keep records of the change requests, but also keep the requirements documentation up-to-date.

In software engineering, the development is usually done in multiple releases. The first release realizes the basic functions. In the development of the second, learning-effects take place and additional functions are realized. By developing in incremental release, the risk of the overall project is minimized.

All companies see the customer as the main source of requirements. A customer can be an external client or an internal department of the same company. Most companies distinguish between internal requirements and laws/standards. In the field of medical technology, the companies rely heavily on market analysis to elicit requirements, a field in which laws and regulations are particularly important.

Two interviewees reported that existing systems were an important source of requirements and needed to be analyzed thoroughly. All companies distinguished between functional and non-functional requirements. Nine (out of 15) of the interviewees reported that in the majority of cases the changes concerned functional requirements. In contrast, only 4 (out of 15) considered the majority of changes to be in the non-functional area. Two (out of 15) were unable to give detailed statements about changes.

Requirements elicitation: All interviewees reported using workshops and interviews to elicit the requirements. Only one interviewee reported observing the users to elicit requirements. Two (out of 15) of the interview partners did not have an explicit elicitation phase because they received the prepared requirements from their customers. The prevalent problems during the elicitation were that the stakeholders did not express their requirements clearly and the different stakeholders had different goals regarding the system in development. A further problem was the different understanding of the requirements by the stakeholders and the developers. In many cases the customer has unrealistic visions regarding the time and costs of the development.

Requirements analysis: With respect to requirements analysis, 14 (out of 15) of the interview partners had a process model. Only one company had no analysis phase since they received sufficiently detailed requirements from their customers. Two interviewees used tools to support the analysis. Three interview partners prioritized the requirements according to a “must-have/should-have/nice-to-have” scheme. The priorities were assigned according to law restrictions, financial aspects and the importance of the requirement for the customer. According to the iterative development described earlier, it is also possible to postpone the requirements until a later release. Only one company did not prioritize the requirements.
Requirements documentation: All interview partners used office tools to document the requirements, and thus the requirements were documented in natural language. Most interview partners used templates that precisely defined the structure of the requirements. One interview partner used EPKs (Scheer 1997) to document the business processes. For documentations, special tools were also used: five (out of 15) of the interviewees used Doors, two (out of 15) used CaliberRM. Two (out of 15) had other tools supporting the RE, especially the documentation. Twelve (out of 15) interviewees reported using Office-tools, among others, for the documentation of requirements. Office-tools have the big-disadvantage of not supporting multi-user operations and not supporting version-control. These tools were thus not adequate for the requirements management in large interdisciplinary projects.

Requirements negotiation: All interview partners reported conflicts between requirements, the reason for the conflicts being that customers and contractors held different views on the developed product. The common methods of resolving the conflicts were negotiating and holding workshops. The following aspects needed to be considered when resolving conflicts: dates, costs, and performance if the requirements were to be realized. Only one interviewee stated that conflicts were impossible because he had a consistent set of requirements from his customer.

Requirements validation has the task of asserting that the developed solution conforms to the requirements. The decision that requirements needed to be realized was taken mostly by customers. The realizing company only gave recommendations. If the company saw that it was impossible to realize a requirement, the customer was consulted. In the area of medical technology, the validation is mandatory by law.

Change management: Most interview partners used the change-request process to manage changes of requirements. This process was used when new requirements were added and when existing requirements were changed. Customers often dictated how the change management process was to be done. The expenditures for new requirements and changes were fixed in advance. When deciding whether a change-request should be realized, costs factors and time factors needed to be considered.

Prototyping: Prototypes are predominantly used to communicate with customers. The main audience for prototypes was therefore the customer and other stakeholders. The most common form of prototypes were Mock-ups created on paper or with office-tools. Two interviewees also used software-prototypes. The decision as to whether prototypes were to be used depended on the cost-benefit ratio due to the high cost of the development of a prototype.

Integration of services: Most companies of our study offered services offered in combination with software or the implementation of the software as part of the service. Typical services were management consulting and IT consulting. Management consulting included process- and organisation consulting, project management, business process analyses. IT consulting can include conception and implementation of software, security assessments, rollout tests, architecture definition, infrastructure and network consulting. An important aspect is the quality assurance of the services that is done according to standards.

5 DISCUSSION

On the basis of the interviews, we identified the following demands on the RE for hybrid products.

Coordinated RE process for the components of a hybrid product: The RE of the different components of a hybrid product has to be coordinated. Only by doing this can the requirements for the solution as a whole be found, and only in this way can the product fulfil the customer requirements for the entire solution. Today, the RE and development of single components, both in literature and practice, is done separately (Leimeister and Glauner 2008). This empirical study revealed that the RE for products and services was done in parallel but without coordination. The development processes do not consider the changes of single parts. Only at the end of the development the single components are brought together.
Support incremental development in releases: In software engineering, incremental development in multiple releases is applied successfully. For hybrid products, a support of incremental development is necessary. Incremental development has the advantage that the fully integrated product is tested not only at the end of the development, but also during different stages of the development process so errors can be found early and the risk of the product is minimized.

Improvement of the interdisciplinary work between the disciplines: The single disciplines often have only the overview over their own area and have a distinctive technical mindset with specific terms. These differences in terms and notions lead to interdisciplinary communication difficulties. It is very important to come to a common understanding of the problem to be solved. Only in this way is it possible to realize a harmonized communication between the disciplines. Therefore, already during the RE a common set of terms have to be defined in a glossary. That glossary is a central repository which will be used by all disciplines. In this way, misunderstandings are excluded a priori.

Stronger integration of the customer into the RE process: The interviews revealed that the customer was not sufficiently integrated into the development process. The companies did not entirely realize that the customer and his wishes were pivotal for the success of a product. The analysis of the requirements was done without customer-involvement. Thus, the requirements negotiation confronts the customer with concepts and terms of the realization domain which are unknown to him. On account of this, we requested the customer to be integrated into further RE processes, such as analysis and negotiation. The communication with the customer during these phases needs to be done in a way that it bears the context and domain of the customer in mind. Prototypes have good potential for this by visualizing the main concepts of the system in development.

Careful selection of stakeholders: The requirements on the developing system have multiple sources, such as customers, users, standards and laws. These sources of requirements are called stakeholders. “A stakeholder is a person or organization who influences the system’s requirements or who is impacted by that system” (Glinz and Wieringa 2007). The importance of selecting the right stakeholders is recognized by many authors (Pouloudi and Whitley 1997, Sharp et al. 1999). These authors agree that the selection of the right stakeholders remains a difficult area. A wrong selection of stakeholders or the selection of stakeholders which are not relevant for the system can influence the development negatively. At worst, phases such as requirements elicitation or negotiation have to be repeated. The selection of stakeholders is particularly important in the interdisciplinary context of hybrid products.

Better tool-support of the requirements management: Even though there are many tools for documenting and managing requirements, they are rarely used in practice. Instead, office tools (Excel, Word, etc.) are used.

Better traceability and a version control: In the context of interdisciplinary work, it is very important to have traceability and a version control system that offers the possibility of tracing which person changed a requirement and restoring old versions of the requirements documentation. The simultaneous work of different people on the requirements specification has to be supported. This so called multi-user capability is not realized by office tools. It therefore becomes necessary to introduce tools that are capable of coping with these demands. What is required is an analysis of whether tools on the market are sufficient or whether individual tools need to be developed.

Improving requirements negotiation: In highly interdisciplinary projects, the requirements negotiation phase is especially important because many different stakeholders with different backgrounds have to agree on a common set of requirements. This empirical study has shown that in practice, this phase is not supported by advanced methods. New approaches for the requirements negotiation using win-win methods and tool support by groupware (Boehm et al. 2001) could support this activity more effectively. What is required is an analysis of how these approaches could be adapted for hybrid products and introduced successfully.
Consideration of changes when the development is complete: Services should be provided after the development of the hybrid product is complete. The change of the product-lifecycle providing support has an impact on other aspects of hybrid products. The RE for hybrid products must be able to handle the requirements that emerge during the use of the product. Therefore, special methods and processes have to be developed in order to support further development of hybrid products.

Thorough documentation of the source of the requirements: The requirements on components of hybrid products are treated differently according to the discipline for which they are intended, and thus need to be documented for relevant disciplines.

6 OUTLOOK AND FUTURE STEPS

This study reports on 15 interviews carried out in industry in order to ascertain the state of practice of RE for hybrid products. This empirical study shows that RE for hybrid products is very important, but in practice the single components of hybrid products are developed independently, which is also true for the requirements. From these results, the requirements for the RE of hybrid products were derived. These requirements represent the main fields that have to be addressed in order to establish a comprehensive RE for hybrid products.

As a next step in research, we propose taking a closer look at the RE and development process of one single company. Based on our findings and this analysis, an integrated process for the RE of hybrid products should be developed. The single methods of the RE process also need to be adapted.

ACKNOWLEDGMENT

We thank the German Research Foundation (Deutsche Forschungsgemeinschaft – DFG) for funding this project as part of the collaborative research centre ‘Sonderforschungsbereich 768 – Managing cycles in innovation processes – Integrated development of product-service-systems based on technical products’. For further information, please visit http://www.sfb768.de.

References


VDI-Richtlinien2221 VDI-Richtlinien 2221.

**IOS project motivation as a determinant of project activities and business capabilities**

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0176.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>interorganizational systems, IT Project Management, Technology adoption, E-commerce (B2B / B2C / B2G / G2C)</td>
</tr>
</tbody>
</table>
IOS PROJECT MOTIVATION AS A DETERMINANT OF PROJECT ACTIVITIES AND BUSINESS CAPABILITIES

Smith, Stephen P., Monash University, Department of Accounting and Finance, Australia, and The University of Melbourne, Department of Information Systems, Australia, stephen.smith@buseco.monash.edu.au
Shanks, Graeme, The University of Melbourne, Department of Information Systems, Australia, gshanks@unimelb.edu.au
Johnston, Robert B., University College Dublin, School of Business, Ireland, robert.johnston@ucd.ie
Rahim, Md. Mahbubur, Monash University, Caulfield School of Information Technology, Australia, mahbubur.rahim@infotech.monash.edu.au

Abstract

Interorganizational systems (IOS) are a relatively common technology in mid-to-large organizations, and much research has been conducted into why firms adopt these systems, and the types of benefits obtained. However, IOS development projects, which is where the adoption becomes crystallized, and where specific ways to use an IOS are made possible, has been largely ignored and so are poorly understood. In this paper, we describe the IOS Motivation Model that explains how motivation determines project activities within the limits imposed by environmental affordances. We argue that motivation — the reason an IOS project is initiated — is a key determinant of the activities performed in a project and, through these activities, the types of business capabilities obtained. Two motivation types, activity control and market position, are explained in detail. Two case studies are then presented to exemplify IOS implementation projects that correspond to these two motivation types, and illustrate more clearly how motivation influences the pattern of activities performed in an IOS project. This is important for firms planning IOS implementation projects and for researchers in understanding the importance of IOS project-level motivation and its interaction with organizational strategy.

Keywords: Inter-organizational system, strategy, motivation.
1 INTRODUCTION

Interorganizational systems (IOS) are increasingly a standard component of the IT infrastructure for mid-to-large organizations. These systems automate information flows between organizations, typically to streamline the operations of a supply chain management system, but they may also be used to coordinate other boundary-spanning activities, such as product design. The large body of research into these systems has particularly looked at adoption influences and post-adoption impacts. Adoption studies explain the importance of selected environmental and organisational factors, such as partner pressure or IT readiness, on the intention to implement an IOS (Markus & Tanis 2000). Post-adoption impact studies, on the other hand, examine the relationship between system attributes (e.g. integration) and business outcomes (Subramani 2004). Few studies, however, have examined the intermediate step, the implementation process, which is where intentions, designs, and capabilities become actualized.

Understanding this intermediate stage helps explain IOS impacts, but involves more than just looking for general relationships between IOS characteristics and outcomes (e.g. integration and cost). As Subramani (2004) observes, to explain why particular outcomes are achieved, one must first know what the system was intended to achieve. A conceptualization of IOS projects that incorporates intentionality is therefore required. The idea that technologies are appropriated, proposed by DeSanctis and Poole (1994), is useful in this context because it helps to explain why diverse outcomes may be observed, even when the technology and the context of use are held constant.

We develop the concept of intentionality in IOS projects by proposing that project motivation determines project activities, and through them, business outcomes. Our theory builds considerably on descriptive exploratory work by Peffers et al. (1998), Rahim et al. (2006, 2007) and Smith et al. (2007) to provide detailed theory about the types of motivation that are possible, as well as specific project activities likely to be observed for these motivation types. In this paper, we focus on exchange (transactional) systems, a common type of IOS application. Collaborative IOS projects are described briefly, but an in-depth analysis is beyond the scope of this work.

This paper is organized as follows. We first theorize how motivation affects the implementation process for an IOS and describe in detail a typology of IOS project types, with that theory constituting the primary contribution of this paper. We then present two exemplar systems, in the form of case studies, to explicate these concepts. In other words, like Hirchheim and Klein (1989), cases are a used as a rhetorical tool rather than as a formal test of the theory. Finally, we describe the implications of this work for theory and practice, and directions for future research. The article provides a new vehicle for theorizing about the nature, purpose, and practice of information systems development.

2 THEORY AND PROPOSITIONS

Organisational motivation, the central concept in this theory, refers to the type of issue or deficiency that an organisation is trying to address when initiating an action. Motivation is related to technology capabilities in that a technology will usually be implemented to help solve a specific problem. IT solutions often provide capabilities beyond those specifically required, but one should not confuse project intent with technical potential. Project intent refers to the specific goals of a project, whereas potential refers to all capabilities regardless of whether they are actualised through the project.

Diverse reasons have been cited for using IOS technology, including enhancing cooperation, increasing bargaining power, improving process efficiency, and even survival (Johnston & Vitale 1988). These reasons are not surprising given the capabilities of IOS as a technology. Many reasons are “transactional” (in the broad sense of the term) in that they are concerned with improving the efficiency of organisational systems, particularly sales or purchase process; others are strategic in nature and aimed at changing or maintaining relationships with external bodies, particularly the relative amount of bargaining power held by partners. Dividing motivations in this way is common in the lit-
erature, with influential studies including Johnston and Vitale (1988) and Subramani (2004) modelling IOS impact as a function of whether the application is relational and transactional or just transactional. IOS technology, however, is useful for more than just transacting and managing supply chains. Some systems support cooperation between groups without any exchange of planning information, and the relational focus of these systems can also vary. We now explicate these issues.

2.1 Relational Investment

Because IOS technology is a communication tool that requires cooperation between organisations just to be implemented, an IOS project can be thought of as an investment in a relationship-based resource. This relational investment can be low (i.e. arms-length relationship), or could be high, meaning that the relationship is very close with significant integration, perhaps even to the extent that partners see each other as an extension of their own firm (Sahay & Mohan 2006). For a low relational investment, some idiosyncratic technology will be present (e.g. EDI template), but the primary focus of projects is on developing internal systems and processes to achieve an operational benefit such as cost reductions. These projects do not aim to expand business opportunities, only to improve internal operations. By contrast, when investment is high, the focus is on supporting more externally focused goals such as creating business opportunities or preventing the loss of existing business (Madhok & Tallman 1998). The motivation focus is therefore on developing resources that are embedded in (specific to) the relationship and non-salvageable should the relationship end. This type of electronic integration between organizations using inter-organizational systems has even been discussed as an alternative to actual financial ownership of adjacent value-added stages of the supply chain (Zaheer & Venkatraman 1995).

2.2 IOS Purpose

The inherent relational nature IOS technology (to communicate with a business partner) also means that any classification of motivations must take into account the cooperative intent of the system. Computer-based collaboration research commonly classifies cooperation on the basis of whether it is more transactional (market exchange) or collaborative (Markus & Christiaanse 2003). An exchange-oriented IOS enables arms-length trading and supply-chain management (e.g. EDI), whereas a collaboration system supports cross-organizational work-flow, particularly development or design work, and so will involve some type of cooperative work system. Web 2.0 tools are particularly suited to this application (Lee & Lan 2007). IOS use for transacting and IOS use for collaborating are therefore conceptualized as distinct patterns of appropriation for interorganizational technologies.

2.3 IOS Motivation Types

We combine relational investment and IOS purpose in Figure 1 to form four motivation types, which we label activity control, market position, long-term partnership, and short-term project. The activity control motivation type is internally oriented, primarily concerned with transactional efficiency via automation. For the market position motivation type, the project is concerned with securing business opportunities and developing a relationship, although project work is directed at facilitating exchange rather than supporting cooperative development or design. The short-term project motivation type is concerned with supporting cooperative development or design, but the relationship is not expected to endure, and so relational investment is kept to a minimum. Finally, the long-term partnership type is both cooperative and involves a significant relational investment. For example, the project may be intended to provide or support a significant level of operational integration between firms.

\begin{center}
\begin{tabular}{|c|c|}
\hline
Relational investment & IOS Purpose \\
\hline
High & Exchange \\
& Market position \\
& Activity control \\
\hline
Low & Collaborate \\
& Long-term partnership \\
& Short-term project \\
\hline
\end{tabular}
\end{center}

\textit{Figure 1: IOS Motivation Types}
3 MOTIVATION – PROCESS – OUTCOME CHAINS

We propose that the motivation behind an IOS project (the reason a project is initiated) determines the pattern of activities performed, but within the limits imposed by environmental affordances, and that each pattern will result in a firm obtaining or enhancing particular business capabilities. Figure 2 presents the model examined in this research. For IOS projects, it shows relationships among patterns of motivations, affordances, activities, and business capabilities.

In our model, motivation is not the same as business strategy. A business strategy is a general pattern of organizational behaviour (Miles & Snow 2003), whereas project motivation is the justification for a specific project. Because motivation expresses (or is closely related to) a specific deficiency, it is reasonable to assume that activities will be structured accordingly, but this is not to say that organisations always achieve what they intend, only that intentionality influences technology outcomes. This reasoning is consistent with evidence from studies of IT project behavior, which show that the specific problem and the way in which it is framed influence the types of issues considered in system investigation and the analysis activities performed (Keil 1995). Investigation and analysis, in turn, affects how the business problem is actually solved, and the specific solution implemented (Markus & Tanis 2000).

![Figure 2: IOS Motivation Model (Motivation, affordances, activities, capabilities)](image)

3.1 Affordances

Environmental affordances are possibilities for action created the environment, and so both allow and suggest particular actions. An affordance can be embodied in objects or social practices, but should not be confused with the object/practice. Rather, it is an opportunity presented by an environment to perform specific actions (Gibson 1977). For a project, environmental affordances can be internal or external (to the project). Internal affordances are created by characteristics of the project, such as experience or budget, but may also emerge through project activities (Markus & Tanis 2000). External affordances are circumstances and practices outside of the project including business strategy, the strategy and capabilities of external partners, and industry standards.

Business strategy is perhaps the most important external affordance (although partner strategy and capabilities will also play a role). The Miles and Snow (2003) typology of organizational strategy is used widely to describe four patterns of strategic behaviour: defenders, prospectors, analyzers and reactors. Defender organizations specialize in a market niche, and defend that niche to ensure a stable and secure market position. It is common for these organizations to be cost-focused, and have a high reliance on formal procedures. Prospectors focus on diversification and growth, and so are always on the lookout for new market opportunities. Analyzers combine defender and prospector strategies. Efficiency and flexibility are both important for these organizations. Reactors, the final category, do not have a stable strategy, but instead respond to market conditions on an ad-hoc basis.

3.2 Business Capabilities

Business capabilities, the final component of the model, represent project outcomes. At this point, specific appropriations are embedded in the technology but still latent (to be crystallized when the IOS is
used), and particular business capabilities enabled. The specific technical capabilities enabled depend on the technology implemented (or enhanced) during the project. However, because it is here that the IOS moves from development to the production environment, research into the relationship between IOS attributes and business outcomes is relevant. The relational investment dimension primarily describes whether these outcomes will enhance internal or external capabilities. A small relational investment implies an internal focus, for example improving efficiency, monitoring, and coordination within the firm, whereas a large relational investment implies a more external focus on enhancing inter-firm relationships, or improving monitoring, and coordination between firms (Child et al. 2005).

The IOS purpose dimension, at the top of Figure 1, describes activities as exchange focused or collaboration focused. Exchange-focused projects tend to result in operational improvements, particularly more efficient sales and inventory management processes, and better information which improves monitoring of performance and better coordination of supply-chain activities (Chatterjee & Ravichandran 2004). IOS-based communication here will be mainly unidirectional, and involve exchange of structured transaction-related data such as EDI documents (Barrett & Konsynski 1982). Collaboration focused projects attempt to integrate processes, and so involve collaboration and coordination work activities, not just outcomes (Markus & Christianse 2003). IOS-based communication here will be bi-directional, and may involve rich unstructured communications (Sánchez & Pérez 2005).

4 APPLICATION OF THEORY

Two motivation types, activity control and market position, are now explained in detail. They represent extreme relational investment positions, and so allow us to describe contrasting effects. These motivations are shown in Figure 1 within the exchange segment, and so represent motivations for exchange applications, perhaps the most common of all IOS applications.

4.1 Activity control

The activity control motive is a long-established reason for entering cooperative relationships (and IT projects in general. Essentially, organizations experiencing financial pressures try to reduce the cost of administrative activities to offset perceived weaknesses (Child et al. 2005). Activity control has been found to motivate many IOS implementations, with these systems designed to reduce the cost of document handling and storage, and to create indirect savings by lowering the incidence of data-handling errors and data-processing related delays in business processes (Riggins et al. 1994).

Initiation: If an IOS project is initiated as a cost-saving initiative, the business case for the project will be based on financial benefits, particularly cost-savings (Peffers et al. 1998). Johnston and Vitale (1988, p. 160) characterize the approach as being driven by the question “how do I automate what is being done?” rather than the more strategically driven question of “how could our firm's objectives best be accomplished?” The immediate consequence of this focus is that the project is a short-term investment requiring a quick payback. Examples of this approach are described in detail by Peffers et al. (1998) and Rahim et al. (2007), with IOS investment motivated by the goal of lowering data entry errors and labor costs. Automating achieved these goals, but long-term concerns such as improving information quality or strategic positioning were ignored.

Development: This type of project is concerned with short-term cost-efficiency, and so is operational rather than strategic. Project activities, costs, and the payback period will be specified in detail, and there will be a clear allocation of project responsibility. In the development/execution phase, project performance measures and rewards will also emphasize cost-control, with the result that project monitoring will involve frequent meetings and focus on the consumption of rather than delivery of outputs (Lockamy & Hit 1995). The emphasis on detail suggests that a predictable and highly structured methodology (e.g. SDLC) may be preferred, although due to the low-cost, short-term emphasis of this motivation, a highly time-focused light methodology, particularly Agile methods such as Scrum or Extreme Programming, may be more common (Rico 2007). Post-implementation, this motivation is likely to involve a formal review, but mainly in terms of whether the project was run efficiently.
Post-implementation: This motivation involves a transaction-focused project and a small relational investment. Business capability improvements will therefore be internally oriented (e.g. improved activity monitoring or process efficiency), and focus on operational systems, particularly inventory management and exchange systems. In addition, IOS-based communication will be essentially a unidirectional transfer of well-defined and highly structured transaction-related data such as EDI documents.

Risks: A problem with the activity control motivation, however, is projects that cannot be shown to reduce costs in the short-term will not be approved even if they will ultimately provide strategic benefits. In addition teams may be rewarded for completing a project quickly at the expense of quality.

4.2 Market position

Under the market position motivation, cooperation is a mechanism for changing relative dependencies within an alliance. The purpose of cooperation here is to increase or protect the “top line” revenue stream of the organization (and to build long-term profitability), as opposed to the expenditure control focus of the activity control motive. This motivation is therefore externally-focused. This type of cooperative motivation was first described by Emerson (1962) in his seminal paper on power-dependence relationships. It holds that the power-dependence status within an alliance is changed when an organization invests in relationship-specific assets to make its business partner dependent on it, thereby producing codependence in the relationship. This power strategy is also often described using transaction cost economics terminology, which characterizes these partnerships as constituting: (1) non-salvageable, relationship-specific investments such as customized processes, (2) substantial information exchange via more direct access to partner systems, and (3) interdependent technological and functional systems (Child et al. 2005, Zaheer & Venkatraman 1995). Subramani (2004) illustrates this motivation in his description of the development by a mattress manufacturer of a make-to-order and quick delivery IOS to help a retailer to reduce inventories while improving customer service. The retailer ultimately became reliant on the manufacturer’s system, and so helped to replace an arms-length relationship with a more idiosyncratic partnership.

Initiation: Overall, this type of project is concerned with long-term strategies (e.g. reduce underlying inefficiencies) and so framed as a strategic issue. Activities conducted in the investigation phase will therefore be concerned with how the project contributes to the firm’s ability to achieve its strategy.

Development: In the development/execution phase, an iterative methodology may be preferred to allow for input from both internal and external parties, and project performance (and rewards) will be based on results. Due to its strategic importance, this type of project will be championed by a senior figure, and external orientation means that the business partner will have substantial involvement.

Post-implementation: Post-implementation, this motivation is likely to involve progressive identification and elimination of problems, particularly by prototyping new ideas and enhancements to further enhance the system. As with the activity control motive, IOS-based communication will support well-defined and highly structured transaction-related data such as EDI documents. However, communication between partners about the IOS (and related activities) will be ongoing.

Risks: This motivation also carries risks, however. Firstly, a long-term focus makes it easier to excuse poor project management, such as an extending deadline or a project that devours resources without ever reaching a conclusion (Keil 1995). Secondly, although long-term relationships tend to be highly beneficial, investing in the relationship may actually reduce long-term flexibility (Child et al. 2005).

Activities for these two motivation types are summarized in Table 1. Each is an ideal type, in that each describes project activities and outcomes that are theorized to be associated with a particular motivation. No actual project should therefore be expected to conform exactly to a theoretical account. Furthermore, the projects we describe as resulting from a given motivation are complex processes with multiple dimensions, but are theoretical exemplars, not categories of actual projects. The processes and outcomes we propose are, in effect, hypotheses regarding relationships between constructs. However, unlike the traditional hypothesis-testing approach, a typological theory such as ours does not
specify relationships between a unidimensional first-order construct and a dependent variable. Rather, it describes patterns among the constructs for each ideal type, and how each pattern produces specific observations for dependent variables (Doty & Glick 1994). Our theory should therefore be thought of as one that allows a researcher to assess the similarity of an actual project motivation to an ideal type, and on that basis to test relationships between the motivation and project activities.

<table>
<thead>
<tr>
<th>Project phase</th>
<th>Activity</th>
<th>Activity control</th>
<th>Market position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation</td>
<td>Senior management approval required</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Focus of business case</td>
<td>Short-term payback</td>
<td>Business position</td>
</tr>
<tr>
<td></td>
<td>Work with multiple business areas</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Clear allocation of project responsibility</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Calculate project cost in detail</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prepare a thorough project plan</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Specify project success criteria</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Development</td>
<td>Work closely with multiple business areas</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Project owner or “champion”</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Confront and resolve conflict</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Frequent progress meetings</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Monitoring method</td>
<td>Inputs</td>
<td>Output/Result</td>
</tr>
<tr>
<td></td>
<td>High-level of partner involvement</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Post-implementation</td>
<td>Formal review</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Formal procedure to identify &amp; fix flaws</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prototype new ideas/enhancements</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Table 1: Summary of project activities for each IOS motive

5 RESEARCH METHOD

A multiple case study approach (two cases) was used in order to explicate theory using real IOS projects (Yin 2003). The unit of analysis is an IOS implementation project within an organisation, concentrating on the construct IOS motivation type. We focus in this paper on the "exchange" purpose within the IOS Motivation Type model, on the basis that transactional systems, such as EDI, are a dominant IOS application, and therefore of great practical significance (future work will report case studies for other motivation types). The two cases assessed represent the two motivation types theorized for the exchange purpose, and therefore provide theoretical replication (Yin 2003).

To allow us to assess whether differences in IOS project activities were a result of differences in motivation, we ensured that external and internal affordances to the projects were as similar as possible. This meant selecting firms that were as similar as possible. In particular, each had recently implemented a new IOS, operates at a similar level in the same industry (Tier 1 suppliers to automotive assemblers), and employs a similar organisational strategy (defender). Profiles of each of the two firms involved in the case studies are provided below.

In the data gathering phase, we were given full access to key decision-makers closely associated with the IOS implementation project (three senior managers with technical and/or business roles in each firm) who explained the specific issues that management was hoping to address, the implementation process, outcomes, and challenges faced. Data collected was first coded to indicate the construct discussed, and codes then analyzed using pattern matching logic (Yin 2003) to assess the whether the patterns predicted for each motivation corresponded to the patterns actually found in the case data. The case studies are considered to be exemplars of each of the two motivation types.

Carolina is a small-to-medium sized enterprise that operates out of three locations, two in a southern suburb of Melbourne, Australia, and one in South Carolina. It has a turnover of around $US 10M, is privately owned, employs approximately 120 staff, and manufactures dies, metal-cast products, and injection-molded plastic components, primarily for the automotive industry. Information processing is
conducted using desktop computers, with most of the materials planning and other quantitative work performed using in-house developed Excel-based applications. Carolina’s production system is primarily based on just-in-time principles, with most production lines employing Kanban principles and supported by lean inventory management. Carolina produces for all major automotive assemblers, but is dependent on Toyota for the bulk of its business. The case study is concerned with a major IOS project involving the implementation of a Kanban-based supply components system for Toyota.

Clayton is the regional subsidiary of a major Tier 1 supplier and has business relationships with every local assembler as well as component retailers. The automotive division that is the focus of this study is located in Melbourne, and employs around 1,500 staff. Information processing is conducted using a sophisticated SAP-based system (installed in 2000), with many custom-developed routines and reports. EDI capability was initially implemented around 1995 at the request of major customers. This capability has been progressively extended in response to further requests from major suppliers and customers, and the company is gradually introducing EDI communications with smaller customers and suppliers. The production system is also based on just-in-time principles, with a mixture of Kanban production lines and MRP demand forecasting. Clayton is not completely dependent on any one customer, but neither does it see itself as being the dominant party in its relationships. The case study is concerned with a recent major IOS project involving a consignment stock system.

6 CASE STUDY DESCRIPTION

The business strategy as described by senior managers in both case firms corresponds closely to the defender type. Both businesses are specialized and operate in a mature industry (primarily motor vehicle parts manufacturing). For a defender, the main entrepreneurial challenge is to ensure a stable and secure market position in the long-term. They tend to protect their position by implementing and refining established and standardized technical processes to ensure that activities are highly efficient. This approach is evident in statements from managers of each company:

“We need to embrace lean manufacturing to become more cost-effective, and if we are cost-effective, Toyota is looking at its list of suppliers now. We have been told that they’re looking at us supplying on the next model, which is beyond 2011, so it’s almost a situation where we’re doing it for two reasons: for the ongoing success of the business, but also to appease Toyota. We want to be on the shortlist [of the ten preferred suppliers].” [Carolina Managing Director]

“In this company, you can’t get too far without convincing management that what you want to do will save some money. You will find that is probably the case in any automotive business because it is the toughest. Therefore the margins are minimal, but the turnover is huge. It is immediate and longer-term cost-savings we are after. So cost savings must be direct, like on every sale, and indirect, like in terms of running the business cheaper and responding to a customer’s requirements quickly. A happy customer means potential more customers, and more sales.” [Clayton Manager]

Defender projects are characterized by a focus on improving the efficiency of existing operations, and careful top-down planning. However, each firm is achieving this objective in a different way. Carolina has a “whole of business” approach that requires multiple areas (e.g. warehouse, manufacturing, purchasing, accounting, sales) to make complementary changes to processes. In addition, it is more concerned with eliminating waste in the long-term than controlling short-term project expenditure:

“[Toyota] tell us to do all this, but we have our own plans. We started on lean [manufacturing] before Toyota got involvement with our company. When they found out we had spent quite a lot of money putting lean [manufacturing processes] in, they brought us into their top 10 preferred companies and they’re now spending a lot of money with us [on process improvement projects].” [Carolina Managing Director]

“We use the 5S methodology [an iterative waste-elimination method]. …With just-in-time you have to be really tight. That’s why we can’t rush [the implementation] or we’ll fall apart. …it has cost us a fair bit so far. I don’t think we will recover the cost soon. The previous system was simpler for us. Eventually, when we go completely Kanban we will recover it.” [Production Manager]

Clayton, on the other hand, is very concerned about controlling project expenditure, and relies on frequent progress meetings, and strict adherence to a defined project methodology to ensure that projects finish on time and on-budget. There is little evidence of outside partner involvement.

“For every project we have to specify the benefits. This includes qualitative benefits and cost-recovery.
Then we have to provide the estimate of the cost as part of the budget preparation. Once everything has been approved, we create an internal order, with a budget. It says "this is how much we allocate for this project, this is the cost-centre". As the project is running, I will check cost against budget to see how we are going. How often we do a major review depends on the project. As part of the project preparation we actually specify how often we are going to review, not only cost, but any activity on the project compared with the plan. So for this project, which is an 8 million dollar project, we have what we call four gates: the critical milestones for the project. We review the plan, what was achieved, issues, risk analysis, and costs. We also do a weekly review where the team meets to discuss the implementation." [Clayton Manager]

Post-implementation, each has a different review and follow-up mechanism, and each even assesses the benefits in quite different ways. Clayton’s review mechanism, like project execution, is highly formal, but the benefits are seen in operational terms (not strategic), and so project success and evaluation of future opportunities are similarly operational:

“We also have post implementation follow-up 5 years after the project. What we do, for example, we look at technical issues. What sorts of errors do we get, what is the reason for those errors? Second part is looking at the benefits to the business. When we started EDI, we started seeing the benefits of it straight away. It is much more reliable, less error-prone than manual entry. Now when we finish an EDI process with one of our customers and they see the benefits of it, we go out and advertise it to our other customers, and say 'look we will both benefit if we go EDI.'" [Clayton Manager]

For Carolina, the continuous improvement philosophy adopted (5S) means that the project is under constant review, with frequent re-evaluation of performance, and regular prototyping of innovations. Benefits from the IOS are seen quite differently. For example, improved information and internal control, and a more secure customer relationship were cited at various points during the interview, but controlling the cost of administrative procedures was not mentioned at all:

"Definitely we have learned lessons from it. The idea of visual management and all that sort of stuff. I think we are taking as much as we can from it. The information we are getting is a lot better. I mean it has benefited us because it has smoothed our production out. It has gotten rid of the fluctuation, and as we slowly increase in demand, we can see what we need to do to cope.” [Carolina Production Manager]

7 CASE STUDY ANALYSIS

Table 2 summarises the above observations. Each case has been selected because it conforms closely to an ideal type: Clayton to *activity control*, and Carolina to *market position*. Both firms are trying to maintain good relationships with customers (to defend market share), and have justified projects on the basis of efficiency improvements, as one would expect of defenders. However, Clayton’s IOS project approval process required a detailed justification and to show immediate cost-savings with a short-term payback, whereas Carolina’s project was expected to cement the firm’s relationship with a major customer, and so protect or expand future revenue streams. In other words, Clayton’s initiation process is inward looking, attempting to meet objectives by minimizing process costs, and Carolina has an external focus in which it attempts to strengthen a relationship in the short-term, while hoping that “cost-downs” generated by the project will recover expenditure eventually. On this basis, Clayton’s motivation is more like the “activity control” type, and Carolina’s like the “market position” type.

Development activities are similarly distinct, and consistent with the respective pattern established in the initiation phase. That is, consistent with the internal focus of the activity control motivation, Clayton has little business partner involvement, inventory-control and IT were the only departments with any significant involvement (and only IT was involved in the actual development phase) and treats the project as more of an operational concern with stringent controls over time, expenditure, and activities. For Carolina, by contrast, the project is a strategic concern, evidenced by the seniority of the project owner, the significant (and expensive) preparatory work to ensure that the project does not fail, and the iterative method employed, which focuses monitoring efforts on results rather than inputs.

Finally, in the post-implementation period, Clayton is continuing its reliance on formal procedures, with a defined review cycle, and a well-defined process for identifying and fixing problems. Carolina, on the other hand, consistent with its outcome orientation, has no formal review procedure, but is instead monitoring production continuously to both address problems as they occur, and enhancing the system by prototyping new ideas (mainly adapting control techniques used by business partners).
Some risks are evident, however. For Clayton, the internal focus of projects (only cost-reduction projects are approved) may be problematic in the long run. In the short-term, the project reported may help Clayton to differentiate itself on the basis of cost, but that strategy is rarely sustainable (Miles & Snow 2003). Carolina’s long-term focus is also a problem, particularly given the difficulties facing the automotive industry as significant parts of the world experience a severe economic slowdown. In particular, by adopting a long-term focus, Carolina risks over-investing in efficiency and in binding its future to the success of a single business partner. A severe economic slow-down could mean that long-term expectations of growth will not be met, making recovery of short-term expenditure much less likely. In addition, any problems experienced by the business partner will now affect Carolina far more directly (although the alternative may nevertheless have been even worse).

<table>
<thead>
<tr>
<th>Project phase</th>
<th>Activity</th>
<th>Clayton</th>
<th>Carolina</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation</td>
<td>Senior management approval required</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Focus of business case</td>
<td>Short-term payback and intangible benefits</td>
<td>Business position</td>
</tr>
<tr>
<td></td>
<td>Work closely with other business areas</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Allocation of project responsibility</td>
<td>Large operational project run by project manager</td>
<td>To senior manager</td>
</tr>
<tr>
<td></td>
<td>Calculate project cost in detail</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prepare a thorough project plan</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Specify project success criteria</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Development</td>
<td>Work closely with other business areas</td>
<td>No. Project confined to a single business area</td>
<td>Yes, significant from most business areas</td>
</tr>
<tr>
<td></td>
<td>Project owner or “champion”</td>
<td>Senior manager as sponsor &amp; champion</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Confront and resolve conflict</td>
<td>In regular team meetings</td>
<td>No conflict reported.</td>
</tr>
<tr>
<td></td>
<td>Frequent progress meetings</td>
<td>Weekly meetings</td>
<td>Only occasional</td>
</tr>
<tr>
<td></td>
<td>Monitoring method</td>
<td>Inputs vs. budget</td>
<td>Output/Result</td>
</tr>
<tr>
<td></td>
<td>High-level of partner involvement</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Post-</td>
<td>Formal review</td>
<td>Major review after 5 years</td>
<td></td>
</tr>
<tr>
<td>implementation</td>
<td>Formal procedure to identify &amp; fix flaws</td>
<td>Yes</td>
<td>No. Ad hoc fixes</td>
</tr>
<tr>
<td></td>
<td>Prototype new ideas/enhancements</td>
<td></td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 2: Summary of project activities observed for each case study

8 DISCUSSION, LIMITATIONS, AND CONCLUSION

Our motivation model (Figure 2) proposes that project activities and outcomes depend on the motivation and environmental affordances. According to this model, process and outcome are predicted by what the motivation is able to determine within the constraints of the environment. The two cases help to illustrate the effects of environmental affordances on project activities and business capabilities because both firms have essentially the same business strategy (defender) in the same industry (Tier 1 suppliers in the automotive industry), and both were engaged in a major IOS project. Both firms are highly cost-focused, but, consistent with the typology each has enacted that strategy differently.

A number of limitations must be acknowledged. Firstly, all typologies necessarily simplify reality, particularly ones based on ideal types because these ideals are theoretical abstractions. Indeed, the specific cases reported here were selected because they correspond quite closely to the ideal types described, and so help explicate concepts without any need to explain differences between theory and observations. This research should therefore not be thought of as a test of theory in the hypothetico-deductive tradition, but instead as a form of theory development with detailed explication in the form of multiple case studies. More complex situations that do not clearly belong to any one cell are bound to occur, and these are a challenge to interpret. The simplified view of reality is nevertheless beneficial because it can help explain the principles underlying more complex situations.
A second limitation relates to the top-down (theory-driven) nature of the analysis. Because case details have been gathered and then compared to existing theory, retrospective sense-making (in which people interpret facts selectively according to what is known about the world) is a potential threat to validity. To counter this threat, motivation has been identified on the basis of explicit statements by managers (i.e., the technology application and depth of relationship were apparent from the outset), all cases have been assessed using the same template (thereby limiting the ability to exclude contradictory facts), and recording a rich description of activities has allowed us to assess the manner in which each is consistent with theoretical principles (assessment is not based on an arbitrary yes/no decision).

A third limitation is our assumption that intent is fully clarified and specified (or clarifiable and specifiable). That this assumption will not hold in all instances is a boundary condition on the application of the theory, rather than a conceptual limitation. If intent is unclear or unknown, application of the principles will be difficult, but this does not render the ideal types invalid.

Our findings have several implications for practice. Firstly, the motivation to adopt a system appears to affect implementation activities, including the design of the system and how it will be used. This is significant for managers because such design choices, once implemented, are usually expensive and difficult to modify. Secondly, relying on the experiences of other firms as a basis for system implementation decisions is clearly risky without in-depth knowledge of a project, particularly of the motivation. That is, simply knowing that a given firm has a similar structure, strategy, or technology is not a sufficient basis to assess another firm’s implementation project.

The major theoretical contribution of this work is our framework for understanding the role of motivation in the activities conducted when implementing a system. Our research model illustrates how motivations affect activities and outcomes for an IOS project, and so informs decision makers about how low-level concerns (project aims) can influence which technology capabilities are actualised, and even long-term plans. In addition, in contrast to the common views of project alignment, our theory does not assume that if project activities are aligned with business strategy, business goals should be met. As a general rule, one would expect project goals and corporate strategy to be aligned, and for project impacts to be highest when this is the case. However, the cases presented here indicate that, although strategy is an environmental affordance that suggests particular activities and constrains others, a given project could be structured in a wide variety of ways that are all consistent with the strategy and yet present quite different outcomes, opportunities, and risks to the business. In other words, the high-level nature of strategy means that the boundaries for project outcomes are broad. It is therefore not enough to align a project with corporate strategy; managers must also understand the specific implications that a given motivation has for project activities.

The case evidence presented explicates our argument that motivation affects the pattern of activities performed in an IOS project. The theory applied here is the basis for a wider research program, currently in progress. In future work we will describe other motivations and provide case examples for each. It is important to note that although IOS technology is the specific focus of our program, there is no reason to suppose that the principles we propose do not apply equally to other IT projects.

References
WHAT WE NEED: PROJECT MANAGERS` EVALUATION OF TOP MANAGEMENT ACTIONS REQUIRED FOR SOFTWARE DEVELOPMENT PROJECTS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0333.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Critical Success Factors (CSFs), Information Systems Development (ISD), IT Project Management, IT/IS management</td>
</tr>
</tbody>
</table>
WHAT WE NEED: PROJECT MANAGERS` EVALUATION OF TOP MANAGEMENT ACTIONS REQUIRED FOR SOFTWARE DEVELOPMENT PROJECTS

Madanayake, Ochini, The Australian National University, 224, Sir Roland Wilson Building, Canberra, ACT 0200, Australia, ochini.madanayake@anu.edu.au

Gregor, Shirley, The Australian National University, 2056, Pap Moran Building, Canberra, ACT 0200, Australia, shirley.gregor@anu.edu.au

Hayes, Colleen, The Australian National University, 2045, Pap Moran Building, Canberra, ACT 0200, Australia, colleen.hayes@anu.edu.au

Fraser, Steven, The Australian National University, 1013, Pap Moran Building, Canberra, ACT 0200, Australia, Steven.Fraser@anu.edu.au

Abstract

Software development projects still fail at an unacceptable rate although prior studies have identified critical success factors needed for project success. This study contributes to the project management literature by providing further insight into the nature and role of `top management support` (TMS), which is widely recognised as one critical factor in the success of software development projects. The study seeks insight into the nature and role of TMS from the perspective of software development project managers and their perceptions of actions required by top management in facilitating project success. A qualitative case-based approach was employed. Sixteen top management `actions` are identified, and subsequently framed by a conceptual model consisting of three top management roles: strategy, facilitate and lead. The study represents the first stage of an ongoing research program. The model will be tested in the Asia-Pacific region in the second stage. The expected final outcome of the research program is a framework that will support project environments by defining top management actions needed to support a software development project in different stages of the life of a project.

Keywords: top management support, critical success factors, project success, project manager, top manager, software development projects
1 INTRODUCTION

Software development is a highly complicated task with the involvement of many stakeholders and previous studies have indicated that low software development project success rates are a major concern (PMBOK, 2004; Reel, 1999; Meredith & Mantel, 2006). Many studies have been conducted to identify the factors that affect the success of projects, leading to a body of knowledge referred to as the Critical Success Factors (CSFs). Although the rate of project failure has declined with the application of accumulated knowledge, software development projects continue to fail at an unacceptably high rate (PMBOK, 2004; Schwalbe, 2006).

Academics and practitioners alike acknowledge the importance of CSFs (Reel, 1999; Butler & Fitzgerald, 2006). However, mere acquaintance with the factors does not bring about the skills or mastery needed to address them. There is evidence, for example, of critical success factors not being handled properly, leading to project failure (Young & Jordan, 2008). This evidence suggests that we do not yet understand well enough how we can lead a project to success by applying CSFs. This line of thought, in turn, raises questions, as to what constitutes a critical success factor, in terms of its component parts, and how the components should be successfully applied across the life of a project.

Top management support (TMS) is one important CSF. As revealed by many studies (Loonam & McDonagh, 2005; Nah et al., 2001; Krumwiede & Lavelle, 1998; Young & Jordan, 2008; Zwilkael, 2008a-b; Zwilkael et al., 2008) it is a main ingredient in the recipe for project success. The major goal of this study is to gain insight into what constitutes TMS from the perspective of project managers (PMs), by asking them what actions they consider are needed from top management (TM) to facilitate project success. This research theme has not been adequately addressed in prior studies.

The study described in this paper is the first stage of a two-stage research project. Here, we look at software development projects from the perspective of five project managers, using extensive interviews. The qualitative findings are built into a conceptual model. The study was conducted in Sri Lanka, the home country of the lead author.

The remainder of the paper is organized in the following way. First, the existing literature related to TMS is reviewed as theoretical background. The methodology is then presented, followed by the findings from the five case studies. The conceptual model and its future development is addressed in the discussion section, which is followed by the conclusion.

2 THEORETICAL BACKGROUND

2.1 Software development projects, project success and CSFs

Software development is pursued as projects (Schwalbe, 2006; Meredith & Mantel, 2006) and therefore draws on knowledge from both the software development discipline and project management. Project management refers to the application of knowledge and skills in the project environment to successfully complete project tasks (PMBOK, 2004). At present we have a very pressing problem in that most software projects are reported to have been unsuccessful in one aspect or another. Software projects are said to be high risk because they involve changing requirements, a variety of business domains, a variety of technical platforms and large amounts of monetary investments (Ropponen & Lyytinen, 2000; Cockburn, 2000; Schwalbe, 2006; Reel, 1999; Scott et al., 2006).

A project involves many stakeholders (Hartman & Ashrafl, 2002). Each one will have their own success criteria, and, therefore, project success is a multi faceted issue (Shenhar, Dvir & Levy, 1997;
Lim & Mohamed, 1999). It is interesting to note that some projects do not meet all predetermined criteria such as time, cost or scope, yet, are deemed to be successful as the client was happy with the project’s product. Previous studies describe project success as multi dimensional. Shenhar, Dvir, and Levy (1997) maintain that project success can be measured in terms of how well the following project outcomes are met: internal project efficiency, impact on the customer, business and direct success and preparing for the future. Others take a two dimensional, macro and micro, view of project success (Lim & Mohamed, 1999; Agarwal & Rathod, 2006). At the macro level the organization looks at project completion and customer satisfaction. At the micro level only project completion is deemed important.

CSFs are the factors identified as critical to the success of a project (Reel, 1999; Nah et al., 2001; Young & Jordan, 2008; Zwilkael, 2008; Hartman & Ashrafl, 2002; The Standish Group, 1995). Most prior studies portray the CSFs as derived from a practitioner’s perspective. However, there is evidence that the CSF concept is also a valid academic concept (Butler & Fitzgerald, 2006). However, very little work has been done to investigate the exact nature of each CSF. Some CSFs have been identified when projects were live (in progress), and others have been identified as important at the post-mortem phase of the project. Knowledge of CSFs is essential, as not attending to them when necessary may prove disastrous to the project. An example in this context would be if top management support was not rendered when needed. The lack of support may take the project down a perilous path, possibly ending in project failure (Young & Jordan, 2008; Schwalbe, 2006; Zwilkael, 2008a-b).

2.2 Top Management Support (TMS)

Agreement on a definition for TMS has yet to be achieved (McLagan, 1998; Loonam & McDonagh, 2005). Some authors define TMS as devoting time in proportion with the cost and potential benefits of a project (Young & Jordan, 2008). Others, however, define TMS as the degree to which top management understands the importance of the project function (Ragu-Nathan et al., 2004). Further, there is still a lack of consensus in the literature as to who comprises TM. Identification of TM may vary according to the organizational structure and with the size of the organization (Green, 1995; Sabherwal et al., 2006). Therefore the understanding of the term TM ranges from immediate superior to departmental manager, to director, CIO or even to the CEO. In the current study, we refer to the management one hierarchical level above the project managers as the TM.

Projects are managed by project managers and their definition in the management hierarchy is given as operational managers (Meredith & Mantel, 2006; Turner & Muller, 2005). They manage work on the project. However a project is part of an organization (Turner & Muller, 2003) and there is much interaction between the organization and the project. Project managers may require TMS for direction, advice or for escalation (Loonam & McDonagh, 2005; Nah et al., 2001; Krumwiede & Lavelle, 1998; Young & Jordan, 2008; Zwilkael, 2008a-b; Zwilkael et al., 2008) during the life of a project.

Prior studies indicate that project failure is strategic rather than technical (Cicmil & Hodgson, 2006). Many studies believe that top management support is essential and will most certainly increase the probability of software development project success (Schwalbe, 2006; Young & Jordan, 2008; Zwilkael, 2008a-b). Related findings from previous studies which concentrated on TMS include the critical success processes (CSPs) of TMS (Zwilkael, 2008a-b) and the introduction of a maturity model for TMS (Zwilkael et al., 2008). Although the CSPs are an interesting finding, the origins of these processes are not clear. For instance, some processes could be traced back to the top manager (e.g. project manager assignment), some to the project manager (e.g. use of new project tools and techniques) and others to the organization (e.g. project based organization). The maturity model for TMS is a subsequent development, depicting five stages of top management growth in an organization against the CSPs. This model would invariably inherit the attributes discussed above.

As interesting as the above findings are, there is a lack of critical analysis of actions required of top management, and how these actions could be understood at a theoretical level. Therefore, the authors
find it timely to undertake such a study from the perspective of project managers, who are of course ultimately responsible for project success.

3 METHODOLOGY

Five case studies were conducted to examine how TMS is viewed by project managers. Small, medium and large sized organisations were targeted in order to accommodate a range of views. The project managers were first asked to complete a short questionnaire about a project they could relate to including its level of success. The project success part of the questionnaire was based on the multidimensional project success model by Shenhar, Levy and Dvir (1997). The project managers took approximately ten to fifteen minutes to complete the questionnaire. The aim of the questionnaire was to draw attention to the project to be discussed. The projects we discussed had already been completed. Most of the project managers had multiple projects in their accounts. Therefore, we used this exercise to refresh their memory and assist in the subsequent discussion, as a road map teasing out specific practical examples, rather than discussing generally what was needed as TMS.

It was explained to the PMs that for the purpose of this endeavour ‘top management’ referred to their immediate supervisor(s) and the support needed from them. A semi-structured interview followed with emphasis on TMS requirements for project success. The notion of project success was also discussed, including what was meant by the term and other factors which were deemed to be important for project success. Supporting information regarding the project, from one hierarchical level above or below the project manager was sought for clarification of information and to rule out any bias in responses. In Cases 1, 3, 4 and 5 we had the opportunity to speak with a key team member, in Cases 2 and 3 we were able to undertake interviews with the departmental manager (the TM) in addition to the project manager. Ten such interviews ranging from 45 minutes to 1 hour 10 minutes were conducted. Table 1 provides details of the organizations and the projects studied.

Interviewees were encouraged to freely convey their perceptions. These discussions were recorded and later transcribed.

4 FINDINGS

Top management support was identified as important for project success. Three out of five project managers stated the fact explicitly. The two remaining project managers saw top management support as built into the organization’s business model. They said that the industry approved model adopted by the organization ensured TMS. So in the discussion they did not try to explicate the factor, but agreed that it was indeed very important. Table 2 presents a comparative analysis of key attributes of each case.

Interestingly, the understanding of project success differed according to each interviewee’s position in the organization. A key team member when asked why he thought the project was a success answered “we were able to give a product which satisfied the customer”, which relates to the micro level of success. The same question was answered by the project manager as “the customer is happy and we met time, cost and scope constraints”. The latter criteria relate to the macro level of success.

The project managers’ perceptions of what was required of TM to facilitate project success were analysed by the researchers and agreed upon after multiple passes through the transcripts. We also sought evidence that the understanding of these requirements was congruent across an organization, since we had access to either one hierarchical level above or below a project manager at each of the organizations studied. Sixteen requirements of TM were identified, and are described below.
1. **Participate in scope definitions**

   Project managers (Case 1) expected the top management from both the client and the performing organization to be involved in the definition of the scope of a project. PMs said that this prevents conflicts regarding requirements over the life of a project. The project manager of Case 3 described the top managers as gate keepers of scope who prevent scope creep.

2. **Build support in the organizational model**

   Having preferred or standard methods have helped project managers to successfully conduct project activities. As one project manager (Case 2) put it “when the customer realizes that we work with proven methods, they just fall in line”. One major aspect of these methods (Cases 2, 4 and 5) is to ensure client management participation alongside the client.

3. **Achieve a sustainable business model**

   One project manager (Case 5) who described his project’s product as not meeting customer satisfaction, pointed out that it was important that the top management looks at the sustainability of the business model employed, both in terms of revenue and workability. He pointed out that in the particular project under discussion three parallel versions were simultaneously developed and released to the customer and the customer was billed accordingly, providing good revenue. He went onto say said that, “because of this model the developers had to be constantly pulled out and plugged in where necessary making it difficult for them to concentrate, a sequential release mechanism would have been a more sustainable model, and the project would have had a better chance at success”.

4. **Provide guidance**

   Project managers, as middle level managers, are consistently pressured by operational constraints. They are well aware that they carry the responsibility of the project on their shoulders. So they expect and welcome guidance (Cases 1, 3 and 4). As one project manager put is “it is not just passive evaluation, but active participation. For instance they might say to fine tune the resource allocation”. More importantly one key team member pointed out that the TM should not let projects be orphans, but a part of the whole organization.

5. **Supply resources**

   Project managers saw it as very important that the TM supplied the required quantity and quality of skilled resources when necessary. One project manager (Case 1) said that it was helpful that the TM was able to get experts from different departments when they faced unforeseen technical issues. He said “all in all we were able to get help from others when we needed it”.

6. **Boost employee morale**

   TM attendance at team meetings, commending good work and offering opportunities to travel on project work was identified as having a positive effect on morale and project managers welcomed such support (Cases 2 and 3).

7. **Balance project assignments**

   Some project managers (Cases 2 and 5) said that they have multiple projects in their accounts and that this may sometimes get in the way of success of projects that are of lower priority.

   One project manager (Case 2) explained in relation to a project which did not meet the desired level of success, “I was involved in another major project and could not give this project the attention it needed”.

   Another project manager (Case 5) said that a normal working day for him lasted twelve to sixteen hours, and that he felt overloaded since he had many projects in his account. He said that “I would start with a stand up meeting and see what has to be done today, then I will attend to the mail which will take me up to lunch, by afternoon I would get feedback from the team about progress, I would
then attend to any communication needs and then update the tracking documentation. In the evening I would get an update when hiccups are shown, then I have to liaise and facilitate, for example hardware problems, HR or admin problems”.

8. **Prioritize**

Project managers (Cases 1 and 2) found that when a project is prioritized, it is much easier to receive required support from an organization. Soft and hard resources flow in and top management is available for any further requests and escalations.

9. **Watch status**

TMs are expected to remain vigilant in relation to the status of a project. This, according to Case 3, was expected from the top management of both the performing and client organizations.

10. **Having clear business objectives and stating them**

Project managers and in some cases key team members maintained that it was important for the TM to have an understanding of what the company is hoping to achieve from the project and to communicate the objectives to the project team (Cases 3 and 5).

11. **Make necessary information available**

According to one key team member (Case 4), it would have been easier to work with better knowledge/information than what was specified. He said “this would have prevented ambiguity of tasks and would have helped promote the success of the project”. This action and point 10 above are somewhat related. However they are elaborating on two different levels, i.e. the project level (action 10) and task level (action 11).

12. **Provide challenging work**

Project managers of Cases 2 and 3 brought out the fact that TM was expected to provide challenging work. This motivated staff and was also key to retaining skilled employees in the long term.

13. **Retention of key employees**

Project managers value and depend largely on capable skilled resources. Removing these skilled resources to other projects or not retaining them in the organisation, what ever the reason may be, is detrimental to the success of an ongoing project. As one project manager (Case 5) put it “I felt some attrition, for example when I wanted to retain some personnel and the management was not supportive”. He indicated that this was a cause for project failure.

14. **Review project plans**

Project managers expected TMs to review and formally accept project plans. Some project managers pointed out that this was beneficial in a number of ways, including securing TM buy-in and ownership at the top level for the project (Cases 1, 2 and 3).

It was also maintained that when revision of time or other constraint is needed, it is helpful to have top management involved in communicating the revisions to the client.

15. **Liaise with customer**

Customer perception of project involvement differs according to the level of management involved, in both the performing and the client organization. PMs reported that when action is needed and does not seem to happen, an escalation to the TM followed by a discussion between the peer levels of management of the performing and client organizations gets things moving (Cases 1 and 4).

In some cases top management was reported to have had close business relationships with the client prior to obtaining the project (Case 2), and this relationship had been beneficial in executing project tasks.

Proceedings ECIS 2009
16. Accept ownership and gain better understanding of project work

Case 2 brought out the fact that when TM from both the performing and the client organizations take ownership interest in the project, it helps project success. In some organizations the TM had a technical background and the project managers said that this was immensely helpful.

<table>
<thead>
<tr>
<th>Case 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>The organization is a campus, with own software development centre, using company standards. They develop software for internal and external customers. The development centre consists of 14 developers. Two developers were dedicated to the project and 1-2 quality assurance engineers were used as per the need. The application system was developed to assist tea auctioning. The project was given priority by the top management. The project was seen as a success because the customer was satisfied. The scope and budget were met. However, the time frame had to be revised but was accepted by the customer.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>The organization is a software development centre with many specialised areas branching out as departments. They develop software for external customers. The product strategy is to market a core product and bridge the gap between new customer requirements and the product. A project is employed for this purpose. The project was for the very first foreign customer and was prioritized. The organization has around one thousand (1000) developers employed. This project had 15 dedicated developers. Separate quality assurance was carried out with the involvement of 4 people. Implementation was done by 5 engineers. The application was an insurance system. The project was a success; the top manager describes meeting the customer requirements 100%. The customer has returned for new business and is now a reference site. The constraints for scope, cost and budget were met. The organization has level 4 certification in Software Engineering Institute’s (SEI) Capability Maturity Model Integrated (CMMI) and has also obtained certification by International standards organization (ISO 9001:2000).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>This Case is somewhat different from the others. We spoke to the software development department of a mobile service provider. This software department consists of 35 employees and is the customer. The project was conducted when the mobile service provider was switching mobile platforms, which is a rare occurrence. In order to provision the requested system to the new software environment the mobile service provider and the vendor had to work together on one project. This project is an example for involvement of multiple stakeholders. In its development phase alone four (4) parties were involved. Those being the supplier (4 developers), in-house developers (3-4 developers), telecom engineers and marketing and customer care. The project was declared a success by the customer. They said that the functional requirements and the time, cost and quality constraints were met.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>The organization is a highly reputed software development centre. They are CMMI level 4 certified and have a large employee pool, i.e. over 3500. The project concerned, handles requests at disaster situations and then handles bills and payments related to the actions taken on those requests. The product was for a foreign customer. This involved 12 developers and 3 quality assurance (QA) engineers. The project was declared a success, primarily because the customer was happy and the internal constraints such as time, cost and scope were met. The organization also has metrics with set indexes to monitor a project. The operations manager defined the project success as acceptance of the product by the customer, smooth rollout of the production system, high value on the client score card and repeated business with the customer.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case 5</th>
</tr>
</thead>
</table>
| This project is from the same organization as Case 4. The project concerned is a document retention management system. The project manager spoken to was a senior consultant for delivery, a higher level manager playing the role of project manager. The project involved 10 developers and 2-3 QA...
Engineers. The project was declared a failure from the customer satisfaction perspective. However, the performing organization’s requirements had been met.

Table 1 - A brief summary of Cases studied

5 Discussion: Conceptual model, future plans & limitations

We are able to group the TMS actions identified above into the three categories of strategy, facilitate and lead (see Table 3). Knowledge gathered from past studies was used to drive the categorization. The researchers sat through a brainstorming session to come to a consensus on this exercise. The term strategy has been defined and spoken of by researchers as early as the 70’s (Porter, 1979). However, at the core of the various definitions that have ensued since that time, is the fact that strategy constitutes the actions to ensure long term success of the organization. Leaders are expected to take ownership and pride in their work and set examples (Green, 1995; Viswesvaran et al., 1998). A facilitator assists an existing process, and prior studies state that this is what is expected of top management at times (Sabherwal, 2006; Kearns, 2007).

Using the three categories introduced above, referred to as TMS ‘roles’ from here on in, a conceptual model (Figure 1) is presented. Project success is introduced into the model as the outcome of the successful execution of the three TMS roles, which are considered to be the explanatory factors. We refer to the combination of a project’s process and a project’s product success as project success. Our research is congruent with findings of previous studies, regarding the fact that project success is multifaceted.

This model is an important milestone in the context of the overall research program. It is of interest in the future to see how these TMS actions are applied in the project management processes (initiation, planning, executing, monitoring / control, and closure) and how they impact project success.

The information we grouped and present is from the perspective of practitioners i.e., project managers. We believe that socio-organizational theories from the academic world can be used to better understand the three TMS roles of strategy, facilitate and lead and when they are required. Further, we believe that different theories come into play with each of these roles. At present it is considered that the following theories may be important for the reasons given. Organizational theory may be related to strategy, as it is based mainly on human, physical, work and coordination attributes (Hodge, 1988). The facilitate role involves creating the conditions for ordered rule and collective action (Stoker, 1998) and, therefore, may be related to project governance (Forcadell, 2007; Ezzamel & Reed, 2008). Human management theory may underpin the lead role (Leskiw & Singh, 2007) since people, however skilled they are, have to be led in situations such as project environments where people are brought together for a limited timeframe to achieve a specific goal or goals (PMBOK, 2004).

The research question to be addressed at the next stage of our project is as follows: “how do the top management roles of strategizing, facilitating and leading apply to the project management processes, and what impact do they have on project success?”

We plan to identify constructs which are related to the above roles from current theory, and use this information to operationalize the conceptual model. Then we plan to go back into the industry with a questionnaire compiled using the conceptual model and survey, firstly project managers and then top management in relation to the model. Information will be sought from the perspective of each project management process. Each role may be important in every one of the processes and one or more may be more important in different processes.

Our aim is to contribute a framework; where the three roles, strategy, facilitate and lead will be used as one dimension and the project management processes initiation, planning, executing, monitoring / control, and closure as the other.
<table>
<thead>
<tr>
<th>Attributes</th>
<th>Case 1 (organization 1)</th>
<th>Case 2 (organization 2)</th>
<th>Case 3 (organization 3)</th>
<th>Case 4 (organization 4)</th>
<th>Case 5 (organization 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organization</strong></td>
<td>Campus with separate SWD centre</td>
<td>SWD centre with many specialised departments</td>
<td>Mobile service provider</td>
<td>Software development centre</td>
<td>Software development centre</td>
</tr>
<tr>
<td><strong>Customer</strong></td>
<td>External (local)</td>
<td>External (foreign), first international customer</td>
<td>Self</td>
<td>External (foreign)</td>
<td>External (foreign)</td>
</tr>
<tr>
<td><strong>Total no of developers</strong></td>
<td>14</td>
<td>Around 1000</td>
<td>35</td>
<td>Over 3500</td>
<td>Over 3500</td>
</tr>
<tr>
<td><strong>Developers in project</strong></td>
<td>2</td>
<td>15</td>
<td>Internal 3-4, vendor 3-4</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td><strong>Application System</strong></td>
<td>Tea auctioning system</td>
<td>Insurance system</td>
<td>Provisioning system for changing mobile platform</td>
<td>Disaster recovery handling system</td>
<td>Document Retention System</td>
</tr>
<tr>
<td><strong>Successful? Why</strong></td>
<td>Yes; Customer is satisfied</td>
<td>Yes: Customer is satisfied, more business given</td>
<td>Yes: customer satisfied, constraints were met. (discussion with customer)</td>
<td>Yes: Customer is satisfied</td>
<td>No: Customer is NOT satisfied and system is not in production.</td>
</tr>
<tr>
<td><strong>Prioritized?</strong></td>
<td>Yes</td>
<td>Yes: first international project</td>
<td>Yes: but not in isolation. Process was prioritized</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Separate QA team</strong></td>
<td>Yes: (1-2 when necessary)</td>
<td>Yes : 4</td>
<td>Thorough testing, both vendor and customer</td>
<td>Yes: 3</td>
<td>Yes: (2-3 as required)</td>
</tr>
<tr>
<td><strong>Experienced PM</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>No, new recruit</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Model/ Method</strong></td>
<td>Company standards</td>
<td>CMMI Level 4; ISO certified</td>
<td>Company standards</td>
<td>CMMI Level 4</td>
<td>CMMI Level 4</td>
</tr>
<tr>
<td><strong>TM Actions</strong></td>
<td>1,4,5,8,14,15</td>
<td>2,6,7,8,12,14,15</td>
<td>1,4,6,9,10,12,14</td>
<td>2,4,11,15</td>
<td>2,3,7,10,13</td>
</tr>
</tbody>
</table>

Table 2 – Comparative view of the Case attributes
It is shown in the literature (Correll, 1994) that top managers are not particularly available for operational management support on request, since they have busy schedules themselves. So it is of importance that both top managers and project managers realize that support will be needed during the life of the project. The framework that we aim to define will assist them in advance to identify and understand, when and what kind of action may be required. A theoretical insight into each element will further elaborate and justify the requirement for top management support. Therefore, this intended framework could be used as a tool to ensure top management support.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Facilitate</th>
<th>Lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workout a sustainable business model</td>
<td>Supply resources</td>
<td>Accept ownership and gain better understanding of project work</td>
</tr>
<tr>
<td>Have clear business objectives and state them</td>
<td>Make necessary information available</td>
<td>Review project plans</td>
</tr>
<tr>
<td>Provide challenging work</td>
<td>Retain key employees</td>
<td>Provide guidance</td>
</tr>
<tr>
<td>Balance project assignments</td>
<td>Liaise with customer</td>
<td>Watch status</td>
</tr>
<tr>
<td>Build support in the organizational model</td>
<td>Boost employee morale</td>
<td>---</td>
</tr>
<tr>
<td>Prioritize</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Participate in scope definitions</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

*Table 3- Analysis of the top management actions*

*Figure 1 - Conceptual model*

We see the following as limitations in this study. We were interested in researching TMS actions required for software development projects. Therefore, the majority of the organizations selected were from the software industry and as such our investigation did not investigate the relationship between TMS and the core business of the organizations. Although references to other projects were brought into the discussion, the factors needed as TMS are mostly limited to the projects that were studied in the five organizations involved. The investigation took place in one country only and, therefore, the findings may not be a generic representation. However, our plan is to extend the research to other Asia-Pacific countries, in the second stage. We have also not investigated in this stage, how the level of economic development and cultural aspects of a nation may affect TMS.
6 CONCLUSION

Many academics and practitioners believe that top management support is an important factor for the success of software development projects. This study examined five software development projects and is the first stage of an ongoing research program. We identified sixteen (16) specific actions needed by project managers as support from their immediate management. These actions were categorized into three groups (strategy, facilitate and lead) with the aid of prior literature and are identified as three important top management roles. The findings led to the development of a conceptual model with the three roles (strategy, facilitate and lead) identified as explanatory factors, and project success as the outcome.

The unique contribution from this paper is the conceptual model (Figure 1) based on the findings to date. There is promise in this framework, which not only looks at top management support roles, but can be built upon to examine the TM roles over the life of a project. The expanded framework planned for the second stage of this research program will have the capacity to explain and justify the top management actions required in each project management process using socio-organizational theories.

References

Kearns, G. (2007), How the internal environment impacts information systems project success: an investigation of exploitative and explorative firms, Journal Of Computer Information Systems, Fall 2007, 48(1), 63-75
Krumwiede, D., Sheu, C., Lavelle, J. (1998), Understanding the relationship of top management personality to TQM implementation, Production And Inventory Management Journal, second quarter 1988, 39(2), 6-10


From Business Case to Value Case - Assessing the Organizational Value of IT Investments

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0469.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Evaluation, IT investment, Management control, interpretivist research</td>
</tr>
</tbody>
</table>
From Business Case to Value Case - Assessing the Organizational Value of IT Investments

Frisk, Elisabeth Jane Applied Information Technology, Gothenburg University, Sweden. elisabeth.frisk@ituniv.se

Abstract

Managers continually invest in new information technology (IT) but the question of organizational value still seems vague. One explanation is poor evaluation. In practice the Business Case including Return on Investment (ROI) still dominate. Information System research has noted for a long time that the Economic Approach is not sufficient and instead the Interpretative IT Evaluation Approach has been put forward. However, the approach has reached limited acceptance in practice and it has been noted that what to evaluate is a far more complex process than might first appear. The aim of this study is to articulate factors and criteria that are important to consider when assessing the organizational value of IT investments. This study is part of a Collaborative Practice Research project that took place 2005-2008 at three public organizations. The findings indicate that it is time to take a step from a Business Case to a Value Case. The Value Case is a pluralistic, a formative and a formalized approach that includes factors and criteria that have its base in prior research and have been further discussed and analyzed by the respondents. The Value Case also put management’s attention to effectiveness and efficiency, the task of management.

Keywords: IT assessment, IT investment, Management, Interpretative IT Evaluation
1. INTRODUCTION

IT investments are still high-risk projects. The newspapers repeatedly describe IT project failures costing the organizations millions of dollars. In Sweden, one organization in the health sector had to cancel one large IT project after the cost had exceeded the budget enormously (Järräng 2007). Also the regional social insurance office had giant runaway costs when implementing SAP (Järräng 2009). There have been similar findings in the UK public sector and the failures there have been related partly to insufficient management skills (Brown 2001). One explanation is the overreliance on financial business cases (Ward & Daniel 2006). Therefore, new knowledge is needed that can create increased understanding for other effects than the economic when introducing a new IT investment to the organization. Today the Business Case mostly including a Return on Investments is still the dominating approach when assessing the organizational value of an IT investment (Ward & Daniel 2006). Several researchers have criticized the economic approach due to that IT’s role has changed from automated processes to increased individual and group effectiveness, to organizational transformation and to collaborative partnership (Pearlson 2001). Instead interpretative IT evaluation, based on stakeholder groups’ perception of reality and a consideration of the context (why), content (what) and a process (how), has been put forward (Hamilton & Chervany 1981, Symons 1991, Walsham 1995, Jones & Huges 2001, Ward and Daniel 2006, Stockdale & Standing 2006). However, interpretative IT evaluation is rarely used and the reasons can be several. Stockdale and Standing (2006) note that; “A decision on what is to be evaluated is a far more complex process than might first appear and is significantly influenced by the stakeholders and by the context of the organization” (p. 1092). Due to lack of management skills and a complexity of what to evaluate it can be of interest to create increased understanding for what management perceives as useful to reflect upon, when assessing organizational value of IT investments. In the research field IT evaluation, methods and models based on economic and interpretative theory are the most discussed (Berghout and Remenyi 2005). The question is if these theories are sufficient or a step forward is needed. The aim of this study is to articulate factors and criteria that are important to consider when assessing the organizational value of IT investments. The goal is from a management perspective, to create increased understanding for different factors affecting organizational value of IT investments. The goal to IS research to develop an analyze tool that can support the content of IT evaluation in the interpretative IT evaluation approach. The following research question is raised: When assessing the organizational value of an IT investment, what factors and criteria should be reflected upon from a management perspective?

This study has a Collaborative Practice Research approach. It is based on managers’ perception about what should be reflected upon when evaluating organizational value of IT investments. The managers involved have experience from decision support systems aiming to save time in the operative work, give information about how to handle chemical accidents and provide better analysis of the operative work etc. The main purpose of these IT systems is foremost to provide the organization with better information related to the operative work. Further, IT systems are considered as social systems and “…has some recognized functionality but needs to be considered as a set of social objects. As a social object, its influence on organizational function and performance cannot be separated from expertise, jobs, processes or structures.” (Zammuto et al. 2007 p. 753).

This study involves three non-public organizations in Sweden. The organizations have a rather unique situation since the grant of money has its origin in a cost budget process that gives the frame for the approved costs. If one organization spends less money one year, they risk receiving fewer grants next year. In addition, the evaluation of IT investments in the public sector can be extra problematic since they work in collaboration with many different organizations, such as the police, ambulance, and health services.

In the next section, I will give a literature review. Then the organizational settings and research approach will be presented. Followed by the result of this study, a discussion, and a short conclusion.
2. ASSESSING THE VALUE OF IT INVESTMENTS

The problem of evaluating IT investments is framed as the productivity paradox. It originates from studies during the 1980s that found no connection between IT investments and productivity in the U.S economy. Productivity is a measure of efficiency, the use of resources (output, doing things right). The productivity paradox was originally stated by Solow and further developed and discussed by Brynjolfsson (Dedrick et al. 2003). Today there is evidence that IT provides positive impact on productivity and instead the attention has switched to IT and profitability (ibid). Profitability is described as the fulfillment of goals and measurement of effectiveness (outcome, doing the right things) (Lewis et al. 2007). The task of management is according to Lewis et al. (2007) to “administrating and coordinating resources effectively and efficiently in an effort to achieve the goals of the organization”. This is a rather rational view of the organization. However, both effectiveness (do the right things) and efficiency (do things right) are relevant measures to consider since the role of the involved organizations is described by Fire Rescue Agency as to deliver services (pre-determined goals) to citizens in an efficient way. Also, non-profit public organizations in Sweden have no traditional income and balance sheet that can evaluate the outcome of the organizations. It should be important to initially reflect upon the effectiveness (goals achievement, doing the right things) and the efficiency (use of resources, doing things right) before an IT investment is accepted. However, effectiveness and efficiency can include other factors than the economic since value is pluralistic (Guba & Lincoln 1990, Bannister 2001). This will be discussed in the next paragraph.

2.1 The Economic and the Interpretative Approach

The two most discussed IT evaluation approaches within the research field IT evaluation are the Economic and the Interpretative IT Evaluation Approach based on economic and interpretative theory (Berghout & Remenyi 2005). The Economic Approach is the most used method by management when assessing IT investments (Ward & Daniel 2006). There exists a plethora of different economic methods such as the capital budgeting methods: Pay Back, Internal Rate of Return (IRR) and Net Present Value (NPV). These methods is described as “the process of analyzing potential capital expenditures and deciding which investments the firm should undertake”. (Brigham & Gapenst 1996). To use these methods requires a consideration of aspects such as estimating expected cash in- and outflows and calculating the sum of the present values of the expected cash flows. Different kinds of information is attained such as, how long it will take for the investment to return invested capital, and calculating the interest costs for not borrowed capital in order to give a more accurate result of the investment etc. The methods are summative and do not give guidelines how to follow-up an investments along its life cycle. Thus, evaluate organizational value of IT investments with economic methods only gives information about efficiency and effectiveness from an economic view.

The Interpretative IT Evaluation Approach sees IT systems as both technical and social entities (Walsham 1995) and as the starting point for evaluating organizational value is the stakeholders’ perception of reality (Guba & Lincoln 1990, Symons 1991, Stockdale & Standing 2006). Examples of interpretative IT evaluation models are the CCP (content, context, process) framework and Benefit management (BM). The CCP framework focuses on the context (why), the content (what) and the process (how) (Symons 1991, Stockdale & Standing 2006). These concepts derive from Pettigrew’s conceptualization of organizational change (Symons 1991). Stockdale and Standing (2006) have described the context as explaining why the IT system is of importance. The content is based on the stakeholders’ perception of the value so their choice of criteria determines the content (ibid). The process is formative and follow up the IT investment along its life cycle. BM focuses on how to follow-up benefits in different steps along the IT investments life cycle. Thus, the Interpretative IT evaluation approach is contextual, the content is dependent on stakeholders’ choice of criteria, and the process is formative. Next it would be of interest to find out what prior research has been put forward as important factors and criteria to be reflected upon when assessing IT investments.
2.2 Beyond the Economic and the Interpretative IT Evaluation Approaches

The research field IT evaluation can in overall be described as fragmentary (Berghout & Remenyi 2005). IT evaluation includes a plethora of different factors and criteria directed to IT evaluation of IT investments. A well-known multi-criteria method is Information Economics developed by Parker and Benson (Robson 1997). Information Economics focuses on the business and the technology domain and include factors such as strategic match, competitive advantage, competitive response and organizational or project risk. Other factors suggested to reflect upon is strategic match. Strategic match is important since it creates awareness of whether the IT investment matches the strategic context (Kefi 2002). In addition, Ballantine & Stray (1999) argue for strategic alignment i.e. the IT investment should align with the IT strategy that in turn should be aligned with the business strategy. If strategic match not is done the IT, investment could be waste of money in the end. In addition, the effect of IT investments inside and outside the organization is consider as valuable input to the discussion of the organizational value. For instance, when considering collaborative and inter-organizational systems the surroundings of the organization should be analyzed (McCalla et al. 2003). Stefanoue (2001) writes that, understanding of organizational impact and change is required if any benefits are to be realized, since poor fitting of the system to the organization relates to an inability to respond to changes. In addition, costs of organizational change are required if any benefits are to be realized (ibid). Benefits are described as hard and soft, costs as direct and indirect (Love et al. 2005). Indirect benefits can occur due to further IT investments. Indirect costs are difficult to see and examples of such costs are user resistance; personal training; external consultants; additional applications and system downtime (Stefanou 2001). In addition, optional theory and uncertainties are mentioned as input to IT evaluation (Toffolon 2001). Stakeholders’ involvement gives a better understanding of benefits, value and suitability than traditional economic methods (Irani & Love 2001) and the functionality of the system (Khalifa et al. 2001). However, when it comes to stakeholders it is mainly the stakeholders of the system that are discussed and not the stakeholders of the organization. Still several IS researchers see the IT system primarily as a technical system and therefore the most important factor to consider (Gemmell and Pagano 2003, Choenni et al. 2003). Further, Clay et al. (2003) write that a project manager who has authority and experience will influence the outcome of an IT project in a positive way. Thus, prior research has provided knowledge about many factors and criteria that should be reflected upon when assessing IT investments.

3 RESEARCH APPROACH

3.1 Research site: The Fire Rescue Services

The Fire Rescue Service (FRS) is responsible for providing the population with services such as prevention, preparation, and response. The FRS is structured as a Fire Rescue Service (FRS) or as a Fire Rescue Alliance (FRA). The difference is that the former co-operates with the municipalities where it is situated. The latter consists of several FRS and acts as one municipality that assists the municipalities involved. The directorate of the confederation gives the economic pre-condition for FRS and FRA. The confederation is composed of politicians from one or several municipalities. The three involved organizations differ in structure and amount of employees. Organization A was a FRA, acted in a large city, included five municipalities, has 1000 employees, and 9 Fire Stations. Organization B was a FRS, acted in a large city, has 650 employees, and 9 Fire Stations. Organization C was a FRS, acted in a middle-sized municipality, has 100 employees, and 3 Fire Stations.

IT investments were initiated inside or outside the organizations. Initiatives outside could be initiated due to a new law or from the Rescue Services Agency, the government authority. Initiatives inside the organizations could come from anyone within the organization. In addition, initiatives come yearly when the budget process is running, all kind of requirements pop up, running track, IT system etc. Decisions on IT investments were decentralized, and were often taken on the departmental level. However, if the costs exceed a specific amount the Chief Manager of the FRS became responsible. Before such a decision, the board often discussed the IT investment and if it should be approved. If the
IT investment costs exceed regular budget restrictions, it becomes a political issue and the decision has to be supported by the local government committee, appointed by the local municipality. The individual decisions on IT investment were based on factors such as benefits, costs and technology. The benefits are often argued from the requester view and not from an organizational view. The cost calculations included mainly costs for IT hardware and software since the costs for using the organizations own resources was consider already taken. Thus, the focus is on the IT system and its possibilities and not on the need of the organization. Historically the organizations have had good economy and that has created a culture where the organizational value is not always in focus. Times have changed and in one organization, large savings are needed. So, evaluating organizational value of IT investments were perceived to be more important in the future.

3.2 Collaborative Practice Research

The research project was performed from 2005 to June 2008. It was a well-known problem that initiatives on IT investments are based on personal interest and have a technique focus. Instead, the perception among several of the respondents was that it was time to put the need of the organization in focus. After the initial interviews were the problem area was appreciated it was evident that IT evaluation was a management, IT evaluation and a conceptual problem.

The next step was to investigate the content, what should be evaluated, and how should the process be accomplished. The research approach was Collaborative Practice Research (CPR). CPR tries to fulfill the dual objectives, improving practice and contributes to the body of knowledge within the research (Mathiassen et al. 2002). CPR is pluralistic since it is difficult to control the research with only using action research (Mathiassen 2002). In addition, the researcher is dependent on how practice evolves and it is not easy to control the focus of the research outcome. Therefore, diverge activities support each other and give deeper understanding for the investigated phenomenon. The CPR-based process is developed by Mathiassen (2002) and is influenced by Susman and Evered, Checkland and McKay, and Marshall’s action research approach (ibid). The project has been done in collaboration between the researcher and management in the organizations involved. The CPR includes different stages, see table 1, and the stages of the research project are presented in the second column. This study focuses on point C in the second column.

<table>
<thead>
<tr>
<th>Collaborative Practice Research</th>
<th>The research project at Fire Rescue Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Appreciate the problem situation</td>
<td>A. Appreciate the problem situation (X 2006).</td>
</tr>
<tr>
<td>2. Study literature</td>
<td>B. Study literature, and select a theoretical framework (X 2007)</td>
</tr>
<tr>
<td>3. Select approach</td>
<td></td>
</tr>
<tr>
<td>4. Develop framework</td>
<td>C. Design a new evaluation framework (This study).</td>
</tr>
<tr>
<td>5. Design process</td>
<td>D. Design evaluation process</td>
</tr>
<tr>
<td>6. Apply approach</td>
<td>E. Apply approach</td>
</tr>
<tr>
<td>7. Evaluate experience</td>
<td>F. Evaluate experience</td>
</tr>
<tr>
<td>8. Exit</td>
<td>G. Exit</td>
</tr>
<tr>
<td>9. Assess usefulness</td>
<td>H. Assess usefulness</td>
</tr>
<tr>
<td>10. Elicit research result</td>
<td>I. Elicit research result</td>
</tr>
</tbody>
</table>

Table 1. This research project, in comparison to Mathiassen et al. (2002) research project.

Thus, the findings are based on the respondents’ perceptions of reality and the interpretation of the researcher. My role was an “outside researcher” and during the workshops and interviews, I did try not to influence the participants in any direction. In the workshops, findings from prior IT evaluation research were discussed. I also took part of several meetings and documents such as a project model, an activity plan, introducing a new IT project, information concerning Balance Score Card etc.

3.2.1 Data collection and Data analysis

In organization A and B the IT manager selected the respondents that should attend the workshops. In organization C it was the manager for the operative work. These managers had awareness of who had
the experience from IT investments or assessing IT investments. The study involved at each organization six to seven managers from different levels. Managers attending the study were at top level: Chief Manager, Vice Chief Manager and IT Manager. At functional level, Managers of department were attending. At the operative level, Managers of the operative work force were involved. By management is meant someone who has an overall responsibility for the organization, the department, or for a team alternative the turnout.

The first workshop included two different meeting points. First, the findings from the interviews (see table 1, point A) were presented. Theses findings indicated that the missing value of IT investments was caused by lack of management control (lack of strategy and co-ordination), lack of understanding of the concept value (each stakeholders has his own perception), and lack of an appropriate formalized IT evaluation approach that consider context, content and a process. The findings were recognizable by the participants. Next, the IT evaluation, the economic and interpretative IT evaluation approach, and factors and criteria put forward by prior research were introduced. The presented factors and criteria were received in a positive way and were in accordance with management perceptions of what should be reflected upon when assessing the organizational value of IT investments. Along the discussion, I wrote down the factors and criteria that by the participants should to be part of an IT evaluation. It was also an agreement among the participants about the selected factors and criteria. I perceived that this exercise created learning to the participants and put structure on a complex issue. I finished the workshop by asking if they wanted to structure the selected factors and criteria in a word document, a power point presentation or in some other form. Two of the organizations wanted to structure the findings in a word document and the third organization wanted a power point presentation. The aim of the documents was to support the evaluation process of IT investments. In the second workshop, the respondents discussed the results from the first workshop. The workshop resulted to further development and modification of factors and criteria selected by the participants. I perceive that the modification occur due to that the respondents have had time to reflect upon the discussed factors and criteria and had developed an improved understanding for organizational value. Thus, the workshops trigger their own learning by reflecting upon factors and criteria related to organizational value. The last workshop was performed according to the second.

Each workshop and lasted approximately 2 hours. During the workshops, I took notes and after the workshop, I wrote down spontaneous reflections from the meetings. For the analysis, I used the notes from the paperboard and from my own reflections. For two of the organizations I summarized the chosen factors and criteria in a word document, called the Value Case. This Value Case was further modified during the workshops. For the third organization, the result was summarized in a power point presentation, named the Value Case. I also compared the results from all three organizations in order to analyze patterns of similarities and deviations. This analysis resulted in a Meta Value Case that included all factors and criteria suggested from the three organizations. Organization A and B wanted to learn from each other and took part of the Meta Value Case.

Care was taken to ensure that the findings were interpreted in accordance with the respondent’s perception regarding what factors and criteria that should be included in the value case. In order to validate the findings, the Meta Value Case were also presented for top managers at two others Fire Rescue Services not involved in the study. The results were favorably received in both organizations. The managers perceive that the document could support the IT evaluation process and increase the understanding for the organizational value of an IT investment.

4 ASSESSMENT FACTORS GENERATED BY PRACTICE

Management has in literature been described as “... the process of administrating and coordinating resources effectively and efficiently in an effort to achieve the goals of an organization” (Lewis 2007, p. 5). Such management approach can be extra important in non-profit public organizations that are goal driven.
In order to increase the understanding of the relevance of these factors they have been categorized into effectiveness (do the right things) and efficiency (do things right). The respondents suggested that the selected factors and criteria should be evaluated in an iterative process along the IT investments, life cycle and give continuously feedback about the relevance of the IT system.

4.1 The Rationale for Effectiveness, doing the right things

The respondents considered strategic match vital since the lack of strategic match has contributed to an ad-hoc development of IT. Strategic match can give management increased understanding of the relevance of new IT investments (doing the right thing). The respondents (functional and operational level) perceive that IT initiatives today are often based upon individual interests and not on the needs of the organization. “We don’t know the plans for the coming two years. This allows persons who are most anxious about new information systems or information technology to get their requests approved. The arguments are often based on personal interest and not on the needs of the organization” (Operational managers). “Unfortunately we give priority to individual desires instead of the total picture. We can’t agree on one brand of digital cameras or digital calendars” (IT manager). In addition, to include strategic match as a factor can give management increased understanding of the need of explicit strategies for the organization in order to be able to govern the public sector towards its mission. By the time of the workshops none of the organizations had evident strategies for the business, organization, or for IS and IT. In one of the organizations the IT Manager was waiting for the organizational strategy while the Vice Chief Manager was waiting for the IS and IT strategy. In two of the organizations, the Balance Scorecard was on the agenda. I perceive that two of the organizations tried to agree on what perspectives to include in order being able to doing benchmarking, but this seems like a problematic process. To introduce a Balance Score Card could facilitate the strategic match of an IT investment. In one organization, the IT manager refused to take responsibility for the operations of new IT investments since IT was introduced into the organization without discussing or informing the IT department. Not coordinating IT will give increased cost for running IT and the burden of work becomes unmanageable for the IT department. In addition, the managers perceive that IT is power and IT decisions are a political act, and a focus on strategic match can reduce power situations since the need and the value of the organization will be in focus.

Today FRS collaborates with organizations such as the police, the ambulance and the municipality so impact on the surroundings is an essential factor to reflect upon. Also, the impact on the citizens are vital since FRS/FRA work for the citizens’ best, but is often forgotten in the discussions about value (functional and operational level). Criteria such as politics, economy, society, and technology i.e. what is going on outside the organization should also be reflected upon in order to ensure that the organization would do the right things. The selected factors and criteria in the workshops concerning effectiveness are presented in table 2.

<table>
<thead>
<tr>
<th>Effectiveness</th>
<th>Organization A</th>
<th>Organization B</th>
<th>Organization C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategic match to:</td>
<td>The Business, Organization and IS and IT. Support the Balance Scorecard. Safety and Health Act.</td>
<td>The Business, Organization and IS and IT</td>
<td>The Business, Organization and IS and IT If not, motivate WHY?</td>
</tr>
<tr>
<td>Impact on the surroundings:</td>
<td>Politics, Economy, Society, Technology (PEST), Citizens, Collaborative org. or other actors. Dependencies to other regulations or other projects.</td>
<td>Collaborative organizations or other actors.</td>
<td>Other Actors.</td>
</tr>
</tbody>
</table>

Table 2. Factors and criteria related to effectiveness.

4.2 The Rationale for Efficiency, doing things right.

Organizational impact is also essential since it opens up the understanding for the changes needed in the organization, and from that, quantitative and qualitative benefits could more easily be deduced. Also, a criterion like power is of interest since IT changes power structures. The respondents regarded
benefits to be difficult to make explicit and benefits should not only be an enumeration of quantitative and qualitative aspects but also be described with who receives the benefits. Managers perceive that the technology get to much attention: “The internal discussion often concerns technology, technical platform and the IT system, but questions should be raised such as, what needs to be developed, what do we want to achieve, and how should we proceed?” (A manager at functional level). Today cost calculations include costs of the hardware and software but not for the employee’s involvement, education, maintenance, support, for the project etc. According to some of the managers this can explain why the decisions-makers seem to have a habit of buying new IT instead of upgrading (IT manager). Factors as risks are necessary and in particular risk for the project, the organization and the technology. Also not involving the stakeholders can make it worse instead of improving. The IT system and criteria such as security, architecture, support, flexibility etc. are important. In addition, the respondents added project organizing in order to increase the IT project chance of succeeding, since the IT project is dependent on right timing and the ability to involve the right persons. Selected factors and criteria related to efficiency are summarized in table 3.

<table>
<thead>
<tr>
<th>Efficiency</th>
<th>Organization A</th>
<th>Organization B</th>
<th>Organization C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describe costs for the:</td>
<td>Project, Purchase, Implementing, Education, Required changes, Running costs and Negative effects.</td>
<td>Project, Purchase, Implementing, Education, Required changes, Running costs and Negative effects.</td>
<td>Purchase, Running costs, Education and Licenses.</td>
</tr>
<tr>
<td>Reflect over following risks:</td>
<td>The decision process not deep enough. Key persons in the project? Too much technique focus? New directive? New policy from the municipality? Software, Supplier, Security, Costs or other risk.</td>
<td>Dependencies to other projects, financiers, and suppliers. Technology. Operating the project. The organization. Environmental factors. Competence.</td>
<td>No risks were put forward by the organization.</td>
</tr>
<tr>
<td>Stakeholders opinion:</td>
<td>What are the opinions among the stakeholders? Is the pre-knowledge ok for this system? Will the system affect the use of resources?</td>
<td>What are the opinions among the stakeholders? Is the pre-knowledge ok for this system?</td>
<td>What are the opinions among the stakeholders?</td>
</tr>
<tr>
<td>Project organizing:</td>
<td>Describe the operating and the administration of the project.</td>
<td>Describe the operating of the project and resources needed?</td>
<td>Describe timetable and responsibilities.</td>
</tr>
</tbody>
</table>

Table 3. Factors and criteria related to efficiency.

5 DISCUSSION: TOWARDS A VALUE CASE

Lewis et al. (2007) have described the role of management as “the process of administering and coordinating resources effectively and efficiently in an effort to achieve organizational goals” (p.5). This is very much in accordance with the opinions of several of the managers, a demand for an IT
evaluation approach that could shift attention from a technique and personal interest focus to organizational value based on effectiveness (do the right things) and efficiency (do things right).

One way to support managers’ consideration of effectiveness and efficiency in the IT evaluation process of IT investments is to introduce a Value Case. The Value Case takes one-step further from the Business Case and takes a formalized, pluralistic, contextual and formative view. The Value Case includes factors and criteria put forward by prior research and is further discussed by the involved managers.

5.1 Reflections on the Economic and Interpretative IT Evaluation Approach

The respondents argue as well as several researchers, Symmons (1991), Stockdale and Standing (2006), and Ward and Daniel (2006), that an economic perspective is not sufficient. However, the Interpretative IT Evaluation Approach was not sufficient either since the content needed support by pre-determined criteria that could increase the understanding of factors and criteria affecting effectiveness and efficiency i.e. the organizational value. Stockdale and Standing (2006) has also noted that what to evaluate is a far more complex process than first appear. It can be worth reflecting upon the consequences of using an Interpretative approach if the involved stakeholders lack experience and pre-knowledge of a new IT investment. Also, if pre-knowledge and experience of the IT system differ among the different stakeholder groups, how will that affect the discussions about organizational value. Since evaluating IT investments is described as a political process there is also a risk that the discussions of the organizational value of IT investments deal with individual interests from some stakeholders and other stakeholders lack arguments due to limited experience. A Value Case including a formalized, pluralistic, contextual and formative approach can support managers in a discussion of the organizational value and the appropriateness of the IT investment.

5.2 Towards a formalized, formative, contextual and a pluralistic Value Case

The involved managers argue that the understanding of organizational value can be improved by assessing factors and criteria related to both efficiency and effectiveness. The factors and criteria generated by the managers are presented and discussed in next section.

Effectiveness

- **Strategic match.** In accordance to Kefi (2002) the respondents said that before starting an IT project it is vital to match an IT-investment with the strategic context since it contributes to an increased awareness if the IT investment is “doing the right thing” for the organization. The respondents added criteria related to internal goals included in the Safety and Health Act.

- **Impact on the surroundings.** This factor was also considered important since the organizations cooperate with several other organizations such as the police, ambulance etc. McCalla et al. (2003) note that this can be particularly important when collaborative and inter-organizational systems are evaluated. The respondents also consider the traditional strategic criteria, politics, economy, society and technique relevant in order to create awareness of doing the right things.

Efficiency

- **Impact on the organization.** The respondents agree with Stefanoue (2001) that organizational impact must be considered if any benefits are to be realized. However, the respondents also emphasize the importance of analyzing the impact of an IT investment on the organization since that will facilitate the understanding of needed changes and in turn to whom the benefits relate.

- **Benefits.** In literature, several discussions concerned tangible, intangible, direct and indirect benefits (Love et al. 2005). The respondents were more interested in discussing quantitative and qualitative benefits and to whom the benefit was directed. The respondents said that it is important to consider the stakeholders of the organization in order to get a rich picture of who benefits from the new IT system. Due to individual interest of IT, it can be a risk of internal focus and the citizen perspective cast aside.
- **Costs.** Costs were of great interest since a better calculation could contribute to better understanding of total investment costs, running costs and for the selection between buying new or upgrading. The literature that discussed direct and indirect costs (Love et al. 2005) received less attention.
- **Risks.** The respondents discussed risks related to the organization, the IT project, the IT system, costs etc. Uncertainties and optional theory did not catch the respondents for further discussions.
- **The IT system.** Gemmell and Pagano (2003), and Choenni et al. (2003) have suggested some criteria, and those were received with recognition. Also, questions like, what will happen if the IT system breaks down and stop were discussed as important input to the value case.
- **Stakeholders’ opinion.** Irani and Love (2001) noted that stakeholders’ involvement incorporates increased understanding of benefits, value and suitability. Khalifa et al. (2001) write that the functionality of the IT system will also be improved. These statements are very much in accordance with the opinions of the respondents. They argue that stakeholder groups’ affected by the new IT system should be involved and interviewed so they can give their view of the new IT system.
- **Project organizing.** This perspective was not so well discussed in the literature; Clay et al. (2003) note the importance of a project champion. The respondents said that it is central to involve the right people and have the right timing in the IT project, also to clarify dependencies on other IT projects.

Factors and criteria discussed by the respondents as valuable input to the evaluation process when assessing the organizational value of IT investments are presented in table 4.

<table>
<thead>
<tr>
<th>The Value Case</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effectiveness:</strong></td>
</tr>
<tr>
<td><strong>Strategic match:</strong></td>
</tr>
<tr>
<td><strong>Impact on the surroundings:</strong></td>
</tr>
<tr>
<td><strong>Efficiency:</strong></td>
</tr>
<tr>
<td><strong>Impact on the organization:</strong></td>
</tr>
<tr>
<td><strong>Benefits:</strong></td>
</tr>
<tr>
<td><strong>Costs:</strong></td>
</tr>
<tr>
<td><strong>Risks:</strong></td>
</tr>
<tr>
<td><strong>The IT system:</strong></td>
</tr>
<tr>
<td><strong>Stakeholders’ opinion:</strong></td>
</tr>
<tr>
<td><strong>Project organizing:</strong></td>
</tr>
</tbody>
</table>

*Table 4. The Value Case*

The empirical findings of this study indicate that only using an interpretative IT evaluation approach don’t match the need of management. A step forward is needed since methods or models based on economic or interpretative theory are not sufficient from a management view when evaluating IT investments. That economic methods are not sufficient has been known for a long time. Instead the interpretative IT evaluation approach has by prior research been suggested. However, the content, what to evaluate, in the interpretative IT evaluation approach is based on stakeholders’ perception on reality. But if the managers lack experience or pre-knowledge of a new IT system, such approach can be limited. The risk is that factors related to efficiency and effectiveness, the task of managements is not on the agenda. In non-profit organizations in the public sector the consideration of efficiency and
effectiveness can be extra important since several of the organizations are foremost goal driven. The empirical findings showed that several managers lack understanding of what should be assessed in order to achieve organizational value. The Value Case gives managers a possibility to discuss and reflect upon important factors related to the organizational value of IT investments. Hopefully, a Value Case can support managers and reduce the risks to IT projects failures since increased understanding will be achieved about how the IT investments match the strategy, the effects on and outside the organization.

As Stockdale and Standing (2006) write; “A decision on what is to be evaluated is a far more complex process than might first appear and is significantly influenced by the stakeholders and by the context of the organization” (p. 1092). In future research, it would be of interest to continue to discuss the content of IT evaluation and hopefully provide an analyze tool. Such analyze tool can for managers increase the understanding of how an organization will be affected by a new IT investment, and essential factors and criteria to reflect upon when assessing organizational value of IT investments? It is time to take a step from an IT system focus, focusing on the stakeholders of the IT system, to a organizational focus and foremost create understanding for how they will interact.

6 CONCLUSION

Prior IS research has concluded that only an economic view is not sufficient when evaluating organizational value of the IT investments. Instead an interpretative IT evaluation approach has been put forward. The content is based on the stakeholders’ perceptions on reality. But if the stakeholders lack experience or the strategic view and organizational goals are not evident, how will organizational value be evaluated? What to evaluate has by prior research been described as more complex than it might appear. The empirical findings show that the involved managers perceive that assessment of organizational value of an IT investment can be improved if the content of IT evaluation can be supported by pre-determined factors related to effectiveness and efficiency. Therefore, in this study a Value Case is developed that puts management’s attention to effectiveness (doing the right things) and efficiency (doing things right). The Value Case has a pluralistic, a formative and a formalized view. Gives managers a possibility to discuss and reflect upon factors related to how the IT investment will affect the organization and its surroundings and thereby create better understanding of the organizational value of the IT investments. The discussed factors are strategic match, impact on the organization and its surroundings, benefits, costs, risks, the IT system as such, the stakeholders’ view of the system, and project organizing. Hopefully future research will continue to discuss the content of IT evaluation and what should be reflected upon in order to facilitate for managers the assessing organizational value of IT investments.

References


CURRENT TRENDS AND FUTURE DIRECTIONS IN THE PRACTICE OF HIGH-LEVEL DATA MODELING: AN EMPIRICAL STUDY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0040.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Conceptual modelling, Practice, IS modeling and design, IS professionals</td>
</tr>
</tbody>
</table>
CURRENT TRENDS AND FUTURE DIRECTIONS IN THE PRACTICE OF HIGH-LEVEL DATA MODELING: AN EMPIRICAL STUDY

Barbara Anglim*, Simon K. Milton^, Jayantha Rajapakse*, Ron Weber*
*Faculty of Information Technology, Monash University, Australia.
^Department of Information Systems, The University of Melbourne, Australia.

Abstract

Many organizations now purchase and customize software rather than build information systems. In this light, some argue that high-level data modeling no longer has a role. In this paper, we examine the contemporary relevance of high-level data modeling. We addressed this issue by asking 21 experienced data-modeling practitioners to reflect on their work and to give their opinions on trends and future directions in high-level data modeling. We analyzed transcripts of our interviews with them using Klein and Myers’s (1999) framework for qualitative research. We found considerable variation in the practice of high-level data modeling. We also found that high-level data modeling is still considered important, even though organizations ultimately may purchase off-the-shelf software. The reason is that high-level data modeling assists organizations to obtain clarity about IT project scope and requirements, thereby reducing the risk that costly implementation mistakes will be made.

Keywords: high-level data-modeling practice, conceptual data-modeling practice, logical data-modeling practice, enterprise systems, package selection.

1 INTRODUCTION

Historically, organizations have built information systems in-house to solve business problems, exploit business opportunities, support business processes, and enable new products or services to be delivered. In this context, part of the practice of building an information system involved analyzing the data needs of an organization so that databases could be designed and implemented to support the system. This requirement gave rise to the traditional database design life cycle of steps reported in contemporary textbooks (e.g., Teorey et al. 2006; Hoffer et al. 2005) and attempts to develop relevant theory to support database design (e.g., Batra 2007; Codd 1970; Wand and Weber 2002).

The steps commonly articulated in database design (e.g., Teorey et al. 2006; Hoffer et al. 2005) and more broadly addressed in systems analysis (e.g., Kendall and Kendall 2004) include two that are often called conceptual data modeling and logical data modeling. Together, these steps are sometimes called high-level data modeling. They are undertaken before physical data modeling, which involves taking into account a target database management system and a hardware/software platform.

In this research, our focus is on conceptual and logical data modeling (i.e., high-level data modeling). Both steps involve constructing and using data models to represent, clarify, define, and relate important business phenomena. Often the distinction between the two steps revolves around the need in logical data modeling to consider key challenges with and to avoid common flaws in database implementation. In contrast, conceptual data modeling focuses much more on the meaning, use, and definition of and relationships among key business entities.

The motivation for our research is that the need for, purposes of, and activities involved in high-level data modeling appear to be changing significantly. For instance:
Many information systems are now considered to be a commodity. As a result, organizations often no longer build information systems. Instead, they buy ready-made packages to support their information systems needs (Davenport 1998; Shang and Seddon 2002). Often these packages require extensive configuration and customization (Markus and Tanis 2000).

To support managerial decision-making, many organizations now build data warehouses and use various business-intelligence tools that obtain data taken directly from business transactions (Kimball et al. 1998; Wixom and Watson 2001).

The strategic importance of data is increasing, especially at the corporate level (e.g., Shanks 1997). For instance, acquisition and merger activities are unlikely to be successful if information systems cannot be integrated.

Patterns are playing an increasingly important role in contemporary high-level modeling (e.g., Batra, 2005; Hay, 2006; Silverston 2001a, 2001b). Furthermore, a number of industry-standard data models are now appearing (e.g., in health services, http://www.hl7.org/, telecommunications, http://www.tmforum.org/, and supply-chain management, http://www.supply-chain.org/).

These changes affect the practice of high-level data modeling. Historically, some research has been done on data modeling practices (e.g., Batra and Marakas 1995; Floyd 1986; Necco et al. 1987). A paucity of research exists, however, in relation to contemporary high-level data modeling practice. Some work has been done on design and development methodologies (e.g., Kautz et al. 2004) and process modeling (e.g., Chang et al. 2001), but little work has been done on data modeling broadly or high-level data modeling specifically.

In response to a paucity of research on conceptual modelling practice, Davies et al. (2006) conducted a survey of conceptual modeling practitioners in Australia. They had two objectives (p. 359): (a) “to obtain empirical data that conceptual modeling is indeed being performed now and into the foreseeable future in IS practice in Australia”; and (b) “to find out what are the principal tools, techniques, and purposes for which conceptual modeling is performed currently in Australia.” Their focus was conceptual modeling broadly, including process and workflow modeling additional to data modeling. They found that (p. 376) “[d]atabase design and management remains the highest average purpose for use of modeling techniques,” and that the reasons for continued use of all conceptual-modeling approaches included “communication…to/from stakeholders, internal knowledge…of techniques, user expectations management, understanding models integration into the business, and tool/software deficiencies.” The limited qualitative data that came from the survey constrained the depth with which current practice and future trends in data modeling practice could be studied.

In this research, our goal was to explore current practices and trends in high-level data modeling in some depth. We conducted field research in which we undertook semi-structured interviews with 21 highly experienced practitioners to obtain their views about high-level data modeling. We sought a deep understanding of the reasons behind the answers they provided to our questions. Specifically, we used the following broad questions as a “scaffold” for the interviews we conducted with them:

- Why is high-level data modeling undertaken?
- What high-level data-modeling activities are undertaken?
- How are high-level data-modeling activities undertaken?
- Who is involved in undertaking high-level data-modeling activities?
- What quality-assurance measures are used in high-level data modeling?
- What is the future of high-level data modeling?

We canvassed these questions with our interviewees in a fairly informal way. Our concern was to use the questions to elicit and trigger their opinions and reflections rather than to impose a rigid structure on our interactions with them. We wanted to understand the nature and purposes of the activities they undertook and the meaning they ascribed to high-level data modelling (Neumann 2007, pp. 276-278).
The remainder of the paper proceeds as follows. First, we discuss our research method. Next, we present our results in relation to each of our research questions. We then summarize our findings. Finally, we present some implications of and some strengths and limitations of our research.

2 RESEARCH METHOD

Because our research is exploratory, we adopted an interpretative research method. In this light, we followed Klein and Myers’ (1999) principles for the conduct of interpretative research (Table 1). During 2007 and 2008, we carried out interviews with 21 experienced data-modeling professionals engaged in high-level data modelling. Each interview lasted about 90 minutes. All interviews were transcribed. Using techniques from Miles and Huberman (1994), we then sought to make sense of interviewees’ experiences with and views about high-level data modeling. Specifically, we employed a descriptive matrix to relate comments made by our interviewees to our broad research questions.

<table>
<thead>
<tr>
<th>Klein and Myers Principle</th>
<th>Approach to Addressing the Principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contextualization</td>
<td>Our experience allowed us to understand interviewees’ responses and reflections and to engage actively with interviewees.</td>
</tr>
<tr>
<td>Interaction between researchers and interviewees</td>
<td>We have substantial academic and practical data-modelling experience. This experience enabled us to establish rapport with interviewees, which facilitated disclosure and a deeper understanding of issues that the questions surfaced.</td>
</tr>
<tr>
<td>Hermeneutic circle</td>
<td>Analysis of interview transcripts followed multiple cycles. We constantly reflected independently and as a research team on the transcripts until we felt we understood underlying themes that interviewees were seeking to convey.</td>
</tr>
<tr>
<td>Abstraction and generalisation</td>
<td>By applying principles in Miles and Huberman (1994), we linked emerging themes back to our research questions.</td>
</tr>
<tr>
<td>Dialogical reasoning</td>
<td>A research assistant with data-modelling training but no commitment to a particular theory or method of data modeling collated an initial list of key themes from the transcripts before we explored them deeply.</td>
</tr>
<tr>
<td>Multiple interpretations</td>
<td>We recognized that interviewees sometimes held different views about data-modeling phenomena and sought to understand these conflicting views from multiple perspectives.</td>
</tr>
<tr>
<td>Suspicion</td>
<td>We used our experience to see when interviewees brushed over (or avoided) questions and rephrased our questions to try to elicit more-helpful responses.</td>
</tr>
</tbody>
</table>

Table 1. How we followed Klein and Myers’ (1999) principles for interpretive research

Table 2 provides an overview of our interviewees. All had constructed at least five major data models in at least two different industries or areas of government. Some were highly regarded among their peers (both nationally and internationally) because of their substantial data-modeling expertise.

<table>
<thead>
<tr>
<th>Interviewee</th>
<th>Number of High-Level Models Completed</th>
<th>Industries Where Work Undertaken</th>
</tr>
</thead>
<tbody>
<tr>
<td>I#1</td>
<td>10+</td>
<td>Banking, health, telecommunications.</td>
</tr>
<tr>
<td>I#2</td>
<td>Unknown, but many</td>
<td>Many; over 30 years’ experience.</td>
</tr>
<tr>
<td>I#3</td>
<td>30-40</td>
<td>Half-dozen different industries.</td>
</tr>
<tr>
<td>I#4</td>
<td>100+</td>
<td>Many, including manufacturing, telecommunications, banks, insurance, social security, community services, land title offices, defense</td>
</tr>
<tr>
<td>I#5</td>
<td>6+</td>
<td>Banking, insurance, telecommunications, higher education.</td>
</tr>
<tr>
<td>I#6</td>
<td>200+</td>
<td>Many, including government, construction, manufacturing, retailing.</td>
</tr>
<tr>
<td>I#7</td>
<td>10-12</td>
<td>Health insurance, life insurance, property, telecommunications, finance, retail.</td>
</tr>
<tr>
<td>I#8</td>
<td>100+</td>
<td>Many; more than 20.</td>
</tr>
<tr>
<td>I#9</td>
<td>50+</td>
<td>Many; over 30 years’ experience, including fire management, game</td>
</tr>
</tbody>
</table>
Table 2. Overview of research interviewees

We sought interviewees by making presentations to local chapters of the Data Management Association (DAMA) and Data Warehousing Association (DWA) in three cities. We explained the nature of our research, its purposes, the assistance we required, and the qualifications of individuals from whom we sought assistance. We obtained contact details from individuals who indicated they were willing to assist us. In some cases, to comply with ethics restrictions, we used ‘snowballing’ whereby individuals asked other data modelers to contact us to indicate their willingness to participate in our research. Subsequently, we scheduled interviews with all who had agreed to assist us.

3 RESULTS

The following subsections present our findings in relation to each of our major research questions. We provide illustrative quotations to support our findings. Quotations are labelled as, for example, I#4, which means a quote from Interviewee 4’s transcript.

3.1 Reasons for Undertaking High-Level Modeling

Previous research has considered conceptual modeling as a precursor to logical modeling, which in turn leads to physical data modeling. This sequence assumes that an application supported by a database will be built. Interviewees had alternative views on reasons why conceptual modeling was now undertaken by organizations.

Scoping for proposed projects was the most-common reason given for preparing a conceptual model. I#4, who is a highly experienced modeler, stated that rigor is unimportant in conceptual modeling. He commented that the primary purpose of the conceptual models he prepared was to scope the project. Furthermore, he argued that conceptual models are not prepared with the objective of “getting towards implementation or working out business rules.” I#6 also considered that articulating rigorous business requirements was unnecessary at the conceptual-model level. Instead, conceptual modeling was done just to “scope the next level of activity.” I#2 describes a model as a “means to the end” with a project’s scope report as a likely associated key deliverable. I#1 described “scope as possibly the key part of the conceptual model.”

I#5 said she created conceptual models for two purposes. First, she used them to build an enterprise data model, which she described as not necessarily leading to the immediate purchase of a package or development of a bespoke system. Instead, a conceptual model was a resource that had “many
purposes.” Second, sometimes she used an application conceptual data model for a “particular application” development. Others reinforced this view. For example, I#3 described how a quality conceptual model and organizational overview helped direct an organization’s selection and purchase of package software and development of applications. In particular, the model and overview helped avoid the pitfalls of systems that evolve independently. He argued that it was better to have the applications “logically conforming to some overarching model than to have them evolve independently and then try and put them together, which is where we’re at today.” Similarly, I#1 spoke of using high-level data models to find “pain points” with proposed systems and thereby to manage risk.

The need for integrated information systems, particularly data warehouses, was the driver for some to undertake entity definition work at the conceptual level. In these contexts, interviewees stressed the need for entity definition resolution. Interviewees found potential in conceptual models for transcending individual projects. They were long-term resources that could be used to help integrate and plan an organization’s informational needs, projects, applications, and data-warehouse activities.

Without a conceptual model to guide configuration, I#3 argued that configuring a package becomes more expensive and error prone. He commented that lack of a conceptual model to provide guidance during package configuration led to arguments about the meaning of data among different departments in an organization. He gave an example where an organization found it could migrate only 30 percent of data from six billing systems to a new system because of the lack of agreed data definitions. The result was seven systems that lacked interoperability. He was not alone in this observation.

I#1, who has supervised many teams on major projects for decades, argued that high-level data modeling is needed to help avoid major problems for organizations by guiding the evolution of data needs. “I see the design documents as a key resource for the company.” While business evolves over time, I#7 argued that conceptual modeling done correctly often needed only fine tuning over time.

3.2 Types of High-Level Data Modeling Activities Undertaken

Interviewees had different views about the nature of conceptual versus logical models. Each particular view appears to be linked to the depth of the project they were undertaking, the domain in which they were modeling, and the complexity or size of the organization they were modeling.

Although interviewees varied in the activities they undertook at the two levels (and sometimes varied in the modeling activities they undertook at these two levels between projects), most considered preparation of conceptual models to be a phase that was distinct from preparation of logical models. Specifically, they indicated the logical level involved adding more details to the model (e.g., keys and attributes) and resolving many-to-many relationships. During the preparation of a logical model, I#4 said that he added more detail to a conceptual model and “nailed down the rules.” I#6 commented that rigor in a conceptual model was unimportant. He indicated that his logical or “subject” models had keys and relationships, but his conceptual or “enterprise” models had no keys or relationships.

Those interviewees who constructed enterprise conceptual models did not fully articulate their models. For instance, a fully articulated ER (entity-relationship) diagram would have keys, attributes, and fully resolved many-to-many relationships. On those occasions where he used ER diagrams, I#1 was clear that relationship resolution did not occur at the conceptual-modeling level. Instead, his focus was on identifying “scope” and “pain points” and achieving some agreement about entity definition.

Most interviewees indicated that the largest amount of work at the conceptual-model level involved getting agreement on definitions. I#4 said this task as “difficult.” I#3 commented that obtaining agreement on name definitions was not always possible, and he described it as an ongoing job. Nonetheless, he argued that obtaining agreement on name definitions was important because it meant that business requirements between departments could be understood better and more easily integrated. I#15 thought it important to get the right stakeholders with expertise because it was “a question of confidence in the answer” and he could “manage expectations.” I#7 endeavors to have one agreed definition for each entity and attribute to ensure “multiple departments are speaking the
same language.” She said it is a mistake to allow different departments or applications to have different definitions. I#10 spoke about the importance of consensus about data definitions: “Build up a glossary and say if we’ve got two different terms for them and maybe let’s think of a term that covers the general, both of them but just trying to agree on the same language.”

I#5 indicated she focuses on definitions needing resolution. She also looks for key performance indicators, as these often needed definitional agreement. I#5 indicated she had recently completed a project to help the client get a “structured business vocabulary” as a first step. The client then followed through with an articulation of business rules using the agreed vocabulary.

Whether entity definition occurred at the conceptual or logical level varied among interviewees and the types of projects they undertook. I#8 indicated he did not undertake entity definitions during conceptual modeling but rather left this task for “techies.” Instead, he gave a “description of what these areas of information are.” Most interviewees considered that attaining some agreement on entity definitions was a key part of conceptual model work, however. For example, at the conceptual-model level, I#1 argued the main deliverables were the model and entity definitions. At the logical level, he would include some key attributes. Many interviewees thought a large part of conceptual modeling was establishing business vocabulary or metadata in the form of data definitions relating to business rules. The client organization could then use these definitions as a business “language” for communication across divisions and applications.

I#9 had a different view, however, about the link between conceptual models, logical models, and physical models. He commented, “they (a client) took that as a logical design which they put into physical code and they got it straight from the conceptual model. Now is that blurring the boundaries?” Even during his conceptual modeling work, he included a high level of detail in the models he prepared, primarily because he relied substantially on previously developed data modeling patterns. I#12 also included high levels of detail in the models he developed. I#15 had changed his mind about the role for logical data models by saying that “the physical needs to be supported.”

Model completion usually was not expected. I#5 said she only represented a few types of phenomena in a conceptual model. Sometimes she might prepare just a simple conceptual model with a few key entities and definitions as a basis for communication among stakeholders in some domain. At other times in small projects, she prepared a conceptual model that included attributes and relationships (which she described as eventually constituting a logical model). She described her most-recent project as being supervised by an experienced modeler who expected rigor. On this project, she had to include detail like relationships at the conceptual level. Moreover, her conceptual model had to link coherently to a logical model that included keys and other “technical information.”

Some interviewees said they undertook sub-typing of entities in their models. For instance, I#3 described sub-typing at the logical level as time consuming and expensive. Nonetheless, he felt the clarification of business rules saved much time in implementation. I#14 avoided sub-typing, however, because when he explained “sub-class, sub-typing, and bringing in sort of universal data modeling type concepts” to business personnel, “they get lost.” By way of contrast, he emphasized the need for abstraction: “So you’ve got to move up a level there and think of that level and that’s important.”

### 3.3 Ways in Which High-Level Data Modeling is Undertaken

All except one interviewee used workshops to garner the business information required for the data models they needed to construct. Attendees at the workshops were almost always representatives of the business. Some interviewees emphasized that workshops also helped to obtain agreement among client representatives about priorities and definitions of business objects or metadata. For instance, I#11 commented: “Often it’s good for us to get into the workshop and try and thrash out some of the nitty gritties with a lot of business units because often they have conflicting priorities.”

Helping participants feel comfortable during workshops was a reason behind some interviewees’ choice of material or methodology. For instance, rather than using software tools to build a model,
I#7 indicated she uses whiteboards and hand-drawings to help client representatives feel relaxed and to build the model in conjunction with them. I#10 follows a similar approach: “Really, to be truthful I tend to start off with a whiteboard and some people sitting around and saying what are the concepts.”

Some interviewees use standard data model patterns that they customize to suit the context in which they are working. For instance, I#1 argued, “You do not need to reinvent the wheel.” He had specialized in banks and telecommunication companies where transactional relational databases are common. He had used the Financial Services Data Model, which is a proprietary data model developed for the banking and finance sector.

I#9 was the strongest advocate for the use of patterns. He commented that he used patterns “I guess, in ever-increasing amounts of detail. I can knock together, and this is going to frighten you, I can knock together an enterprise-wide conceptual model in probably 30 to 60 minutes. And you say you’re joking, but I can interact with senior management and start to talk and their eyes light up when they realize, hey you understand and all I’m doing is patterns.” I#14 is also a strong believer in patterns: “I’m a great believer in utilizing what other people have done.” I#12 did not use patterns, but he reused models he had developed during previous engagements with clients.

I#8 said that he started with a straw-man pattern and looked for disagreement with it, because he frequently had only a few days to complete a model. He commented that he often had already decided what package was likely to suit his client’s business. As a result, he would look to make only slight adjustments to the patterns he used. He saw patterns as a fast means of getting input from end users. I#10 also sometimes used a straw-man approach: “The other way is probably a bit more common…come up with a straw-man yourself and then go through and sort of say well is this right.”

One of the more-experienced interviewees, I#4, replied that he used patterns only occasionally and used them only in the sense of capturing previous experience. He indicated that he discarded them quickly, however, if they did not fit the context. I#7 found dangers when presenting models prior to eliciting business requirements. She warned that end users might simply accept a pattern without sufficient reflection on its accuracy and completeness. As a result, she emphasized the importance of starting with a clean whiteboard and working with end users to create their own model.

I#3 was the most outspoken against patterns. He argued that it was “best practice” not to assume similarities among models and to commence modeling with an open mind. He referred to recent experiences with patterns in which he found that “a US telco is very different from an Australian telco.” Despite the fact that telecommunication companies often offer similar products, he found that they applied different business rules that required different models. He had found that patterns did not save time and led to important omissions in data definitions. I#5 commented that some proprietary data models were too expensive for some organizations to purchase.

I#19, who has worked primarily in government, was keen to identify “reference sets” in his modeling workshops. He indicated that he “would source (reference data) from somewhere else.” Examples he gave included “countries of the world,” “area codes,” and “telephone prefixes.”

3.4 Participants in High-Level Data Modeling

Interviewees were split as to whether IT professionals should be included in workshops. If IT professionals were present, however, interviewees held a common view—namely, that IT professionals should play little or no role in the discussions that ensued in a workshop. Instead, IT professionals should play the role primarily of observers.

I#5 voiced one perspective saying she “does not deal with IT folks… I may have some present but only as observers.” I#18 preferred “to concentrate on the business problem.” Others had regrets like I#6: “I wish I did work with people at that (the technical) level, but consultants come in.”
3.5 Quality-Assurance Practices in High-Level Data Modeling

Throughout our interviews, it became clear that quality-assurance practices like audits, walkthroughs, reviews, metrics, and documentation standards received little attention during high-level data modeling. Our findings stand in contrast to the modeling standards now in force in organizations such as the U.S. Defense Department (where UML has been adopted as the conceptual-modeling standard). Interviewees stressed that model correctness and completeness were important for high-level data modeling, but that standards were not defined clearly. For instance, interviewees selected a modeling methodology depending on the project, personal leanings, or their client’s preferences. Moreover, conformity with standards was often impossible because of time constraints imposed on them.

We asked interviewees how their models were validated and updated. Many indicated they were hired as consultants and were not given time to validate and correct their models. For the most part, any validation that occurred was an outcome of interactions among attendees at modeling workshops.

I#7 aims to get consensus about correct and complete business requirements and data definitions despite the fact that “a lot of companies are not willing to invest that time.” She indicated that some organizations were willing to devote only an hour or two to model validation. I#8 acknowledged that getting agreed data definitions was pivotal to quality and validation concerns. Nonetheless, he argued that trying to get agreement on definitions quickly was “counterproductive.” His view is that during modeling he must “understand business and not techno stuff.” He prefers to get a “description of what these areas of information are” rather than an agreed definition because “we can’t let things drag on.”

Interviewees involved with data-warehousing projects seemed to have more opportunities to validate and update their models. For instance, I#5 said she was responsible for updating her conceptual models if errors and omissions were found during logical-level modeling. Similarly, I#7 remains engaged until the “physical-build” stages of a warehousing project. She oversees quality checks that she suggests for her logical model. Her model is updated when errors and omissions are found.

With governments, I#4 indicated he has been given the chance to undertake quality reviews. He went through many cycles of feedback with stakeholders until differences relating to the models he had prepared were resolved. Nonetheless, for most of his projects, achieving correct and complete high-level data models was not considered important. He commented, “you do not really worry about getting it perfect,” because once the model is used it becomes evident how it should be “tweaked.” I#8 went further and stated that aiming to get most data definitions right at the conceptual level is “counterproductive.” He felt definitions and relationships should be left for “techies.”

Interviewees were asked about the tools they used to enhance the quality of their models, such as tools to check syntax and reverse engineer models based on existing systems. They responded that they used whatever tools their client organization wanted or provided. Few mentioned use of expensive tools. Those who did often had large government organizations as their clients. For instance, I#6 described liking the “Mega” tools, and commented that “System Architect” was “brilliant.” Nonetheless, he indicated that these tools often require full-time personnel to support their use because they are highly complex.

I#6 said that he preferred to use the types of tools that allow his models to be linked with other models (including physical models) and ultimately with the data that populates his models. He described the tools as not so much CASE (computer-aided software/system engineering) tools but as tools to manage business areas on many levels. He stated that use of these tools enhanced the quality of models because they facilitate validation and update of the models—they allow a system’s requirements and architecture to be integrated and considered from several perspectives. In relation to the projects he undertook, I#6 pointed out a trend toward increasing demands of cohesiveness between (a) different organizations models, and (b) business and technology. He observed that some data-modeling tools were evolving to support this trend.
I#5 sees herself as exceptional among modelers in her attempts to provide metadata about each object in her models. She spoke of a procedure whereby a “series of assertions” derived from a model was presented to stakeholders for validation as a thorough, best-practice technique. Her experience with using this approach, however, was that business people were unable to understand it or refused to devote time to it. Instead, like I#16, she recommended using a walkthrough with stakeholders that focused primarily on entities. I#18 used quasi-assertions in which he would “explain each one of those relationships both backward and forward and find out... is this correct or is it not correct?”

Like I#5, I#3 seeks to cross-reference metadata definitions with business rules to test the validity of his models. He was the only interviewee who expressed his ability to undertake model validation in a fairly formal way. He used the most-extensive procedures to validate his models. For him, quality was also about exploring business requirements rather than using presupposed patterns. He uses three validation techniques. (1) He presents his model to stakeholders for feedback and questions and resolves errors, omissions, and ambiguities. (2) He describes the business rules in a spreadsheet and asks stakeholders to check their validity. (3) He asks stakeholders to check his model through scenarios. He particularly focuses on the validity of subtypes in his models.

In short, our interviewees indicated that they had only limited influence over quality-assurance procedures. Unlike other areas of IT practice where quality checks, metrics, validations, updates, and reviews are now standard, they often are not features of high-level data-modeling environments. Moreover, quality problems appear to be compounded by a general lack of stakeholder input and interaction and the distribution of modeling activities across different IT professionals.

### 3.6 Future of High-Level Data Modeling

Interviewees were emphatic and unanimous about the ongoing need for high-level data modeling. Several argued that “erroneous ideas” about the declining importance of high-level data modeling were “driven from the package industry.” I#7 argued that good high-level data models save millions of dollars by avoiding the purchase of inappropriate, sometimes expensive, software systems. I#1 indicated that a package might not fit all business requirements. He believes that knowledge of the overall business requirements is necessary to make an informed decision about package purchases, especially expensive packages.

I#5 believes a resurgence in high-level data modelling is occurring, driven partly by greater use of Service Oriented Architectures (SOAs), which are a relatively low-cost approach to aid application integration across organizations. I#5 is working increasingly with “information architects,” who have been appointed by their organizations to achieve the goal of better data management. I#5 argued that pressures to develop high-level models quickly were counterproductive to the growing need of “providing an enterprise asset for the organization.”

I#3 commented that systems integration and interoperability were now a high priority for many organizations. He argued that high-quality, high-level data models help with the development of applications systems and the purchase of packaged software. They enable organizations to avoid the pitfalls associated with bespoke systems that are developed independently. Moreover, they help ensure that packages can be integrated with other systems in the organization. In the absence of high-quality high-level data models, I#3 argued “some really expensive failures” can result.

I#7 argued that high-level data models were increasingly important to effective strategic planning. Some types of organizational strategy required systems integration. In this regard, I#6 said that the need for integrated information systems required organization-level data management, which in turn was linked to a trend toward increasingly big and complex organizations and applications.

I#6 argued that the assets of an organization are “in the requirements and design,” because nowadays one can “auto generate code” or “select packages.” Therefore, he believes that increasing emphasis will be placed on developing high-quality design documents and data models. He also stressed the importance of governance to support high-level data modeling. He has observed many times that
models “fell apart” when good governance structures and procedures were not in place and working. He explained that a clear governance structure effectively allows discrepancies among high-level data models to be resolved and an organization to own changes that it makes to its data models. He now prefers to work only for organizations that have effective governance in place. 1#6 argued that effective governance occurs, however, only when a leader recognizes the importance of clearly stating business process requirements and is willing to devote resources to support quality business analysis.

I#21 was circumspect about the role that high-level data models could play in package selection by helping to drive a “side-by-side functionality capability check” between the package and requirements. I#19 cited a specific example of how data modeling helped build a requirements document where the package vendor “knew exactly what our requirements were at a data level, at the process level, at the use-cases.” Another project took nine months longer than expected where the vendor “saw us coming a mile away... because the requirements were not clear” after I#19 had warned the internal client “you need to have models for this.”

Where companies continue to build bespoke systems, agile methodologies are now in the ascendancy. Those data modelers who used agile methodologies identified advantages and challenges. I#20 sees a key role for high-level data modeling and observed that sometimes “what falls off the bottom of agile prototyping is the underlying model,” which means it is difficult to extract data into warehouses. I#21, who is also experienced with agile methodologies, said that “we use a tool to gather the stories and exercise the model... on how well it supports the business functionality... (and if it) can actually support the business processes.”

4 SUMMARY OF FINDINGS

Table 3 is a descriptive matrix that summarizes major themes (Miles and Huberman 1994, p. 246) we identified in the responses provided by interviewees to our research questions.

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Findings/Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why undertake high-level data modeling?</td>
<td>To scope projects, build enterprise data models, guide the evolution of an organization’s data needs, select and configure packages, manage project risk, and develop applications.</td>
</tr>
<tr>
<td>What high-level data-modeling activities are undertaken?</td>
<td>Building conceptual data models in which entities and sometimes relationships are defined. Building logical data models in which attributes and keys are defined. Gaining agreement on entity definitions.</td>
</tr>
<tr>
<td>How are high-level data-modeling activities undertaken?</td>
<td>Via workshops and sometimes by using patterns, past models that the modeler has developed, industry-standard data models, and CASE tools (often those tools mandated by the client).</td>
</tr>
<tr>
<td>Who is involved in high-level data-modeling activities?</td>
<td>Modeler and most often key business stakeholders.</td>
</tr>
<tr>
<td>What quality-assurance measures are used in high-level data modeling?</td>
<td>Quality-assurance activities are limited due to lack of client resources and enthusiasm.</td>
</tr>
<tr>
<td>What is the future of high-level data modeling?</td>
<td>To articulate business requirements before package selection, help manage data that needs to be integrated in service-oriented architectures, and improve the effectiveness of strategies that lead to system integration in large, complex organizations by clearly defining data to be integrated.</td>
</tr>
</tbody>
</table>

Table 3. Summary of results

In our view, these themes enable us to characterize contemporary high-level data-modeling practice in the following way:

- Relative to the past, most high-level data modeling is now undertaken with much-less emphasis on database design. Instead, the goals are to (a) identify critical business concepts and their relationships, and (b) provide guidance for projects, especially in terms of a project’s scope.
• Group techniques dominate contemporary high-level data modeling. Experienced modelers play the important role of facilitating stakeholder discourse. They assist stakeholders to share ideas and information, negotiate outcomes, and establish shared vocabularies. IT professionals are seldom included as active stakeholders in group activities.
• High-level data modeling increasingly is used to underpin package selection, data-warehouse design, and enterprise planning.
• Quality assurance in relation to high-level data models often is not undertaken as a distinct activity. Instead, it is an active process that occurs, somewhat subliminally, during workshops. To some extent, it is facilitated by using industry-standard data models and data-model patterns.
• Experienced data modelers see an ongoing and more-important role for high-level data modeling.

5 IMPLICATIONS, STRENGTHS, AND LIMITATIONS

Our findings have implications for research, teaching, and practice. For research, it is clear that the objectives of and practices associated with high-level data modeling are changing. A deeper understanding is needed of the motivation for and nature of these changes and the likely ways they will unfold. Our findings show that variations exist in the views of and activities undertaken by practitioners. A deeper understanding of why these variations exist is required. Ultimately, theories are needed to account for the variations.

For teaching, we believe high-level data modeling still needs to be an important part of the curriculum taught to intending IT professionals. Moreover, it appears that conceptual modeling, logical modeling, and physical modeling increasingly are becoming distinct areas of expertise and practice. While students need to understand how these three types of data modeling are related, it may be advantageous for them to develop higher levels of competence in one particular type of data modeling.

For practice, we believe the management of organizations should revisit the role that high-level data modeling plays in their planning activities and the operations of their organizations. In some organizations, this role seems to have become diminished, mainly because of some of the rhetoric around package software (e.g., that packages obviate the need for detailed requirements specifications and thus data models). Our findings show that high-level data modeling is transcending its traditional roles in system development and playing an increasingly important role in strategy and planning.

A strength of our research is that it is one of few studies investigating how highly skilled practitioners undertake high-level data modeling. It is also novel because it explores the under-researched role of high-level data modeling in the post-bespoke-systems era. Because we have used an interpretive approach to the analysis of data collected in our research, we believe we have obtained rich insights into the current practice of high-level data modeling. Thus, our research should inform both researchers and practitioners.

A limitation of our research is that we have obtained only the opinions and reflections of our interviewees. We did not observe them directly, and we recognize that espoused theories may not reflect theories-in-use (Argyris and Schön 1974). Furthermore, because we interviewed only active experienced data modelers, our results potentially are biased toward emphasizing the ongoing importance of high-level data modeling.

References


Towards Increased Comparability of Conceptual Models - Enforcing Naming Conventions through Domain Thesauri and Linguistic Grammars

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0573.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Conceptual modelling, IS modeling and design, Reference modelling, Domain ontology</td>
</tr>
</tbody>
</table>
TOWARDS INCREASED COMPARABILITY OF CONCEPTUAL MODELS – ENFORCING NAMING CONVENTIONS THROUGH DOMAIN THESAURI AND LINGUISTIC GRAMMARS

Becker, Jörg, University of Münster, Leonardo-Campus 3, 48149 Münster, Germany, becker@ercis.uni-muenster.de

Delfmann, Patrick, University of Münster, Leonardo-Campus 3, 48149 Münster, Germany, delfmann@ercis.uni-muenster.de

Herwig, Sebastian, University of Münster, Leonardo-Campus 3, 48149 Münster, Germany, herwig@ercis.uni-muenster.de

Lis, Łukasz, University of Münster, Leonardo-Campus 3, 48149 Münster, Germany, lis@ercis.uni-muenster.de

Stein, Armin, University of Münster, Leonardo-Campus 3, 48149 Münster, Germany, stein@ercis.uni-muenster.de

Abstract

Distributed construction of conceptual models may lead to a set of problems when these models are to be compared or integrated. Different kinds of comparison conflicts are known (e.g. naming conflicts or structural conflicts), the resolution of which is subject of different approaches. However, the ex-post resolution of naming conflicts raises subsequent problems that origin from semantic diversities of namings – even if they are syntactically the same. Therefore, we propose an approach that allows for avoiding naming conflicts in conceptual models already during modelling. This way, the ex-post resolution of naming conflicts becomes obsolete. In order to realise this approach we combine domain thesauri as lexical conventions for the use of terms, and linguistic grammars as conventions for valid phrase structures. The approach is generic in order to make it reusable for any conceptual modelling language.

Keywords: Conceptual Modelling, Model Comparison, Domain Thesaurus, Computational Linguistics
1 INTRODUCTION

Since IS modelling projects are often large-scaled, the required models are increasingly constructed in a distributed way in order to boost the efficiency of modelling (vom Brocke & Thomas 2006). I.e., different modellers participate in the modelling process developing different model parts – maybe at different places and at a different time. Furthermore, domain experts are increasingly integrated in the modelling process to comply with domain requirements as much as possible – an approach that is commonly known as End-user Development (Wulf & Jarke 2004; Fischer et al. 2004). Distributed modelling becomes more and more mainstream. This applies not only intra-corporately but also in the form of so-called modelling communities, e.g. the Open Model Initiative (Frank & Strecker 2007).

Empirical studies show that model sections developed in the course of distributed modelling can vary heavily concerning terms and abstraction level (Hadar & Soffer 2006). Therefore, the comparability of the regarded models cannot be guaranteed. Comparison conflicts may occur, which are commonly divided into naming conflicts and structural conflicts (Pfeiffer 2008). Consequently, the integration of the different model sections into one total model representing the original modelling issue may be extremely laborious. Usually, model parts are consolidated manually and with attendance of all modellers in order to reach a consensus. Sometimes, external consultants are involved additionally (Phalp & Shepperd 2000; Vergidis & Tiwari & Majeed 2008). Thus, the original sub-goal of distributed modelling – i.e. the increased modelling efficiency – may be missed.

Summarising, a considerable problem of distributed modelling is the insufficient comparability of conceptual models. In order to solve this problem, approaches are required that are able to assure comparability. In literature, there exist many contributions that propose approaches for resolving comparison conflicts subsequent to modelling (cf. Section 2).

The goal of this article is to introduce an approach that ensures the comparability of conceptual models by avoiding potential comparison conflicts already during modelling. This way, we prevent semantic problems that result from the ex-post resolution of conflicts. This article focuses on naming conflicts. We define naming conventions for elements of modelling languages and ensure their compliance by an automated, methodical guiding during modelling. The conventions are set up using the domain terms and the phrase structures that are defined as valid in the regarded modelling context. As a formal specification basis, we use thesauri that provide term conventions not only for nouns but also for verbs and adjectives. These thesauri also include descriptions of a term’s meaning. In order to provide conventions for phrase structures, we use the Head-driven Phrase Structure Grammar (HPSG) (Pollard & Sag 1994), a formal grammar for the specification of natural language syntaxes. During modelling, model element names are validated simultaneously against both the term and phrase structure conventions. Our approach is generic so that it can be applied to any conceptual modelling language.

Our approach is suitable for modelling situations, where it is possible to provide all involved stakeholders with the necessary information about the modelling conventions, i.e. modelling projects that are determined regarding organization and/or business domain. These modelling situations usually apply in companies, corporate groups or modelling communities.

The paper is structured as follows. First, we analyse related work in Section 2 and discuss the research gap that led to the development of the approach presented in this paper. Furthermore, we outline the research methodology used. In Section 3, we introduce a conceptual framework for the specification and enforcement of naming conventions. The feasibility of our approach is shown exemplarily with a demonstrator software in Section 4. Furthermore, an explorative analysis of real-world process modelling projects shows the potential of the approach reducing ambiguities in process models. We conclude the paper in Section 5 and motivate further research addressing e.g. the conflict classes not yet covered by this approach.
2 FOUNDATIONS

2.1 Related Work

Early approaches of the 1980s and 1990s discussing the resolution of naming conflicts address the integration of company databases and use the underlying schemas as a starting point (Batini & Lenzerini 1984; Batini & Lenzerini & Navathe 1986; Bhargava & Kimbrough & Krishnan 1991; Lawrence & Barker 2001). Hence, these approaches focus on data modelling languages, mostly dialects of the Entity-Relationship Model (ERM) (Chen 1976). Names of schema elements are compared and, this way, similarities are revealed. The authors state that such a semantic comparison can exclusively happen manually. Moreover, only single nouns are considered as names. In contrast, in common modelling languages and especially in process modelling languages, names are used that consist of sentence fragments containing terms of any word class. This aspect is not regarded by schema integration approaches. Thus, they are only suitable for the mentioned data modelling languages.

Other approaches make use of ontologies (Gruber 1993; Guarino 1998) in order to address the problem of semantic comparison of names. These approaches act under the assumption that there exists a “generally accepted” ontology describing a certain modelling domain. It is assumed that all considered models of this domain comply with its ontology, i.e. that modellers had a thorough knowledge of the ontology. E.g., Höfferer (2007) connects domain ontologies to the terms that are used as element names in conceptual models. This way, he establishes relationships between elements of different models that are to be compared and identifies similarities. In addition to ontologies, Ehrig, Koschmider, and Oberweis (2007) define combined similarity measures that consist of syntactic and semantic parts. These serve as a basis for the decision whether or not the model elements compared are equivalent. Consequently, it is argued that if identical terms – or those that are defined as synonymous within the ontology – are used in different models and by different modellers, they can be considered as semantically identical as well (Koschmider & Oberweis 2005; Sabetzadeh et al. 2007).

However, solely the circumstance that two modellers act in the same business domain does not guarantee at all that they share the same or an equivalent understanding of business terms. If a “generally accepted” ontology is available, it is suitable for model comparison if and only if it is explicated and can be accessed by all involved modellers already during the modelling process. Furthermore, in order to ensure comparability of the models, modellers have to comply strictly with the ontology. Most approaches make the implicit assumption that these preconditions are already given rather than addressing a methodical support. So far, there exist only few approaches that address such a support. Greco et al. (2004) propose adopting terms from existing ontologies for process models manually. However, due to manual adoption, correctness cannot be assured. Born, Dörr, and Weber (2007) propose semi-automated adoption of model element names. However, they restrict their approach to BPMN models (White & Miers 2008).

In order to avoid the problems of ontology based approaches, Pfeiffer (2008) suggests integrating domain semantics into modelling languages, i.e. the names of model elements are preset. Thus, he proposes domain-oriented building block-based process modelling languages. With the PICTURE modelling approach, he concretises them for the domain of public administration. Since PICTURE consists exclusively of 26 process building blocks with defined semantics, naming conflicts are a priori impossible. As the author states himself, this language class is restricted to mainly linear business processes.

Only few approaches, mainly from the German speaking area, suggest standardised phrases for model element names in order to increase the clarity of process models. The phrase standards are summed up as modelling conventions. E.g., Rosemann (1996) and Kugeler (2000) propose particular phrase structure guidelines for names of process activities (e.g. <verb, imperative> + <noun>; in particular e.g. “check invoice”). Moreover, the authors propose so-called Technical Term Models (Rosemann 2003) that have to be designed previously to process modelling and that specify the terms to be used within the phrases. Therefore, the scope of Technical Term Models is restricted to nouns. Similar approaches provided by Koschmider and Oberweis (2005) and Sabetzadeh et al. (2007) propose preposi-
tioning generally accepted vocabulary. Further approaches recommend connecting names of model elements to online dictionaries in order to establish semantic relationships of terms (Rizopolous & McBrien 2005; Bögl & Kobler & Schrefl 2008). These online dictionaries consist of extensive collections of English nouns, verbs, and adjectives as well as their semantic relationships. Actually, the proposed approaches are promising regarding increased comparability of conceptual models since all of them aim at standardising names for model elements prior to modelling. However, up to now, a methodical realisation is missing.

To sum up, we identify the following need for development towards avoiding naming conflicts in comparing conceptual models: Up to now, methodical support for (1) the formal specification of naming conventions for all word classes and (2) the formal specification of phrase structure conventions is missing. Furthermore, there exists neither methodical support for (3) the integration of conventions in modelling languages nor (4) for guiding modellers in order to comply with the conventions.

In order to realise such a methodical support, we propose an approach that consists of (1) a formalism to specify thesauri covering nouns, verbs, and adjectives, (2) a grammar to specify phrase structures that can hold terms specified as valid within the term models, (3) a framework to integrate these naming conventions with modelling languages, and (4) a procedure to guide modellers automatically in complying with the conventions.

2.2 Research Methodology

The research methodology followed here complies with the Design Science approach (Hevner et al. 2004) that deals with the construction of scientific artefacts like methods, languages, models, or implementations. Following the Design Science approach, it is necessary to assure that the research addresses a relevant problem. This has to be proven. Furthermore, the artefacts to be constructed have to represent an innovative contribution to the existing knowledge base within the actual research discipline. I.e., similar or identical solutions must not be already available. Subsequent to the construction of the artefacts, these have to be evaluated in order to prove their fulfilment of the research goals.

In this contribution the scientific artefact is the modelling approach outlined in Section 1. This artefact aims at solving the relevant problem of the lacking comparability of conceptual models that are developed in a distributed way (cf. Section 1). Related work does not provide appropriate solutions up to now (cf. Section 2). Hence, the approach presented here (cf. Section 3) makes an innovative contribution to the existing knowledge base. In order to evaluate the approach, we have implemented a demonstrator software that shows the general applicability of the approach. Furthermore, we have conducted an explorative study that shows the potential of the approach (cf. Section 4). Additional evaluations concerning efficiency and acceptance issues will be performed in future studies (cf. Section 5).

3 A FRAMEWORK FOR THE SPECIFICATION AND ENFORCEMENT OF NAMING CONVENTIONS

3.1 Procedure Model

In order to provide a framework for naming conventions, we propose the usage of a domain language that is used for naming model elements in a certain modelling context (i.e. a specific modelling domain, project, or company). This domain language is a subset of the respective natural language (here: English) used in the modelling context. The domain language consists of a set of valid domain terms that are allowed to be used for model element names exclusively. The set of domain terms is a subset of all terms available in the respective natural language. Furthermore, every natural language has a certain syntax that determines the set of grammatically correct phrases. In our framework, we restrict the syntax of the respective natural language as well. I.e., the possibilities to construct sentences for model element names are limited. In summary, we restrict the grammar of a natural language in order to provide a formal basis for naming model elements (cf. Figure 1).
Natural language grammars are usually defined by a formalism that consists of a lexicon and a syntax specification (Mitkov 2003). This grammar is complemented with naming conventions, which again consist of term and phrase structure conventions. Term conventions are specified by a domain thesaurus containing domain terms with a precise specification of their synonym, homonym, and word formation relationships as well as a textual description of their meaning. It is connected to the natural language’s lexicon. Moreover, valid phrase structures are specified by phrase structure conventions. Hence, the natural language is customised for the needs of a specific modelling context. This allows for subsequent validation of the model element names and the enforcement of naming conventions. A conceptual overview of the naming conventions’ specification is given in Section 3.2.

The domain thesaurus can be created from scratch, or by reusing possibly existing thesauri or glossaries. Contrary to most common domain ontologies, which consist of nouns (e.g. “invoice”) or atomic domain concepts (e.g. “check invoice”), the thesaurus used here includes single nouns, verbs, and adjectives that are interrelated. Other word classes are generally domain independent. Thus, as they are already included in the lexicon, they do not need to be explicitly specified in the thesaurus. The terms in the thesaurus are linked to their synonyms, homonyms, and linguistic derivation in the lexicon. This additional term-related information can be obtained from linguistic services, which already exist for different natural languages. WordNet (http://wordnet.princeton.edu) is such a lexicon service for the English language providing an online interface. Therefore, in case of a later violation of the naming conventions by the modeller, synonymous valid terms can be automatically identified and recommended. The terms specified should be provided with short textual semantic descriptions, allowing modellers to look up the exact meaning of a term. The thesaurus should not be changed during a modelling project in order not to violate the consistency of application. When online services are used, the option of creating a local copy should be analysed.

Before starting a modelling project, the naming conventions have to be specified once for every modelling context. However, already existing conventions can be reused. Naming conventions are modelling language specific. For example, functions in Event-driven Process Chains (EPC) are labelled with activities (e.g. `<verb, imperative> + <noun>`; in particular e.g. “check invoice”) and events are labelled with states (e.g. `<noun> + <verb, past participle>`; in particular e.g. “invoice checked”) (Scheer 2000). For each model element type at least one phrase structure convention has to be defined. For the sake of applicability, the conventions should be specified in a manner, which is compatible with the formalism of the natural language grammar.

The conventions should be defined by a project team consisting of domain experts and modelling experts. I.e. the stakeholders responsible for the conventions should have thorough knowledge of the actual modelling context in order to reach a consensus. Most commonly, the thesaurus part of the conventions already exists in terms of corporate or domain-specific glossaries (e.g. http://www.automo-
During modelling, the model element names entered are verified simultaneously against the specified modelling context-specific grammar. On the one hand, the structure of an entered model element name is validated against the customised syntax specification. On the other hand, it is checked whether the used terms are allowed. Nouns, verbs, and adjectives are validated against the thesaurus. Other word classes are validated against the natural language lexicon.

In case of a positive validation, the entered model element name is declared as valid against the modelling context-specific grammar. In case of a violation of one or both criteria, alternative valid phrase structures and/or terms are suggested based on the user input. The modellers themselves have to decide, which of the recommendations fits their particular needs. By looking up the semantic descriptions of the terms, modellers can choose the appropriate one. Alternatively, they can choose a valid structure as a pattern and fill in the gaps with valid terms. However, the modeller should have the chance to propose a new term, which is then accepted temporarily. Hence, the modeller can continue without being distracted from his modelling session. It is then up to the modelling project expert team whether they accept the term or not. If the term is accepted, it is added to the thesaurus. Otherwise, the modeller is informed to revise the model element. Hereby, we ensure that equal model element names represent equal semantics, which is a precondition for comparability.

3.2 Conceptual Specification

In the following, we provide a conceptual framework for the specification and the enforcement of naming conventions using Entity-Relationship Models in (min,max)-notation (ISO 1982) (cf. Figure 2). Phrase structure conventions (PSC) are defined depending on distinct element types of conceptual modelling languages (e.g., activities in process models are named differently to events). Phrase structure conventions consist of phrase types or word types. A phrase type specifies the structure of a phrase, which can be used as a model element name. Therefore, a phrase type can be composed recursively of further phrase types or word types. Representing atomic elements of a phrase type, word types are acting as placeholders for particular words. An example of a word type is <noun, singular>, an example of a phrase type is <noun, singular> + <verb, infinitive>. The composition of phrase types is specified by the phrase type structure. At this, the allocation of sub-phrase types or word types to a phrase type and their position in the superordinated phrase type are defined. A word type consists of a distinct word class (noun, verb, adjective, adverb, article, pronoun, preposition, conjunction, or numeral) – and its inflection. Inflections are specialised as case, number, tense, gender, mood, person, and comparative, which are usually combined. E.g., a particular combined inflection is <3rd person, plural>. In respect to specific word classes, not every inflection is applicable. Based on the recursive composition of phrase types, the specification of arbitrary phrase structure conventions is possible.

Figure 2. Specification of Phrase Structure Conventions on Type Layer

Phrase structure conventions restrict the English syntax and thus limit modellers in their freedom of naming of model elements. In order to facilitate the synchronisation between the syntax of the natural
language and the applied phrase structure conventions, compatible formalisms for both syntax specifications are necessary. Hence, it should be possible to verify phrase structure conventions against the underlying natural language and to signal potential conflicts directly during the specification process. For this purpose, we establish the connection to an appropriate linguistic method in Section 3.3.

Independent from their corresponding word class, particular uninflected words are called lexemes (e.g. the verb “check”). Inflected words are called word forms (e.g. past participle “checked” of the lexeme “check”). Word forms are assigned to the corresponding word classes and inflections, i.e. their word types. Thus, word forms represent lexemes of a particular word type (cf. Figure 3).

![Figure 3. Specification of Term Conventions on Instance Layer](image)

In order to specify the domain thesaurus, allowed words are stored in the form of lexemes that are related by different word relationship types. They are specialised as homonym, synonym, and word formation relations. Word formation means that a lexeme originates from (an)other one(s) (e.g. the noun “control” originates from the verb “to control”). In case of synonym relations, one of the involved lexemes is marked as dominant to state that it is the valid one for the particular modelling context. Homonym relations are necessary in order to distinguish lexemes that consist of the same string but have a different meaning and to prevent errors during modelling. Word formation relations are used to search for appropriate alternatives when a modeller has used invalid terms and phrase structures. E.g., if the phrase “order clearance” violates the conventions, the alternative phrase “clear order” can be found via the word formation relation of “to clear” and “clearance”. Based on the word relationship types, lexical services (cf. Section 3.1) are connected to the domain thesaurus. To specify the actual meaning of a lexeme, a textual semantic description is added at least to each dominant lexeme. This way, modellers are enabled to check if the lexeme they have used actually fits the modelling issue.

### 3.3 Specification of Linguistic Restrictions

We make use of formalisms for the syntax specification of natural languages both for assuring correctness of specified phrase structure conventions and for identifying the structure and words of model element names. The latter is necessary in order to detect convention violations during modelling and to suppose alternative valid phrase structures and terms.

Such formalisms are mainly established in the field of computational linguistics. A well-known and widely accepted class of such formalisms are unification grammars (Mitkov 2003). All formalisms of this class have in common that syntax rules are specified in relation to phrase structures and their components. In unification grammars, feature structures are used to formalise syntax rules of a given language. A feature structure is a set of features and their values, which are structured hierarchically. It describes syntactical characteristics of a phrase structure. These phrase structures declare for a given natural language, which word types or sub phrase types can occur at which position within a phrase structure. An exemplary phrase structure consists of a phrase type, which is composed of two word types (cf. Figure 4). The word types are characterised by the features word class and inflection. Depending on the word class, the feature inflection is composed of further features. Possible features are number and case for the word class noun, and tense for the word class verb. The position of word types and phrase types within the phrase structure is defined by the given order. In the class of unifi-
cation grammars, the *Head-driven Phrase Structure Grammar (HPSG)* (Pollard & Sag 1994) is a widely established approach. Several formal specifications of natural languages based on HPSG are already available (cf. http://www.delph-in.net). A formal specification of the English syntax with HPSG was developed at the CSLI LinGo Lab of the University of Stanford (Copestake & Flickinger 2000).

In our approach, we use this HPSG-based syntax of the English language. Hence, the set of possible phrase structure conventions for model elements is a subset of the syntax of the English language based on HPSG. In the modelling process, the model element names have to comply with the conventions. Moreover, the used words have to comply with the domain thesaurus. In order to assure this, the model element names are parsed against the used words, their inflection, and their phrase structure. Corresponding parsing methods are available as a part of the *LinGo English Resource Grammar* (Copestake & Flickinger 2000). The parser is able to detect the lexeme, its word class and its inflection for each word using a lexicon. In our approach, the lexicon consists of the domain thesaurus that is connected to common lexical services. If the terms used within model element names do not comply with the conventions, alternative valid lexemes are searched in the domain thesaurus via the defined word relationships and are proposed in the appropriate inflection form for proper use. If a phrase structure is violated in turn, alternative valid phrase structures are proposed that contain the valid terms.

4 EVALUATION

4.1 Modelling Tool Support

To validate the applicability of our approach, we developed a modelling prototype. The way of navigating through the software and its handling corresponds with the procedures motivated in Section 3. Please note that the screenshots were adapted to English, since the prototype solely exists in German.
As a preliminary but unique step, the team responsible for constituting the domain grammar has to define the terms, which are allowed for the modelling context. Subsequently, the phrase structure conventions have to be specified. If the actual modelling context represents a domain which has been processed before, the existing set of terms and rules can be reused or adapted to the current requirements. Figure 5 shows an exemplary definition of domain verbs. It is sufficient to add the uninflected word, as the inflection can be looked up in the lexical services. Specifying nouns and adjectives is similar to specifying verbs. Word relationships like “synonym” or “word formation” can be defined between words. Besides the word relationships, the semantic descriptions are defined within the thesaurus editor as well.

![Figure 5. Specification of Domain Verbs](image)

Figure 5. Specification of Domain Verbs

In the next step, phrase structure conventions are defined and connected to those language elements for which they are valid. Figure 6 illustrates this with the example of EPCs. Here, phrase structure conventions are created for the model element type event. Concerning their semantics (but not their type), events of the EPC can be distinguished in trigger events and resulting events. The former ones trigger action in activities, the latter ones conclude activities. Different phrase structures can be attached to each of them in regard to their different semantics. An example for a trigger event is “Invoice is to be checked”, hence an appropriate phrase structure convention called “Trigger” is \(<\text{noun}> + \langle\text{"to be", simple present, 3rd person}\rangle + \langle\text{"to be", infinitive}\rangle + \langle\text{verb, past participle}\rangle\). With this phrase structure, a set of trigger events can be named. However, different aspects might require additional phrase structures to be defined. For result events, an adequate phrase structure is \(<\text{noun}> + \langle\text{verb, past participle}\rangle\), allowing phrases like “Invoice checked”.

Once generated, the phrase structure conventions in combination with the domain thesaurus are used during modelling. The modeller gets just-in-time hints as soon as he violates a convention (cf. Figure 7). First, he might have chosen invalid terms (e.g. bill instead of invoice or audit instead of check). As soon as a phrase is entered, it is parsed to determine its compliance with the naming conventions. Every term is transformed to its uninflected form and is compared to the domain thesaurus. If it is not found, synonymous valid terms are searched in the lexicon. If according alternatives are found, they are proposed to the modeller. Otherwise, the modeller has to rename the respective element – optionally by choosing a valid term from the domain thesaurus. Second, violations of phrase structure conventions are signalled, and alternative valid structures are proposed. Summarising, the name audit bill is suggested to be changed to Invoice is to be checked. Phrases corresponding with both the domain thesaurus and the phrase structure conventions are accepted without any feedback.
4.2 Explorative Analysis

Naming practices in process models provide evidence concerning the danger of naming conflicts as well as show the feasibility of approaches that aim at resolving or even avoiding them. In order to empirically evaluate our phrase structure conventions approach we conducted an exploratory empirical analysis. We analysed two process modelling projects consisting of overall 257 EPC models containing in turn overall 3918 elements (1827 functions and 2091 events). Within these modelling projects, modelling conventions were available in terms of glossaries and phrase structures. However, these conventions solely existed as textual recommendations rather than methodical support. All model element names were analysed and revised. We then manually analysed 280 phrase structure types, which occurred at least twice. These structures covered 1770 model element names in total. We were able to match the found structures to 47 identified phrase structure conventions, which had the same power of expression. E.g. we were able to express the same semantics concurrently achieving less naming variability. This analysis shows that having applied our approach would have substantially reduced naming differences and thus supported model comparability.

5 CONCLUSION AND OUTLOOK

Integrating naming conventions into conceptual modelling languages is promising for increasing the comparability of conceptual models. Two characteristics are significant to avoid common problems:

- Defining and providing naming conventions previously to modelling is the basis for avoiding naming conflicts rather than resolving them. Therefore, time-consuming alignment of namings in the course of model comparison becomes dispensable.
- Guiding the modeller automatically during modelling is of substantial importance, since only this way the compliance with the modelling conventions can be assured.

Certainly, specifying naming conventions in the proposed way is time-consuming. Our approach is therefore mainly suited for large-scaled modelling projects. Nevertheless, for every project, business domain or company, the conventions have to be specified only once and are reusable. Moreover, term models, thesauri, or glossaries that may already exist in companies or business domains can be reused.

Future research will focus on the further evaluation of the proposed approach. In the short-term, the approach will be instantiated for different modelling languages, different natural languages and different application scenarios. In particular, the capability of our approach to increase the efficiency of dis-
tributed conceptual modelling and its acceptance will be evaluated in empirical studies. In order to as-
sure the applicability of the approach, the demonstrator software will be enhanced in order to make it
usable in practice. In the course of evaluation, it will also be investigated if ambiguities play a role in
model element names. E.g. the sentence “They hit the man with a cane” is ambiguous, even if the
meanings of all of the used words are considered definite. Thus, we will analyse further conceptual
models and determine if phrase structures are common in conceptual modelling that promote ambi-
guities. A result of this analysis could be a recommendation to restrict phrase structure conventions to
phrases that do not lead to ambiguities.

Middle-term research will address further approaches in order to facilitate the comparison of models
that are developed in a distributed way. Here, the comparison of conceptual models based on the
model structure is a promising research area. Moreover, we will question if modelling conventions on
the basis of structure patterns that are provided at the beginning of modelling are able to increase the
comparability of conceptual models as well.

References

Model Integration or You Say Tomato, I Say Tomahito. ORSA Journal on Computing 3 (2),
pp. 107-120.
of Process Patterns in Engineering Domains. In: Proceedings of the Multi-Conference on Informa-
eling. In: Proceedings of the International Workshop on Human-Friendly Service Description, Dis-
covery and Matchmaking (HF-SDDM 2007) at the 8th International Conference on Web Informa-
tion Systems Engineering (WISE 2007). Editors: Weske, M., Hacid, M.-S. and Godart, C., Nancy,
Chen, P.P.-S. (1976). The Entity-Relationship Model: Toward a Unified View of Data. ACM Trans-
actions on Database Systems 1 (1), pp. 9-36.
coverage English grammar using HPSG. In: Proceedings of the 2nd Conference on Language Re-
Development and Dissemination of Reference Models. Enterprise Modelling and Information Sys-
tems Architectures 2 (2), pp. 32-41.
Greco, G., Guzzo, A., Pontieri, L. and Saccà, D. (2004). An ontology-driven process modeling frame-
work. In: Proceedings of the 15th International Conference on Database and Expert Systems Appli-
13-23.
quision 5 (2), pp. 199-220.


A NOVEL DATA QUALITY METRIC FOR TIMELINESS CONSIDERING SUPPLEMENTAL DATA

Journal: 17th European Conference on Information Systems
Manuscript ID: ECIS2009-0656.R1
Submission Type: Research Paper
Keyword: Data quality, Design research, Data Warehousing, Quality
A NOVEL DATA QUALITY METRIC FOR TIMELINESS CONSIDERING SUPPLEMENTAL DATA

Heinrich, Bernd, Department of Information Systems, University of Innsbruck, Universitätsstrasse 15, A-6020 Innsbruck, Austria, bernd.heinrich@uibk.ac.at

Klier, Mathias, Department of Information Systems, University of Innsbruck, Universitätsstrasse 15, A-6020 Innsbruck, Austria, mathias.klier@uibk.ac.at

Abstract

It is intensively discussed in both science and practice how data quality (DQ) can be assured and improved. The growing relevance of DQ has revealed the need for adequate metrics because quantifying DQ is essential for planning quality measures in an economic manner. This paper analyses how DQ can be quantified with respect to the DQ dimension timeliness. Based on an existing approach, we design a new metric to quantify timeliness in a well-founded manner that considers so-called supplemental data (supplemental data are additional data attributes that allow drawing conclusions about the timeliness of the data attribute considered). In addition, it is possible to take the values of the metric into account when calculating expected values, an advantage that in turn leads to improved and comprehensible decision support. We evaluate the presented metric briefly with regard to requirements for designing DQ metrics from literature. Then, we illustrate the metric’s applicability as well as its practical benefit. In cooperation with a financial services provider, the metric was applied in the field of customer valuation in order to support the measurement of customer lifetime values.

Keywords: Data Quality, Data Quality Metrics, Design Research, Customer Valuation
1 INTRODUCTION

Both the benefit and the acceptance of application systems depend heavily on the quality of data processed and provided by these systems (Ballou et al. 1999, Fisher et al. 2003). Executives and employees need high-quality data in order to perform business, innovation, and decision-making processes properly (Al-Hakim 2007, Even et al. 2007). This in mind, it is not surprising that insufficient data quality (DQ) may lead to wrong decisions and correspondingly high costs. According to an international survey on DQ, 75% of all respondents have already made wrong decisions due to incorrect or outdated data. In addition, the respondents and their staff spend up to 30% of their working time on checking the quality of data provided (Harris Interactive 2006). Therefore, ensuring completeness, correctness, and timeliness of data – these properties are known as DQ dimensions (Wang et al. 1995) – still remains an important problem for many companies (Ballou et al. 1998, Jiang et al. 2007). Non-surprisingly, many scientific papers (e.g. Ballou et al. 1998, Even et al. 2007, Lee et al. 2002, Parssian et al. 2004, Pipino et al. 2002, Wang 1998) deal with the question of how DQ can be quantified. This is essential for analysing the economic effects of poor or improved DQ as well as for realising DQ measures considering cost-benefit aspects (e.g. Heinrich et al. 2007a, Pipino et al. 2002).

In the following, we propose a metric for the DQ dimension timeliness. The reason is that – according to several studies – timeliness is a serious issue in DQ management (Klein et al. 2007, Yu et al. 2007). Therefore, this dimension has already been discussed from both a scientific and a practical point of view (e.g. Al-Hakim 2007, Klein et al. 2007, Knight et al. 2005, Lee et al. 2002, Wand et al. 1996).

Referring to the guidelines for conducting design science research defined by Hevner et al. (2004), we consider the metric for timeliness as our artifact and organize the paper as follows: After discussing the relevance of the problem, section 2 briefly compiles the related work regarding timeliness and identifies the research gap. Our contribution is a novel approach to quantify timeliness. Hence, a metric is proposed in section 3 which is based on probabilistic considerations. This metric enables to quantify timeliness in a well-founded manner and to consider so-called supplemental data (supplemental data are additional data attributes that allow drawing conclusions about the timeliness of the data attribute considered). In section 4, we illustrate the application of the new approach and its practical benefit by means of an extensive real world example in the field of customer valuation at a financial services provider. Section 5 summarizes our findings and critically reflects on the results.

2 RELATED WORK

In literature, there is a wide range of definitions with respect to the DQ dimension timeliness. In some publications, timeliness is also referred to as currency or recency. Table 1 contains some selected definitions.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Term and Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ballou et al. (1985), Ballou et al. (1998)</td>
<td>Timeliness: “the recorded value is not out of date […]. A stored value, or any data item, that has become outdated is in error in that it differs from the current (correct) value.”</td>
</tr>
<tr>
<td>Wang et al. (1996)</td>
<td>Timeliness: “The extent to which the age of the data is appropriate for the task at hand.”</td>
</tr>
<tr>
<td>Redman (1996)</td>
<td>Currency: “degree to which a datum in question is up-to-date. A datum value is up-to-date if it is correct in spite of possible discrepancies caused by time-related changes to the correct value.”</td>
</tr>
<tr>
<td>Hinrichs (2002)</td>
<td>Timeliness: “Property that the attributes or tuples respectively of a data product correspond to the current state of the discourse world, i.e. they are not out-dated” (own translation)</td>
</tr>
<tr>
<td>Price et al. (2005)</td>
<td>Timeliness: “The currency (age) of data is appropriate to their use”</td>
</tr>
<tr>
<td>Batini et al. (2006)</td>
<td>Timeliness: “Timeliness expresses how current data are for the task at hand.”</td>
</tr>
<tr>
<td>Heinrich et al. (2007a), Heinrich et al. (2007b)</td>
<td>Timeliness: &quot;Timeliness can be interpreted as the probability that an attribute value is still up-to-date”</td>
</tr>
</tbody>
</table>

Table 1. Selected definitions of the DQ dimension timeliness
The main issue of most definitions is that timeliness expresses whether an attribute value stored in a database is still up-to-date. This means that an attribute value, which was correct when it was stored, still corresponds to the current value of its real world counterpart at the (later) instant when DQ is quantified. In other words, the attribute value has not become outdated (due to its temporal decline). This is also reflected in the authors’ approaches to quantify timeliness. In contrast to the DQ dimension correctness, quantifying timeliness does not necessarily require a real world test. Instead, a metric for timeliness should provide an estimation, not a verified statement under certainty (which is necessary for correctness), on whether an attribute value is still up-to-date. Heinrich et al. (2007a, b) refer to this fact explicitly. They interpret timeliness as the probability that an attribute value is still up-to-date. For huge data sets and when the shelf life of attribute values is not explicitly known, it seems to be quite reasonable to quantify timeliness by means of such an estimation. This is because comparing attribute values to their real world counterparts (real world test) is often by far too time- and cost-intensive and not practical at all.

In this context the following questions arise: (1) How can well-founded estimations for the timeliness of attribute values be derived? (2) Avoiding real world tests, what kind of data can alternatively be used for quantifying timeliness? Some authors mentioned above consider so-called attribute metadata. Such metadata are the instant \( t_0 \) when the attribute value’s corresponding real world counterpart was created (e.g. for an attribute value “student” of a data attribute “professional status”: the instant of the student’s enrolment, at which the data value “student” became valid), and the attribute value’s shelf life \( T \) (e.g. for the attribute value “student”: the duration of study that represents how long this value is valid). Depending on whether these metadata are known or not, we have to quantify the timeliness of an attribute value under certainty or uncertainty. According to the definition of timeliness given above, we have to quantify whether an attribute value still corresponds to the current value of its real world counterpart at the instant \( t_1 \) of quantifying DQ. In other words, we have to analyse whether \((t_1-t_0) \leq T\) holds. If the instant \( t_0 \) when the attribute value’s corresponding real world counterpart was created and the attribute value’s shelf life \( T \) are known (e.g. based on a temporal rule), it is possible to determine whether the attribute value is still up-to-date under certainty. In this case the value of the metric equals zero (minimum value) if \((t_1-t_0)>T\) holds (attribute value is outdated). Otherwise, we can definitely state that the attribute value considered still corresponds to the current value of its real world counterpart \((t_1-t_0) \leq T\) and the value of the metric equals one (maximum value). In contrast to this scenario of temporal rules, we will focus in the following on quantifying timeliness under uncertainty. This case is much more interesting (in the scenario under certainty you have to check only if \(t_1-t_0) \leq T\) holds) and often more realistic because the shelf life \(T\) of an attribute value is usually unknown. Thus, the question arises of how to accomplish well-founded estimations for the timeliness of data when the shelf life of attribute values is not known.

To improve the estimations for the timeliness of data and to enhance existing approaches, we follow the idea to use additionally other data attribute values \( w_i \) \((i=1, \ldots, n)\) to draw conclusions about the timeliness of an attribute value \( \omega \) considered (e.g. about its unknown shelf life \( T\)). In the following, the values \(w_i\), which are used to improve the estimation, are called supplemental data. A short example illustrates their importance: Figure 1 (based on data from Federal Statistical Office of Germany 2007, Heublein et al. 2003, Heublein et al. 2008) shows that the duration of study (including dropouts) – i.e. the shelf life \( T\) of the attribute value “student” – and the type of university (university vs. university of applied sciences) are statistically contingent. Consider a number of customers stored in a database whose professional status “student” was for example stored 5.5 years ago at the beginning of their studies of economics and social sciences (age \( t \) of the attribute values: \((t_1-t_0)=5.5\) years=11 semesters): Then it is expected that in average about 90% of the customers enrolled at a university of applied sciences have already finished their studies (see Figure 1). This means that the attribute value “student” is up-to-date for only 10% of them. In contrast, it is expected that only 66% of the customers enrolled at a university have already finished their studies (see Figure 1) – i.e. the attribute value “student” is still up-to-date for about 34% of them in average.
Consequently, supplemental data – like the type of university – are relevant for quantifying the timeliness of data. Thus, it seems to be necessary to consider not only attribute metadata, but also supplemental data. Below, we discuss how existing approaches deal with this kind of data.

We found the approaches by Hinrichs (2002), Ballou et al. (1998), Even et al. (2007), and Heinrich et al. (2007a) as these are – to the best of our knowledge – the only approaches which (1) design metrics for timeliness, (2) are formally noted, and (3) are based for the most part on a Quality of Conformance definition. Regarding (3), literature distinguishes two different concepts and definitions of quality: Quality of Design and Quality of Conformance (Heinrich et al. 2003, Juran 1998, Teboul 1991). Quality of Design denotes the degree of correspondence between the users’ requirements and the specification of the information system. In contrast, Quality of Conformance represents the degree of correspondence between the specification and the existing realization in the information systems (e.g. correspondence between a data schemata and a set of stored data values). In the following, we focus on metrics for quantifying Quality of Conformance as they can be applied in many different situations and are more reusable (because they are more independent from particular users’ requirements in a specific business context). Heinrich et al. (2009) have already analysed the approaches mentioned above and compared them with respect to six requirements (due to space restrictions we can not describe this here): Normalisation, Interval scale, Interpretability, Aggregation, Adaptivity, and Feasibility. This analysis revealed weaknesses particularly with respect to Adaptivity, which requires a metric to be adaptable to the context of a particular application in order to enable a goal-oriented quantification of DQ (for a detailed analysis see (Heinrich et al. 2009)).

Referring to Adaptivity, supplemental data are valuable to improve the quantification of timeliness in many fields of application. Just consider the example mentioned above (see Figure 1). The existing metrics by Hinrichs (2002), Ballou et al. (1998), Even et al. (2007), and Heinrich et al. (2007a) use metadata like instant of creation \( t_0 \) and shelf life \( T \). However, they do not use supplemental data. That is, values of data attributes – as in our example type of university and course (field of study) –, which are relevant for quantifying timeliness and improving estimations, cannot be considered at all so far. Concluding, further research is needed to design a metric for timeliness that supports supplemental data in a methodically well-founded way.

3 DESIGNING A NOVEL METRIC FOR TIMELINESS

In the following, we take the metric for timeliness defined by Heinrich et al. (2007a) as starting point. We do so because the metric is based on probabilistic theory and its value can be interpreted as the probability that the considered attribute value \( \omega \) is still up-to-date. Assuming a finite shelf life for an attribute value, the value of the metric and consequently the probability decrease when the age \( t = t_f - t_0 \) increases (and vice versa). In contrast, an infinite shelf life would lead to a situation that an attribute value acquired once is always up-to-date and never changes, which refers to a quantification under certainty and is therefore quite simple.
We generalize the metric proposed by Heinrich et al. (2007a) as follows: The metric quantifies the timeliness of an attribute value \( \omega \), which is characterized by the corresponding real world counterpart’s instant of creation \( t_0 \). Together with the instant \( t_1 \) of quantifying DQ (with \( t_1 \geq t_0 \)), it is possible to determine the age \( t \) of the attribute value \( \omega: t = t_1 - t_0 \). The limited shelf life \( T \in \mathbb{R}^+ \) is unknown and therefore defined as a continuous random variable. Consequently, timeliness is defined as the probability that the shelf life \( T \) is greater than or equal to the age \( t \) of the attribute value \( \omega \). Given the probability distribution function \( F_\omega(t) := P_\omega(T \leq t) \) of the shelf life \( T \), we define the metric for timeliness as follows:

\[
Q_{\text{Time}}^\omega(t) := P_\omega(T \geq t) = 1 - P_\omega(T < t) = 1 - F_\omega(t)
\]  

(1)

In the particular case of an exponential distribution, which is a typical distribution for lifetime and has already proven to be useful in quality management (especially for address data etc.), Heinrich et al. (2007a) define the metric as shown in (2). Assuming that the attribute value \( \omega \) is correct at the instant \( t_0 \) of its acquisition, we may use this instant \( t_0 \) to calculate the age \( t \) instead of the corresponding real world counterpart’s instant \( t_0 \) of creation. This is because the exponential distribution is memoryless in the following way: \( P(X \geq x + t | X \geq x) = P(X \geq t) \); i.e. the conditional probability that the attribute value becomes outdated in the next period of time is independent of its current age.

\[
Q_{\text{Time}}^\omega(t) := \exp(-\text{decline}(A) \cdot t)
\]  

(2)

The parameter \( \text{decline}(A) \) is the decline rate indicating how many of the attribute’s values become outdated on average within one period of time. For example, a value of \( \text{decline}(A)=0.2 \) has to be interpreted like this: on average 20% of the attribute’s values lose their validity within one period of time. Obviously, the definitions (1) and (2) do not support the use of supplemental data \( w_i \) \((i = 1, \ldots, n)\). Therefore, values of data attributes like type of university and course cannot be considered at all when quantifying timeliness of the attribute value “student”, for example.

To solve this problem, we developed the following idea: We redefine the metric for timeliness in term (1) to represent the conditional probability that the considered attribute value \( \omega \) is still up-to-date. Using the supplemental data \( w_i \) as conditions \( W_1 = w_1, \ldots, W_n = w_n \) when calculating the probability is a well-founded way to consider them. The values of the variables \( W_i \) (i.e. \( w_i \)) are known (they are stored in the database) and thus need not be modelled as random variables. However, they usually are subject to temporal decline as well. Hence, it is advantageous to model them – without loss of generality - as random variables, too. Given the distribution function \( F_\omega(t|W_1, ..., W_n) := P_\omega(T \leq t | W_1 = w_1, ..., W_n = w_n) \) with the supplemental data \( w_i \) \((i = 1, \ldots, n)\) we define the new metric for timeliness \( Q_{\text{Time}}^\omega(t, w_1, ..., w_n) \) as follows:

\[
Q_{\text{Time}}^\omega(t, w_1, ..., w_n) := P_\omega(T \geq t | W_1 = w_1, ..., W_n = w_n) = 1 - P_\omega(T < t | W_1 = w_1, ..., W_n = w_n)
\]

\[
= 1 - F_\omega(t | w_1, ..., w_n) = 1 - \int_0^t f_\omega(\theta | w_1, ..., w_n) d\theta
\]  

(3)

The conditional probability (=value of the metric) is calculated based on the complementary probability \( P_\omega(T < t | W_1 = w_1, ..., W_n = w_n) \) – which represents the probability that the attribute value is outdated at the instant \( t \) of quantifying DQ – and the distribution function \( F_\omega(t|w_1, ..., w_n) \). Thereby, the conditional distribution function is defined as the integral over the conditional probability density function \( f_\omega(\theta | w_1, ..., w_n) \). This function, in turn, is determined by the quotient of the combined probability density functions \( f_\omega(\theta, w_1, ..., w_n) \) and \( f_\omega(w_1, ..., w_n) \). As the complementary probability represents whether the attribute value \( \omega \) is outdated before the age \( t \) is reached, the definite integral is calculated for the interval \([0, t]\). This in mind, we can calculate the probability in our example that the stored attribute value “student” is still up-to-date for a certain customer considering supplemental data.
Before, we briefly evaluate whether the novel metric meets the requirements defined by Heinrich et al. (2007b).

The definition as a conditional probability ensures that the values of the novel metric are normalized to \([0; 1]\). Moreover, the metric is equal to one for attribute values with age \(t=0\) (i.e. \(t_1=t_0\): 
\[
Q_{\text{time}}(0, w_1, ..., w_n) = 1.
\]
This is reasonable because the attribute value \(\omega\) is correct at the corresponding real world counterpart’s instant \(t_0\) of creation (see definition of timeliness). Moreover, the values of the metric are limited to zero – due to their limited shelf life \(T\), the following equation holds:
\[
\lim_{t \to \infty} Q_{\text{time}}(t, w_1, ..., w_n) = 1 - \lim_{t \to \infty} F_{\omega}(t | w_1, ..., w_n) = 0.
\]
Thereby, a value of zero means, that the attribute value \(\omega\) is certainly outdated. Based on probability theory, the values of the metric are interval scaled and interpretable (as a probability). Moreover, the aggregation formulas defined by Heinrich et al. (2007b) can be applied as well. As mentioned earlier, the timeliness of an attribute value (e.g. professional status of a particular customer) can be calculated automatically to a large extent by using the formula above as well as SQL DML statements. This ensures that the metric meets Feasibility. The weighting factors in the aggregation formulas and designing the metric according to the shelf life of the attribute values support Adaptivity. However, the Adaptivity of the novel metric could be further improved – related to the metric proposed by Heinrich et al. (2007a) – by integrating supplemental data. In the next section, we illustrate this advantage by an example of a financial services provider.

## 4 PRACTICAL EVALUATION OF THE METRIC

In this section, we evaluate our metric for timeliness by means of a real use situation. Thereby, we analyse its applicability and practical benefit. We applied the metric at a German financial services provider (FSP) and especially focus on the question of how it can support the process of customer valuation. In the context of customer valuation, the customer lifetime value (CLV) is a widely accepted approach to valuate customers and is defined as the present value of all existing and future cash flows generated with a certain customer (cp. e.g. Berger et al. 1998). It is obvious that a high CLV has to be assessed very critically, when the input data used for calculation are (possibly) outdated. Thus, quantifying DQ is an important issue in this context. Due to confidentiality, all data are anonymised and modified. But the principal results still hold.

The FSP acts as an independent company aiming at advising its customers (mostly academics) holistically during a large part of their life cycle. Thus, the CLV is a starting point for many decisions and plays an important role. CLV calculation is based on several input data like customer’s age, current professional status (e.g. student) and course (e.g. engineering sciences). Such data attribute values – which were often acquired many years ago - are usually stored in the customer database.

In order to calculate CLVs, the FSP assumes that every customer passes through different phases of a typical customer life cycle. This customer life cycle starts at the instant of career entry because at this instant the FSP can start selling various products and services. The FSP then tries to determine a customer’s (typical) demand for each of these phases (e.g. retirement). On this base, it is possible to estimate the cash flows resulting from selling products and services in each phase and to calculate the CLV as an expected net present value. In this way, the cash flows and CLVs have already been quantified for several target groups during a former project. The results are a starting point for our example.

However, the FSP also focuses on students in order to acquire them before they become career starters (then their demand is usually especially high). Therefore, the FSP has a very large number of customers being students. These are stored in the database with professional status “student”, including course and instant of enrolment. But how should the CLVs of these customers be quantified? As most of the products cannot be sold to students, the staff of the FSP usually contact these customers quite sporadically (this finally leads to marginal, negative cash flows in this phase). As a result, a student’s instant of career entry is typically not known by the FSP. Nevertheless, it would be much too time- and labour-intensive to contact each customer in order to verify the correctness of the stored professional status.
sional status “student” (real world test). Therefore, it is necessary to quantify timeliness at a high level of automation and to avoid such real world tests. In the following, we describe two customer valuation procedures that were previously used by the FSP. Then, we discuss the application of the metric for timeliness.

According to the first procedure, the FSP calculated the CLV of customers with professional status “student” as follows: The attribute value for instant of enrolment was used to estimate the remaining duration of study by comparing the duration since the instant of enrolment with the average duration of study (in semesters). The latter was determined by using publicly available data about graduates and dropouts (Federal Statistical Office of Germany 2007, Heublein et al. 2003, Heublein et al. 2008): Accordingly, 64% of all students who finally graduated studied about 15.5 semesters on average, while the remaining 36% dropped out after 5.6 semesters on average. The FSP calculated an average duration of study of about 12 semesters (≈0.64*15.5 semesters+0.36*5.6 semesters). For this estimated remaining duration of study, the FSP calculated negative cash flows (note: as the FSP did not put DQ into question, the remaining duration of study was estimated with one semester, if the elapsed time since instant of enrolment was larger than or equal to the average duration of study). Starting from the instant of career entry, the FSP considered the CLV related to its life cycle approach. We briefly illustrate this simple procedure: Considering a customer, who enrolled in economics and social sciences in April 2003, a remaining duration of study of about 2 semesters is assumed by the FSP at the instant of customer valuation in April 2008 (computed by (average duration of study)-(duration since instant of enrolment)=12-10=2). For this period of time negative cash flows (here: € -150 per semester) were calculated. On the one hand, the CLV of a graduate in economics and social sciences is € 4,000 (net present value according to the lifecycle approach). On the other hand, the CLV of a dropout is only about € 2,000. This leads to a weighted average CLV of 0.64*€ 4,000+0.36*€ 2,000=€ 3,280. Table 2 summarizes the results of this procedure:

<table>
<thead>
<tr>
<th>Instant of enrolment</th>
<th>Instant of customer valuation</th>
<th>Duration since enrolment</th>
<th>Remaining duration of study (est.)</th>
<th>Cash flows of students (est.)</th>
<th>CLV at the instant of career entry (est.)</th>
<th>Calculated CLV at the instant of customer valuation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr. 2003</td>
<td>Apr. 2008</td>
<td>10 semesters</td>
<td>2 semesters</td>
<td>€ -150/semester</td>
<td>€ 3,280</td>
<td>€ 2,554</td>
</tr>
</tbody>
</table>

Table 2. Customer valuation in the example (first procedure)

The example illustrates that the FSP (with a discount rate of 5% per semester) calculated a CLV of \[ \sum_{i=1}^{2} \frac{-150}{(1+0.05)^i} + \frac{3280}{(1+0.05)^3} \approx € 2,554 \] (net present value of the expected cash flows), which is quite low compared to the average CLV at the instant of career entry. The result is based on the implicit assumption that the customer will certainly be studying another two semesters before starting his/her career. All in all, this procedure is quite simple, but does by no means consider DQ aspects.

Therefore, the FSP began to modify the procedure and extended it by probabilistic considerations. First, the assumption of a fixed duration of study (of 12 semesters) for all students was avoided. Instead, a customer’s professional status “student” may remain or change (to “career entry/graduate” or “career entry/dropout”) in each period (semester). All three possible transitions are assigned with probabilities, considering the customer’s estimated remaining duration of study. By using this procedure, it is possible that the instant of career entry is before or after the average duration of study. Figure 2 illustrates the modified procedure, which is to large parts based on the concept of homogeneous Markov chains (cp. e.g. Pfeifer et al. 2000). Note: This concept implies that the transition probabilities (e.g. \( p_{stu.,stu.} \)) are identical for all periods considered (e.g. \( p_{stu.,stu.}(i)=p_{stu.,stu.}(i+1) \) for each period \( i \)). The FSP calculated the probability \( p_{stu.,stu.} \) for each customer so that the expected number of semesters, for
which a customer remains in professional status “student” \( \left( \sum_{i=1}^{\infty} i \cdot p_{\text{Stu},\text{Stu}} = \frac{p_{\text{Stu},\text{Stu}}}{(1 - p_{\text{Stu},\text{Stu}})^2} \right) \) was equal to the estimated remaining duration of study. Using this probability-based procedure and calculating the expected value to determine the CLV, it is considered that a customer may finish sooner or later than the average duration of study (variation).

![Diagram](image.png)

**Figure 2.** Possible transitions of the professional status “student” (modified procedure)

Taking into account the customer data in Table 2, the transition probability \( p_{\text{Stu},\text{Stu}} \) was determined to 50\% (this goes along with an expected remaining duration of study of 2 semesters). Interpreting this probability means that 50 out of 100 customers studying in the 10\(^{th}\) semester still remain students in the next semester and so on. The other 50 customers finish after the 10\(^{th}\) semester. They were separated – based on the shares mentioned above – in 32 (\( \approx 0.64 \times 50 \)) graduates and 18 (\( \approx 0.36 \times 50 \)) dropouts to determine \( p_{\text{Stu},\text{Grad}} \) and \( p_{\text{Stu},\text{Drop}} \). Summing up, a CLV of € 2,845 was determined by means of the modified procedure:

\[
\sum_{i=1}^{\infty} \left( \frac{0.50 \cdot \varepsilon - 150 + 0.32 \cdot \varepsilon \cdot 4,000 + 0.18 \cdot \varepsilon \cdot 2,000}{1 + 0.05} \right)^i
\]

Though avoiding the assumption of a uniform fixed duration of study (of 12 semesters), the modified procedure does not address DQ issues of input data yet. What is particularly noteworthy is the underlying assumption that each customer with the attribute value “student” is still a student at the instant of customer valuation (in the example in April 2008). Both procedures are based on this assumption. Therefore, the FSP ignored that the stored customer data could already be outdated at the instant of customer valuation. Exactly this was observed by the FSP when some customers were asked in summer 2008 with respect to their professional status (cp. ex post analysis below). As a consequence, we analysed how the designed metric for timeliness can be used to meet this issue. The basic idea is as follows:

Our metric for timeliness represents the probability that a customer with the attribute value “student” is still a student in the real world at the instant of customer valuation. This in mind, the values of the metric are appropriate for calculating the transition probabilities \( p_{\text{Stu},\text{Stu}}, p_{\text{Stu},\text{Grad}}, \) and \( p_{\text{Stu},\text{Drop}} \). Neglecting the limiting assumption of both former procedures, we first calculated the probability \( Q_{\text{Time}}(t,0) = Q_{\text{Time}}(t) \) that a customer is still studying at the instant of customer valuation (April 2008 = period 0). We did not consider supplemental data in the first step. The value of the metric had to be calculated at the instant of customer valuation considering the probabilities for graduation \( P_{\text{Grad}}(0) \) as well as for dropout \( P_{\text{Drop}}(0) \) up to that point in time: \( Q_{\text{Time}}(t,0) = 1 - (P_{\text{Grad}}(t,0) + P_{\text{Drop}}(t,0)) \). We similarly determined the transition probabilities for the following periods (semester). Since conditional probabilities were needed, we calculated \( p_{\text{Stu},\text{Stu}}(t,i) \) for each period \( i \) with \( i \geq 1 \) as follows:

\[
p_{\text{Stu},\text{Stu}}(t,i) = \frac{Q_{\text{Time}}(t,i)}{Q_{\text{Time}}(t,i-1)}
\]

In the next step, we determined the transition probabilities \( p_{\text{Stu},\text{Grad}}(t,i) \) and \( p_{\text{Stu},\text{Drop}}(t,i) \) taking into account the values of the metric, too. They represent conditional probabili-
ties for graduation and dropout respectively regarding period $i$

$$p_{\text{Stu,Grad}.(t,i)} = \frac{P_{\text{Grad}.(t,i)} - P_{\text{Grad}.(t,i-1)}}{Q_{\text{Time}.(t,i-1)}}$$

$$p_{\text{Stu,Drop}.(t,i)} = \frac{P_{\text{Drop}.(t,i)} - P_{\text{Drop}.(t,i-1)}}{Q_{\text{Time}.(t,i-1)}}$$

Summing up, we get $p_{\text{Stu,Stu}.(t,i)} + p_{\text{Stu,Grad}.(t,i)} + p_{\text{Stu,Drop}.(t,i)} = 1$ for each period $i$ (equivalent to $Q_{\text{Time}.(t,i)} = 1 - (P_{\text{Grad}.(t,i)} + P_{\text{Drop}.(t,i)})$).

Figure 3 illustrates the procedure considering the metric:
Considering dropouts for example, the cumulative relative frequencies by Heublein et al. (2008) for each type of university and each course can be used. Figure 4 shows the cumulative relative frequency distribution of dropouts (in relation to all dropouts) for economics and social sciences at universities. For example, approx. 25% of all dropouts occur within the first two semesters. All in all, the figure shows that the dropout rates with respect to all students of a semester are approx. constant (contrary to the absolute number of dropouts which is obviously decreasing). Hence, we can assume a constant relative decline rate and apply an exponential distribution. Using a least square estimation, we determined an exponential distribution with a decline rate of 0.165 (see Figure 4). The value of the coefficient of determination $R^2$ was calculated to 0.98, which shows a very good approximation. Therefore, $P_{\text{Drop}}(x)=0.19 \cdot (1-\exp(-0.165 \cdot x))$ denotes the probability that a student has already aborted his/her studies after $x$ semesters. Here, the factor 0.19 corresponds to the fraction of dropouts in relation to all students who have enrolled at universities in economics and social sciences.

![Figure 4. Cumulative relative frequency distribution of dropouts](image)

If we compare this result to the function $0.05 \cdot (1-\exp(-0.145 \cdot x))$, which was determined equally for students enrolled at universities in medical sciences, significant differences become apparent. Considering the first five semesters (i.e., $x=5$) for example, the probability $P_{\text{Drop}}(x)$ for economics and social sciences is 10.7% in contrast to just 2.6% for medical sciences. Such differences completely depend on the different supplemental data with respect to the data attribute course. We similarly computed the probability $P_{\text{Grad}}(x)$, which represents the cumulative probability that a customer with professional status “student” graduated after $x$ semesters. Here, it was necessary to assume a Weibull distribution (for this distribution see Heinrich et al. 2009). Then, we defined the metric $Q_{\text{Time}}^\text{grad}(t, w_1, \ldots, w_n)$ based on both probabilities and calculated its values for each customer and period automatically.

Table 3 lists some selected customers with the supplemental data (type of university and course) from the FSP’s customer database as well as the corresponding metric values $Q_{\text{Time}}^\text{grad}(t, w_1, \ldots, w_n)$.

<table>
<thead>
<tr>
<th>Customer</th>
<th>Professional status</th>
<th>Instant of enrolment</th>
<th>Type of university</th>
<th>Course</th>
<th>$Q_{\text{Time}}^\text{grad}(t, w_1, \ldots, w_n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>“Student”</td>
<td>Apr. 2004</td>
<td>University</td>
<td>Engineering sciences</td>
<td>0.58</td>
</tr>
<tr>
<td>B</td>
<td>“Student”</td>
<td>Apr. 2004</td>
<td>University of applied sciences</td>
<td>Engineering sciences</td>
<td>0.24</td>
</tr>
<tr>
<td>C</td>
<td>“Student”</td>
<td>Apr. 2004</td>
<td>University</td>
<td>Medical sciences</td>
<td>0.92</td>
</tr>
<tr>
<td>D</td>
<td>“Student”</td>
<td>Apr. 2004</td>
<td>University</td>
<td>Economics and social sciences</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Table 3. Supplemental data for selected customers and corresponding value of the metric
The differences between the values of the metric highlight the impact of supplemental data (since all other data are the same). For our example customer (see above), we get the following results when applying the procedure based on the new metric: The probability that the customer, who enrolled at a university in economics and social sciences in April 2004, is still a student at the instant of customer valuation (April 2008) is only 46%. In contrast, the probabilities that he/she has already graduated or aborted were calculated to 38% and 16% respectively. When using the previous procedures, it was not possible to determine such probabilities at all. Instead, it was assumed that the customer is still studying in April 2008. A further advantage of the metric is that we do not need to assume that the transition probabilities are constant over time (this is unrealistic). But, it is possible to determine customer-specific transition probabilities for each period automatically. Applying the new metric within the example, we got a CLV of €3,173 which is more than before due to 1), 2) and 3) (cp. above).

The importance of supplemental data can also be demonstrated by an ex post analysis. After the valuation in April 2008, the FSP instructed its sales staff to ask about the current professional status of customers who were stored as “student” in the database. For all customers who had already graduated or dropped out the staff had to acquire the instant of graduation or dropout. Otherwise professional status “student” was confirmed. Similar information was requested by the FSP in a campaign starting in May 2008. The FSP wanted to know from customers with professional status “student” whether and when their status had changed. All in all, the FSP and its staff changed the attribute value “student” for 1,510 customers until the end of August 2008. We analysed these customers by comparing their actual instant of graduation or dropout with the results and estimations of each procedure accomplished. According to the first and simple procedure the FSP assumed an average duration of study of 12 semesters. Thus, for each of the 1,510 customers we could determine when he/she would have been expected to finish his/her studies. Comparing these instants with the actual semester of graduation or dropout, we found that these conformed for only 130 out of 1,510 customers. In other words, in 91.4% of all cases, the estimation was actually incorrect. We also analysed the other probability-based procedures. For every instant of enrolment, we determined the corresponding number of students. On this basis, we calculated how many students would have been expected to graduate and dropout in each of the following semesters using the transition probabilities (see Figures 2 and 3). An example illustrates this: 157 customers out of all 1,510 customers enrolled in October 2002. With the transition probabilities \( p_{\text{Stu.,Stu.}}(0,1)=0.86, p_{\text{Stu.,Grad.}}(0,1)=0, \) and \( p_{\text{Stu.,Drop.}}(0,1)=0.14, \) we get no customer who expectedly graduates in the first semester (until Feb. 2003), 22 customers who dropout and 135 customers who continue their studies. We did such calculations for each probability-based procedure and compared the expected numbers with the actual numbers acquired by the FSP. This way we calculated the difference between the expected frequency distribution of each procedure and the actual frequency distribution of all 1,510 selected customers (number of faults=1,510-number of customers where actual and expected semester corresponded).

The results of the ex post analysis were the following: The second procedure, which does not consider DQ issues, had 1,136 faults (75.2%). For the third procedure, which includes the metric for timeliness without supplemental data, these faults could be reduced to 892 (59.1%). Finally, by using the last procedure, which is based on the novel metric for timeliness considering supplemental data (type of university and course), these results could be further improved to 710 faults (47.0%). Table 4 summarizes the findings of the ex post analysis (the number of faults is quite high for all procedures because we count every “minor” difference in terms of one semester as a fault):

<table>
<thead>
<tr>
<th>Number of analysed customers</th>
<th>Number of faults first procedure without DQ</th>
<th>Number of faults second procedure without DQ</th>
<th>Number of faults existing DQ metric</th>
<th>Number of faults novel DQ metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,510 customers (100.0%)</td>
<td>1,380 customers (91.4%)</td>
<td>1,136 customers (75.2%)</td>
<td>892 customers (59.1%)</td>
<td>710 customers (47.0%)</td>
</tr>
</tbody>
</table>

*Table 4. Ex post analysis*
The analysis shows that using the metric for timeliness and considering supplemental data obviously improve the results. Focusing on DQ issues, we did not evaluate the CLVs, but the metrics for timeliness. Here, using the novel metric (instead of existing approaches) allows estimating transition probabilities better and therefore creates practical benefit to a significant extent.

Summing up, it can be stated that the example was intentionally kept simple in order to illustrate the practical benefit. It is possible to represent much more difficult issues considering further statuses besides “student”, “career entry/graduate”, and “career entry/dropout”. For customers whose professional status was acquired a long time ago other statuses seem to be reasonable (e.g. “junior consultant” or “senior consultant” instead of “career entry/graduate”). In addition, the new metric is valuable for determining the instant of career entry because such information can be used to contact customers in a more goal-oriented way. Figures like average duration of study are bad estimates here, especially considering good students. These students who are often characterised by a short duration of study would be contacted too late, though being more attractive for the FSP in many cases. Nevertheless, the simple example illustrates that quantifying DQ helps to improve customer valuation.

5 SUMMARY AND CONCLUSIONS

In this paper, we presented a novel metric for timeliness that is based on probabilistic considerations. Extending an existing approach, we take supplemental data into account and define the metric as the conditional probability that a considered data value is still up-to-date at the instant of quantifying DQ. Hence, quantifying timeliness can be done in a more accurate and methodically well-founded way. The metric’s practical benefits as well as its applicability were illustrated by a real use situation. In contrast to previous approaches, the new metric allows us to consider supplemental data. Here we work with conditional probabilities using supplemental data as conditions of the probability density function. The example illustrates the importance of supplemental data like type of university and course. Moreover, when designing the metric, we were able to avoid limiting assumptions concerning the probability distribution (a number of possible probability distributions is presented by Heinrich et al. 2009). This assures that the metric is appropriate for many attributes and their individual characteristics (such as constant, increasing or decreasing decline rates). In practice, the metric is valuable for calculating decision variables like the CLV. If companies rely on the CLV in order to manage customer relationships, outdated customer data may result in wrong decisions. The same holds for other fields of application (e.g. mobile services provider (cp. Heinrich et al. 2007a, b, 2009), financial services provider (cp. Heinrich et al. 2008), where such metrics for Quality of Conformance could be applied successfully.

Besides these findings, calculating the probability distribution function can be difficult in some cases: On the one hand, publicly available data (e.g. from Federal Statistical Offices or scientific institutions) can be applied to define the metric. On the other hand, internal data (e.g. from the data warehouse) may be analysed using statistical software such as SPSS to derive the probability distribution function. Moreover, interviews (as the FSP from the example did) and experts’ estimations are further instruments. However, quantifying correctness by means of a real world test for every single attribute value – which is an alternative to quantifying timeliness – is usually much more cost-intensive. Additionally, it has to be considered that a metric, which was developed once, can be reused frequently or adapted to several fields of application. The authors are working currently on a model-based economic approach for planning DQ measures. For implementing such a model, adequate DQ metrics are necessary. The approach presented here provides a basis for those purposes. Nevertheless, further metrics for other DQ dimensions should be developed and further research in this area is encouraged.

References

Due to space restrictions, we provide the list of references on http://www.uibk.ac.at/iwi2/.
**STORYTELLING AS A TOOL FOR KNOWLEDGE TRANSFER IN THE IT INDUSTRY**

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0630.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Culture fit / differences / heritage/ intelligence / issues / theory / values, Cross-cultural issues, knowledge transfer, Offshoring / Outsourcing</td>
</tr>
</tbody>
</table>
STORYTELLING AS A TOOL FOR KNOWLEDGE TRANSFER IN THE IT INDUSTRY

Wende, Erik, University of Zurich, Department of Informatics, Binzmühlstr. 14, CH-8050 Zurich, Switzerland, wende@ifi.uzh.ch

Dr. Haghirian, Parissa, Shophia University, 7-1 KIOI-CHO, CHIYODA-KU, 102-8554 Tokyo, Japan, p-haghir@sophia.ac.jp

Abstract

As organizations increase their offshore software development efforts, they must develop new methods and models for handling the vast amount of knowledge involved in these projects. Successful knowledge management and transfer is considered key to the success of contemporary organizations. When transferring knowledge to other operating units of a multi national company, the overall goal is to successfully implement the knowledge sent to the receiver. Cultural differences however, can interfere with successful knowledge management intentions. This paper investigates storytelling as a tool to transfer knowledge between global corporate units. A case study on how this instrument is used to communicate knowledge between a German and an Indian IT company gives first insights into factors that influence implementation success.

Keywords: Knowledge Management, Storytelling, Case Study, Culture.
INTRODUCTION

Globalization has played a major role in developing businesses processes. Modern technology supports these transformations and allows members of global organizations to be involved in increasingly international cooperation. In the IT industry, more and more software development projects have been geographically distributed and happen in many different countries at the same time. Offshore software development has thus become a notable area of focus in the IT industry. Nowadays, we often speak about global or offshore software development when it comes to globally distributed development teams working together in different time zones, with different local language settings, with different cultural backgrounds and a different educational approach to software engineering.

1 THE NEED FOR KNOWLEDGE TRANSFER

As organizations increase their offshore software development efforts, they must develop new methods and models for handling the vast amount of knowledge involved in these projects (Desouza et al., 2006). ‘Knowledge management’ and ‘knowledge transfer’ become highly prominent in this scenario. ‘Knowledge management’ has many definitions, one of them being the process of continuously creating new knowledge, disseminating it widely through the organisation, and embodying it quickly in new products/services, technologies and systems (Takeuchi and Nonaka, 2004). ‘Knowledge transfer’ is basically giving background information on software projects to people who do not have it (Stellman and Greene, 2005).

The concept of knowledge transfer is difficult to capture, because there is no clear distinction between the transfer of knowledge and the creation of new knowledge (Bresman et al., 1999). It is customary to speak of the ‘transfer’ of knowledge between two distant units of a multi national company (MNC) or between two different functional units at the headquarters, between a vendor and a customer, or even between countries. The use of ‘transfer’ implies flow: knowledge ‘flows’ from its primary holder to the receiver (Doz and Santos, 1997).

Knowledge flows or knowledge transfer refers to the transfer of either expertise or external market information of global relevance, but not to the transfer of internal administrative information (Gupta and Govindarajan, 1991). Transferring knowledge means the transferring of operational knowledge. This can happen in the form of data, information, blueprints, parts, subassemblies, machines or other means to represent knowledge. It can also happen via persons, individual or teams (Doz and Santos, 1997).

Knowledge flows and knowledge transfers are strategically important to organizations for several reasons. They transmit localised know-how, which is generated in one sub-unit to other locations in the organization. Knowledge transfers also facilitate the co-ordination of work flows linking multiple, geographically dispersed sub-units. Furthermore they can enable organizations to capitalise on business opportunities requiring the collaboration of several sub-units. Knowledge flows are also crucial to the orchestrated execution of unified strategic responses to moves of competitors, customers, and suppliers. Finally knowledge flows enable the recognition and exploitation of economies of scale and scope (Schulz and Jobe, 2001).

The MNC faces various challenges with regard to internal knowledge transfer. Subsidiaries should be motivated to access and produce knowledge within the MNC, which means that relevant subsidiary knowledge has to be made accessible to those MNC units that need it. To do so communication needs to be established between those who need and those who possess knowledge. To achieve this goal the organization has to choose the best instruments of control, motivation and context (Foss and Pedersen, 2002).
Successfully identifying, analysing, specifying and documenting better requirements are very crucial; it becomes a higher priority in terms of its effectual transfer across boundaries in offshore software development cases. Differences in location specific work cultures like work ethic, importance of hierarchy and mode of communication can impact the transfer of the software requirements specifications.

When cautiously considering the inherent risks of globally distributed development (Aspray et al., 2007), co-ordination and communication issues are the most intense burdens compared to distance and time (Herbsleb, 2007). Moreover, issues on data and system security, contractual and intellectual property issues as well as concerns about losing domain knowledge play an important role (Carmel and Tija, 2005). But despite those risks, reasons to offshore are still persuasive with cost advantages as the dominant force (Carmel and Tija, 2005). Offshore strategies are further utilised to gain access to enormous skilled labour pools with a certain domain experience and to exploit time shift advantages by expanding the daily development cycle to different time zones.

Conveying this knowledge to counterparts working in a geographically distant, culturally differing country becomes an important issue to focus on. During such scenarios a lot of factors come into prominence that should be well taken care of. The crucial challenges are ‘knowledge transfer’ and ‘cultural’; it is important to inspect these problems and come up with a feasible solution in each case.

2 CROSS-CULTURAL KNOWLEDGE TRANSFER

Culture plays an important role in any team activity’s success (Bhat et al., 2006) and is associated with the knowledge transfer process. Understanding and dealing with the culture of the vendor country for the efficient transfer of specifications is one of the motivations for our research.

Knowledge management literature gives the impression that knowledge management operates in a kind of cultural vacuum. Diversity in terms of language, cultural and ethical background, gender and professional affiliation are considered to be one independent variable, which is in any case pushed to the side. This approach may be convenient for conceptualizing, but is very limited for practical purposes in the modern international business world (Holden, 2002). Hofstede (Hofstede, 1984) furthermore points out that geographical separation and cultural differences can lead to quasi-autonomous sub-organizations which may further lead to numerous problems of communication, co-ordination, control and motivation. Thus cultural differences within an MNC should not be neglected when discussing knowledge transfer and can be regarded as one of the barriers between company divisions and local units of the company (von Krogh et al., 2000, Davenport and Prusak, 1998). Knowledge transfer within units located in the same country can already be troublesome, but it is clear that the problem associated with transfer increases with geographical and cultural distance (Bresman et al., 1999). Li (Li, 1999) shows that communication between individuals in high-context countries and low-context countries differs significantly in the amount of information transferred.

Within knowledge transfer relationships between members of differing cultures interlocutors communicated less information than between members of the same cultural background. His results indicate that low-context/low-context communication relationships do not differ from high-context/high-context relationships in this term. These differences in the communication between high-context and low-context cultures lead to tremendous losses of relevant knowledge within the transfer process between these groups.

Contact and communication between different cultures is an inherent fact of offshoring, thus research on cross-cultural issues in this area is gaining more and more emphasis. Motivated by the immense negative influence of cross-cultural issues on the offshore performance in software development projects (Carmel and Tija, 2005), even information systems research is ‘seeking culture’ nowadays. As a conclusion, the common understanding of culture is that it is learned, associated with values and behaviours, shared by a group and passed from one generation to the next (MacGregor et al., 2005).
To explain cultural differences researchers make use of dimensions of cultural variations. Dimensions in this context are aspects of a culture which can be measured in relation to other cultures (Hofstede and Hofstede, 2004). Triandis provides an overview of the most popular cultural dimensions (Triandis, 1982). Referring to them helps to understand and explain why people from other cultures behave and think differently than we do. Therefore in the context of offshore software development and the necessary transfer of knowledge we need to analyse some of the typical dimensions to understand why the knowledge transfer is so complicated between team members from different cultures.

The findings of Hofstede and Hall are often discussed in the scientific community and based on their work we can build a model of the most important cultural obstacles that impede the performance of knowledge transfer. The following seven obstacles have been summarized and selected from the cultural orientations formulated by Hofstede and Hall (Hall, 1976, Hofstede, 1984). This selection is based on casual expert interviews in preparation for this research initiative.

Firstly the power dimension is one of the most important in any business context. The structure of power accounts for the expression of emotional distance between subordinates (Hofstede, 1984) and superiors where higher power cultures tend to have more autocratic managers (Hall, 1976, Hofstede, 1984). Individuals in such cultures are less likely to express disagreement with their supervisors. Less power-orientated cultures use participatory and consultative management styles. When both extremes have to collaborate in a knowledge transfer initiative cultural obstacles may emerge.

Secondly, relationship dimensions reflect the difference between individualism and collectivism. People from individualistic cultures tend to highly value personal freedom, privacy, and time (Hall, 1976, Hofstede, 1984). They are usually expected to look out for themselves, especially in a business context. For more collectivist-orientated cultures, group harmony is more important than personal ambition. At work they have a higher dependence on organization and a stronger desire for non-financial rewards. Some authors in the knowledge transfer community argue that individuals from collectivist cultures are better suited as knowledge transfer partners, because no financial reward is required.

Different cultures experience time dimensions differently. For certain groups deadlines are firm and literal, in other words people tend to be on time (e.g. stereotypical Germans and Americans). For others the interpretation of time is more flexible. A team of mixed cultures may find it hard to meet knowledge transfer milestones and to dedicate time for joint work sessions when one part of the team has a different understanding of when to meet a given objective. Since knowledge transfer in many cases, as in the transfer of implicit knowledge, requires that two individuals work together, a different understanding of how often and how rigorously to schedule joint meetings may slow the employee from the more ambitious culture. This would naturally be lead to frustration and conflict between the two parties.

Dimensions of uncertainty, as defined by Hofstede, represent the amount of uncertainty an individual tolerates. This is due to the fact that the business environment requires numerous decisions involving doubt and risk. Examining this perspective on the unknown will contribute a description of how people cope with ambiguity. Hofstede, for example has found that British people can handle uncertainty better then Germans (Carmel and Tija, 2005). A similar difference may also arise between German and Chinese workers.

Hofstede defines the ‘future’ dimension as how focused on the future a culture is. East Asian countries, including China, Korea, and Japan, tend to be very forward looking. The central purpose of orienting one’s work around the future or the long-term implies delaying present gratification or gains in return for future prosperity on a grander scale. Naturally, the opposite would be an emphasis on the present, where instant gratification would reign supreme, or on the past, where present ambitions are shaped by former achievements. In the context of knowledge transfer obstacles will present themselves when one group of workers invests much more time into the long-term objective of knowledge transfer than the other.
Such frustrating situations may become worse if the communication dimensions of the parties are also incompatible. Two classifications of communication orientation can be found in the relevant literature on culture in general: high- versus low-context communicators. Low-context cultures listen to what is said rather than how it is said while high-context cultures consider secondary factors such as one’s tone and peripheral and contextual information in order to understand each other. Given the fact that a knowledge transfer requires two individuals to communicate regularly often regarding entirely new concepts, different communication orientations can become significant obstacles for knowledge transfer. For example, a low context communicator might find it difficult to explain something to a high context communicator often interrupting him because he sensed, that his partner is bored and tries to convey interest by asking a confirming question.

Finally we identify information processing dimensions by the way cultural groups process information. East Asian cultures tend to see more relationships and connections between disparate pieces of data. Westerners distinguish more across categories and taxonomies in a rather disconnected approach. As with communication problems, an expert may find it difficult to explain an isolated metaphor to an individual thinking in terms of relationships.

These cultural dimensions help to understand the basic principles of cross-cultural communication and data processing. We concentrate our efforts on the first three dimensions, as we found storytelling to have a strong positive effect on intercultural problems in those areas, but a weaker effect on the other dimensions.

3 STORYTELLING

When transferring knowledge to other operating units of a MNC, the overall goal is to implement the knowledge sent successfully at the receiver’s unit (Sorensen and Snis, 2001). Therefore, a shared, explicit understanding of concepts, categories, and descriptors lays the foundation for effective communication and knowledge transfer in organizations (Zack, 1999).

The knowledge to be sent needs to be transferred in a format that can be understood by the receiver (Thomps, 2002). Unfortunately, most of the time the encoded messages cannot be considered universal, since they are culture-specific and arbitrary (Roth, 2001). This might not always be obvious during the communication process. Messages received from individuals of other cultures might have an outward similarity with messages of the home culture; their culture-specific differences are often ignored. This might also influence the transfer of knowledge negatively. Successful transfer of knowledge must thus be based on a collaboratively established consensus among the participants (Sorensen and Snis, 2001) and can improve relationships among organizational communities if there is a commonly acknowledged context in which the significance given by the users to the symbols are unique (Dupouet and Laguecir, 2002).

Ever since human beings have communicated and socially interacted with each other, stories have played a vital role in exchanging and propagating complex ideas and disclosing knowledge. In every culture, different stories exist and have been used to preserve and pass on knowledge from generation to generation. Stories are in a certain intrinsic sense interesting, because they are an attractive high-priority memory booster. With purpose and a meaning behind it, stories will draw and grasp the attention of any audience and in this sense will outperform any logical argument (Haghirian and Chini, 2003, Papadimitriou, 2003).

Stories and narratives are reports about company related incidents that happened in the past and that have a special meaning for the company. Davenport and Pruzak (Davenport and Prusak, 1998) claim the most efficient way of transferring knowledge is through a convincing narrative. People prefer to talk to their colleagues about their latest ideas (Birkinshaw, 2001). They tell stories to exchange knowledge. So narratives are used in order to transfer the complex contents of tacit knowledge (Snowden, 2002). An organizational story is defined as a detailed narrative of past management actions, employee interactions, or other intra- or extra-organizational events. These stories are usually
communicated informally within the organization. Normally, such stories consist of a plot, major characters and an outcome (Swap et al., 2001). Purposeful stories will be able to capture and hold the attention of the audience. They are rooted in truth and are self-propagating (Snowden, 1999). Snowden distinguishes between two kinds of storytelling: storytelling as a knowledge disclosure mechanism and storytelling to create meaning and understanding that can be a helpful tool in getting hold of the valuable tacit knowledge of members within the organization. Storytelling to create meaning and understanding creates metaphors to transfer knowledge in a more transparent way (Snowden, 1999).

Lately, much emphasis has been placed on stories within the organizational knowledge discussion and especially on stories as a tool for knowledge management. Based on studies on communities of practice, of technical knowledge transfer, e.g. Orrs study on Hewlett Packard technicians (1990), and on organizational sense-making processes, it is claimed that stories may fulfil a variety of functions such as the distribution of uncodified or tacit knowledge within knowledge management (Schreyogg and Geiger, 2005). Furthermore, stories allow the listener to comprehend new experiences and to create impressions about the persons, objects and beliefs of the storyteller. Stories help develop general attitudes and beliefs (Adaval and Wyer, 1998). Storytelling as a mechanism for disclosing knowledge can be a helpful tool to get hold of the valuable tacit knowledge within a project team. It creates a self-sustaining, low cost means by which knowledge can be captured on an ongoing basis (Haghirian and Chini, 2003).

These assumptions are based on an understanding of the knowledge taxonomy and address the socially and contextually-bound nature of knowledge, by which any formalised or explicit knowledge can only be understood through its tacit components. Therefore, knowledge can only be shared and understood successfully among people if, and only if, the participants involved share a general set of meanings, beliefs, values and a socially common interpretation. Stories do address the tacit part of knowledge and thus can be seen as a way to establish coherent structure of meaning and frames of references needed to interpret explicit forms of knowledge ending in an effective exchange (Meyer et al., 2005).

Organizationaly, stories emerge as a natural part of the day-to-day life, the routines, and the ongoing communication between individuals and groups. Not as a tool but rather unconsciously, they develop from events, extraordinary situations, successes and failures and are told and retold in everyday organizational life. Though during offshore software development, teams are geographically distributed and hence informal communications, spontaneous conversation and informal “corridor talks” are eliminated. This informal talk helps people stay aware of what is going on around them, what people are working on, what states various parts of the project are in, who has expertise in what area, and many other essential pieces of background information that enables teams to work together efficiently. In addition, different cultural and social backgrounds exist, resulting in an absent common meaning structure.

To bridge these gaps of culture, trust building, informal corridor talks and collaboration, we propose that storytelling may serve such a purpose. Stories have been used in all cultures to communicate values, norms etc. for centuries (Haghirian and Chini, 2003). Building on the findings of Hofstede and Hall, we can assume that organisations in high-context-cultures emphasise storytelling more. This is especially important for software offshore development because this usually involves low-context cultures in the western hemisphere and high-context Asian cultures.

Storytelling to create meaning and understanding creates metaphors to transfer knowledge in a more transparent way. They help to better transfer any information or formal knowledge in a sequential order, with priorities and including a chain of motivation or justification of the inherent transported tacit parts. This can be especially important in a cross-cultural context. People from high context cultures emphasise interpersonal relationships and developing trust as an important first step to any business transaction. In contrast, people from the low context cultures value logic, facts and directness. To be absolutely clear, they strive to use precise words and intend them to be taken literally (Hall, 1976). These very different styles of communication can more often than not cause misunderstandings and sometimes even failures in the intercultural communications process.
In addition to those communication difficulties, the effective and successful transfer of knowledge between people poses further difficulties. One reason is the ambiguous nature of knowledge itself as a result of the previously mentioned context and social embeddedness.

Especially for software offshore development, storytelling seems a promising tool for transferring tacit knowledge, as other instruments like social interaction between company members, traditions, routines and learning-by-doing are usually implausible due to geographical distance or the impossibility of face-to-face communication.

4 RESEARCH QUESTION AND METHODOLOGY

Storytelling is portrayed as an effective tool to communicate and transfer knowledge within cross-cultural teams. However, the case of transferring knowledge via stories in the IT industry has so far not been investigated.

The goal of our research is therefore to examine how telling successful organizational stories can be applied when communicating technological knowledge between geographically dispersed teams that also have different cultural backgrounds.

Since there is little evidence on knowledge transfer via stories and storytelling in a cross-cultural context we applied a qualitative research approach. This allows us to investigate a contemporary phenomenon within its real-life context where the relevant behaviour cannot be manipulated (Yin, 2002). The research is explanatory in nature and relies on an in-depth case study. The collection of data included interviews as primary sources and secondary information from documents and questionnaires regarding software development were used to assure triangulation.

The interviews were mainly conducted from winter 2007/08 through spring 2008. They involved both the client and the vendor, and were conducted each time with a project manager and developers in charge of the relevant project; in total we talked with 14 people. The interviews lasted 45 to 120 minutes. They were semi-structured to allow flexibility and to ensure that the researchers captured any interesting phenomena. Questions were formulated according to perceived performance of the projects, the project communication, the standards and details of the development process and the appearance of context-relevant information. The interviews were conducted with staff and senior management of each company, in Bangalore, India and Leipzig, Germany, together with a review of company documentation and formal presentation material. A number of telephone interviews were also conducted with vendor staff in the United Kingdom (a branch from the Indian company) and India. Gathered data currently includes approximately 90 hours of interviews.

To achieve an adequate level of validity we used multiple sources of evidence and had key interviewees as reviewers. Internal validity, needed for explanatory case studies, was obtained by using a pattern matching technique after coding the interviews. Causal chains are derived from the data analysis in order to later build a causal model (Miles and Huberman, 1994).

5 STORYTELLING AS A TOOL FOR KNOWLEDGE TRANSFER IN THE IT INDUSTRY

We present a case study in this section which will provide us with the result at the end of our research. With respect to the involved corporate partners we will not mention their real names. From a country perspective, India is still the global leader in providing offshore services (Kobayashi-Hillary, 2005). The subcontinent will continue to fulfill this role in the future due to its low labor costs and an abundance of skilled workers (Gott, 2007). We conducted an in-depth case study research involving an Indian vendor and a software company from Germany (client) in order to develop an understanding of the impact of storytelling during offshore software projects.
Here are some basic facts from the case background: the client team had 3 developers, one project manager and one unit manager. On the vendor site, the team had 6 developers, one project manager and one key account manager. Project duration was 8 months and the project was completed in August 2008. Both companies (client and vendor) could be classified as small to medium enterprises. The Indian company has 900 employees and is focused on software services and the client company has 40 employees and is specialised in IT Services for public companies.

The case involved a software development project that included further development and enhancement of an existing software application. Challenges included simultaneous ongoing development at both the client-site and the vendor-site. Therefore it was important for both parties to share the same vision of the products future roadmap. Both teams were urged to share their work experiences and challenges to be implemented in future work.

While focusing on the beginning and kickoff of the cooperation, storytelling was used mainly during the early phases on the project, using a variety of means of communications, e.g. face to face, via phone, chat, mail, and documents.

At the project start many documents and source code files had to be provided to the vendor. A kick-off meeting with members of both parties was held where the development vision was communicated, milestones and timelines were set and specific development tools were agreed upon. Not only the core facts related to the project had been communicated at this meeting but also soft facts like escalation chains and communication schedules for teleconferencing and instant messaging and especially the preconception of the client concerning the realisation of the project. One month after the project start the vendor provided the client with re-briefing and detailed requirement specifications in regards to the vendor’s processes. After the initial kickoff and during the starting phase, phone conferences were the primary form of communication. Here, stories were developed and transferred mainly for the Indian developers to enhance and facilitate the understanding of the development background of the product as well as the motivation, the history and related problems and solution. The focus was to transfer the client’s preconceptions and to determine a possible solution. In the ongoing process, the emphasis of the stories shifted towards the use of feedback rounds in which soft factors or problems became an issue. Here stories were used to bridge the difference in dealing with different approaches of problem solving, e.g. dealing with direct critique, and used to understand timing issues, e.g. meeting deadlines. Therefore, main themes of the stories were cultural differences, descriptions of different ways of collaborative work styles and team approaches integrating do’s and dont’s.

To give an example of such stories, one project manager told us the following story:

"Once I was the technical contact of an offshore project. The Indian company gave my contact data to one of their programmers and if he had any questions I was the person he communicated with. I’m a programmer myself so we share a certain degree of experience. But sometimes the Indian programmer asked an elementary question or could not solve a simple problem that made me think he lacks some basic programming skills. At one point (after several days on his part trying to solve a particular problem) he declared a certain task impossible when I knew it would be quite easy to accomplish. Using a web browser and typing the three keywords into Google gave the correct solution ranking first. So I sent him the article I found and an example of how to accomplish that particular task. I did not get an answer to that email but the next email merely stated the problem had been addressed."

The project manager used his experience from a former project to prepare the team with such stories to establish an open communication within the developer team. From that on, he could clearly ask the Indian team if they needed help. Further he told us many more stories he used in team communication and project set up. The integration of project experiences into the development of stories of both parties involved, the client and the Indian vendor, helped the Indian side, which at the beginning
seemed rather resistant to this management tool, to accept the stories. However, since only some of the developers were aware of the stories, this hindered the Indian team as a whole to take full advantage of the stories in all areas of the development process.

Stories were used at the beginning of the project to diminish issues related to cultural differences. They were also used to convey clear guidelines to the client’s counterparts for better communication. Facts explaining why things are done the way they are, the client’s expectations etc. were conveyed using stories. Regular meetings were scheduled and took place for tracking the project’s progress, noting any significant hindrances to the process. The client transformed his experiences into a story and conveyed it to the vendor, which made it easy for both sides to work collaboratively. These early data show results from instant messaging chat protocols, voice chats, interviews of involved team members of both client and vendor and project documentation.

6 LESSONS LEARNED

Based on our findings, we propose that storytelling can be used as an appropriate instrument for transferring knowledge especially in cross cultural contexts, where the differences in dealing between low context and high context cultures are remarkable. The following table shows a summary of the utilisation of storytelling during the case.

<table>
<thead>
<tr>
<th>Purpose of storytelling</th>
<th>Used a story that,</th>
</tr>
</thead>
<tbody>
<tr>
<td>Establish culture behaviour</td>
<td>Highlights typical problems in dealing with different cultures</td>
</tr>
<tr>
<td>Introduce collectivistic Teamwork</td>
<td>Describes and explains teamwork and different collaborative styles</td>
</tr>
<tr>
<td></td>
<td>Displays different roles within teams</td>
</tr>
<tr>
<td>Bridging power distance</td>
<td>Shows a variety of escalation chains and means of decision making</td>
</tr>
<tr>
<td></td>
<td>Shows advantages of transparent decision making processes</td>
</tr>
<tr>
<td>Bridging between high and low context</td>
<td>Shows advantages of a culture fostering open discussion at all levels</td>
</tr>
</tbody>
</table>

Table 1. Purposes of storytelling

7 DISCUSSION AND CONCLUSION

With the help of our case study, we are making an attempt to explore and make use of the benefits the storytelling tool can provide in offshore projects. So far, the findings of the study show that it is a practical and beneficial solution in offshore situations to bridge cultural differences between the parties and members involved. Stories are not generic and are highly related to specific organisational and cultural conditions and therefore create a shared vision, sparking action, and fostering collaboration and understanding within the team. A deeper recognition and understanding of this
utilisable tool in IT organisations is still needed. This area of study must be examined closely with respect to its appropriateness in the knowledge management field. The limitation we see so far is a lack of storytelling cases in the field.

Further research aims to develop guidelines for a way of transferring collective experiences of source group to members of a target group by proper co-ordination and co-operation. Furthermore, research should be validated using additional cases and develop a framework for utilising storytelling as an instrument of knowledge transfer. Research may also focus on different aspects of the storytelling method such as the influence of oral vs. written communication on knowledge transfer, limitations of storytelling, and the role storytelling can play in decision making processes during knowledge transfer processes.

References


KNOWLEDGE SHARING IN ONLINE COMMUNITIES

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0730.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Knowledge-based community, Empirical study, Web 2.0, Shared knowledge</td>
</tr>
</tbody>
</table>
KNOWLEDGE SHARING IN ONLINE COMMUNITIES

Meng, Zhaoli, Renmin University, School of Information, Beijing, 100872, P. R. China, mengzhaoli@gmail.com

Gong, Jiong, University of International Business and Economics, School of International Trade and Economics, Depaartment of Economics, Beijing 100029, P. R. China, johngong@gmail.com

Abstract

This paper investigates online knowledge sharing behaviour in Baidu Knows, a platform sponsored by the largest search engine company Baidu in China. We developed a spider engine to collect data from over 2 million questions posted at Baidu Knows. The data collected allows us to profile registered members, to answer questions such as who are the main driving force of those online communities and their attributes, and who are the free-riders seldom posting and answering questions. We also test several hypotheses in explaining the motivations of knowledge sharing, and the quality of such knowledge sharing. Our results indicate that there is a large proportion of users who seldom answer questions, but always ask questions. Although their behaviour seems selfish at the surface, they are actually the key factors driving the growth of online knowledge sharing communities.

Keywords: knowledge-based community, shared knowledge, Web 2.0, empirical study
1. INTRODUCTION

Online communities are virtual spaces on the Internet where a group of people of similar interests communicate and interact collectively or individually to achieve certain goals. Its popularity has grown tremendously in recent years that in many instances online communities have become one of the main resources of knowledge seeking. The most famous and successful example is Wiki, the largest free encyclopedia in the world, where anyone can access its web pages to contribute or modify content. However, as its name indicates, the Wiki website is only useful for acquiring information about well-established encyclopedic subjects and thus lacks the flexibility for knowledge seeking of more general and practical types.

The model of seeking general and practical knowledge on the Internet has evolved substantially over years. Previously, Internet users actively sought knowledge via search engines, such as Google and Yahoo!. However, the search efficiency highly relies on the selection of keywords. Also the information obtained tends to be fragmented, incomplete and may even be wrong sometimes. When users are unfamiliar with the subject matter, or choose inappropriate keywords, the knowledge seeking process tends to be quite inefficient.

The broad class of Web 2.0 technologies attempts to improve upon the search process by establishing a passive way of acquiring knowledge. In this model, users can propose questions in virtual communities, and expect answers provided by other users. Once they find the right place in the knowledge sharing platform to post their questions, the rest is left for others to contribute answers. The questions range from highly specialized technical questions to everyday practical questions. Interestingly, in large knowledge sharing platforms, such as Yahoo! Answers and Baidu Knows, millions of online users are active in the community, spending time to read posts of others, and answer questions if possible.

Launched in June of 2005, Baidu Knows is China’s largest knowledge sharing platform. In Baidu Knows, the questions proposed by users are classified into over 500 categories at the time of this research, which are still increasing at a rapid pace everyday. Over 23.3 million questions and 22.8 million answers have been posted by the end of 2007. As a platform that was established only 2 years ago, Baidu Knows’ phenomenal success arguably plays an important role in Baidu’s competition with Google in the search engine arena in the Chinese Internet.

The success of the knowledge sharing platforms begs two important questions. First, why do people devote their time and effort to sharing their knowledge in virtual communities helping others, and who are they? Second, what is the quality of the answers provided and how useful are they? Obviously without the quantity and the quality of the answers, the usefulness of the knowledge sharing platform would dissipate, ultimately leading to reduced hit rates.

Despite the phenomenal growth of online knowledge sharing communities, there is surprisingly little empirical research about actual knowledge sharing patterns. Most prior studies use survey methods to investigate users’ knowledge sharing motivation. New online knowledge sharing communities enable a new way of knowledge sharing: users freely propose their questions and exchange opinions with members. However, the technology infrastructure alone cannot guarantee meaningful and constructive knowledge sharing behavior (Alavi and Leidner 1999). Its survival totally depends upon the “kindness of strangers” as previously suggested (Constant et al. 1996). This is in sharp contrast to conventional communities, where the main factors driving knowledge sharing are identified as strong ties (Wellman and Wortley 1990), co-location (Allen 1984, Kraut et al. 1990), demographic similarities (Pelled 1996), and status similarities (Cohen and Zhou 1991). However, Wasko and Faraj (2005) argue that these factors are not applicable in virtual communities, where participants are typically strangers, and do not always have demographic or status similarities.

---

1 Baidu is the largest search engine company in China.
Then why do people spend their valuable time and effort to contribute knowledge helping others? Prior studies postulated a variety of theories explaining knowledge contributions. Social capital is typically defined as “resources embedded in a social structure that are accessed and/or mobilized in purposive action” (Lin 2001, p. 29). Wasko and Faraj (2005) identified three forms of social capital related to knowledge sharing behavior in electronic networks. Chiu et al. (2006) also identified the influence of social capital on individual’s willingness to share knowledge. Another theory concerns the collective action. Wasko and Faraj (2005) theorized that in online communities, individuals voluntarily contribute their time, effort, and knowledge toward the collective benefits. Intrinsic benefits such as self-worth, social norm, social affiliation, and extrinsic benefits such as economic rewards, future promotion opportunities, can drive knowledge contributions (Ma and Agarwal 2007, Bock et al. 2005, Kankanhalli et al. 2005). Donath (1999) investigated Usenet newsgroup and concluded that identity plays a vital role in knowledge sharing behavior. The identity definition includes both the establishment of their own reputation and the recognition by others.

In this paper, we address knowledge sharing issues with empirical data collected from a spider engine that we developed. The data enables us to build a profile of those who post questions and those who answer questions. This allows us to address questions such as how many users are free-riders who seldom share their own knowledge while only asking questions, and who are the main contributors who are the driving force of the community? What are their main attributes, and whether knowledge sharing in different information categories exhibits different sharing patterns? With respect to the answer quality issue, we established a simple regression model to investigate the correlation between the quality of knowledge sharing and the characteristics of those who contribute knowledge. By analyzing users’ true knowledge sharing behavior, this study sheds light on how to induce, promote, and manage online knowledge sharing behavior.

The rest of the paper is organized as follows. Section 2 describes the spider engine and the data collection effort, followed by a profile analysis of different classes of users. Section 3 tests several hypotheses of different behavior among user groups in motivating knowledge sharing. Section 4 looks at the free-rider phenomenon. We also established a theoretical model that investigates the correlation between answer quality and the characteristics of those who contribute answers in section 5. Section 6 contains concluding remarks.

2. USER PROFILE ANALYSIS

To collect field data, we choose from two of the largest knowledge sharing platforms in the world, Baidu Knows and Yahoo! Answers. Interestingly, these two platforms adopt different mechanisms of how to reward users’ sharing behavior. Yahoo! Answers makes each question have the same Reward Points as answerers. Once the question is answered, the best answerer will get the Reward Points, which are given by Yahoo! Answers. In Baidu Knows, in contrast, askers could choose how many points they want to pay, and the Reward Points are deducted from askers’ cumulative points. Compared with Yahoo! Answers, Baidu Knows seems to provide more flexibility in motivating knowledge sharing. The comparison of these two mechanisms in driving their respective growth would be an interesting subject, but we will leave this issue to another research paper in the future.

We developed a spider engine to collect data in Baidu Knows as of December 2007. First we identified users who have asked or answered questions in one week from our sample in Baidu Knows’ Computer and Entertainment categories. Overall, we got 2,763 questions and 7,520 answers, with each question having an average of 2.72 answers. For the Computer category, there were 1,383 questions and 3,417 answers, averaging 2.47 answers per question. For the Entertainment category, there were 1,380 questions and 4,103 answers, averaging 2.97 answers per question. Detailed descriptive statistics of questions and answers are shown in the following Table 1:

---

2 This version reports data from Baidu Knows. Data from Yahoo! Answers will be incorporated in a forthcoming version.
The collected data contains past information of each user ID, which specifies three variables: No. of Answers, Reward Point and Best Answer Ratio. For privacy concerns, Baidu does not provide information about how many questions each ID has asked. We then collect all the information about these three variables for each user ID. The exact definition of these three variables are as follows:

- No. of Answers shows how many times each ID shared his or her knowledge with others.
- Reward Point tells how many points won by answering others’ questions.
- Best Answer Ratio is the ratio between the answers rated as the best answers by askers and the total number of questions one member answered.

Each user ID’s past history of asking and answering questions allows us to enlarge the sample size to include 2,114,930 questions. Based on the profile of questions posted and answered, we classify all members into several categories: top contributors, casual contributors and non-registered free riders.

Figure 1 illustrates the distribution of all registered members. Data along the x axis in Figure 1 is rank-ordered by the No. of Answers in the past, and the y axis is the No. of Answers provided by each corresponding registered member. These is clearly a concentration of users who often answered questions. Therefore we classify the top 10% of registered members as top contributors, who answered 1,618,161 questions in total, accounting for 76.6% of the total questions. Casual contributors are those registered members who answer questions occasionally. They account for 90% of the registered members. There are clearly also non-registered members of online knowledge sharing communities outside our sample. We define them as non-registered members who only come to the website to view questions and answers without contributing anything.

### Table 1: Descriptive Statistics of No. of Answers for Each Question

<table>
<thead>
<tr>
<th></th>
<th>Computer</th>
<th>Entertainment</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Questions</td>
<td>1383</td>
<td>1380</td>
<td>2763</td>
</tr>
<tr>
<td>No. of Answers</td>
<td>3417</td>
<td>4103</td>
<td>7520</td>
</tr>
<tr>
<td>Mean of No. of Answers</td>
<td>2.47</td>
<td>2.97</td>
<td>2.71</td>
</tr>
<tr>
<td>Median of No. of Answers</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Std. Dev of No. of Answers</td>
<td>0.06</td>
<td>0.07</td>
<td>0.05</td>
</tr>
<tr>
<td>Minimum of No. of Answers</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maximum of No. of Answers</td>
<td>26</td>
<td>24</td>
<td>26</td>
</tr>
</tbody>
</table>

*Figure 1: Users’ Knowledge Sharing Pattern*
Past literature reports that knowledge contribution in Linux working groups follows a power distribution. Figure 1 seems to indicate the same pattern, as 10% of the members account for about three quarters of all answers. We test the power distribution for this dataset as well in SPSS. The power test is significant at 99% level.

Figure 2: Knowledge Sharing Pattern

Figure 2 above illustrates the distribution of member segments as defined previously. One interesting fact about top contributors is that not only the quantity but also the quality of their answers is higher than others. The quality of shared knowledge is measured by Best Answer Ratio. For each question proposed in Baidu Knows, askers are required to select and mark one best answer, if possible, from all the answers proposed by other members. The profile of each member contains the Best Answer Ratio, which shows the percentage of answers that are rated as the best among all answers provided by each member. The average Best Answer Ratio of top contributors is 20.9%, which is much higher than the 12.5% average for the rest. This indicates that top contributors appear to be the driving force of the knowledge sharing communities. They answer about three quarters of the total questions, and one fifth of their answers have been selected as the best answer by askers. On average, each of them has answered 2,021 questions, which is quite astonishing considering the two year time span of our sample data.

Interestingly, the data shows that top contributors can be further classified into two different subgroups by using the 1,000 Reward Points as the demarcation point. 85.3% of the top contributors’ Reward Points are less than 1,000, indicating that they asked less than 223 questions in the past. This is in sharp contrast to the 2,021 questions they answered. As a result, we call these members pure contributors, as they seem to be inclined to passing knowledge to others without expecting anything in return.

The second group of top contributors not only answers many questions but is also active knowledge seekers by asking many questions. 14.7% of the top contributors’ Reward Points are higher than 1000, with the average at 2,619, which means that they asked 589 questions on average in the past. We call these members active contributors and knowledge seekers. On one hand, they are keen on sharing knowledge with others. On the other hand, they are also active knowledge seekers, who frequently propose questions and give high rewards to answerers.

3 When we calculate the average Best Answer Ratio, we delete those members whose No. of Answer is less than 10.
Casual contributors are 90% of all registered members in the community. They answered 496,769 questions, accounting for 23.5% of the total questions. Both the quantity and the quality of their shared knowledge are inferior to that from the top contributors.

3. KNOWLEDGE SHARING MOTIVATIONS

Social exchange theory states that members engage in social interaction based on the expectation that it will lead in some way to social rewards, such as approval, status, and respect (Blau 1964). Ma and Agarwal (2007) theorized that participants share knowledge in online communities for both extrinsic benefits and intrinsic benefits. Donath (1999) proposed that in Usenet newsgroups, identity, both the establishment of their own reputation and the recognition of others, plays a vital role.

In knowledge sharing platforms, Reward Points could serve both as extrinsic award and intrinsic award. First, Reward Points act as the electronic money in the community. When members propose their questions, they need to pass some of their Reward Points to answerers. Second, Reward Points indicate the status of the member. In Baidu Knows, members are classified into 18 levels. Once members have accumulated certain Reward Points, they could be promoted to a higher level, indicating that they are more knowledgeable and have higher willingness to help others. As a result, high Reward Points can help members move up to higher levels and receive more esteem from others. This is hypothesis 1:

H1: Members share knowledge because of extrinsic and intrinsic motivations. The higher the Reward Points are associated with a question, the more answers it will get from others.

To check whether the Reward Points assigned to a question are related to Number of Answers it is likely to solicit, we use nonparametric test: Pearson’s r, Spearman’s rho, and Kendall’s tau_b. The results are reported in Table 2:

<table>
<thead>
<tr>
<th></th>
<th>Computer</th>
<th>Entertainment</th>
<th>Both Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson’s r</td>
<td>0.516***</td>
<td>0.249***</td>
<td>0.372***</td>
</tr>
<tr>
<td>Spearman’s rho</td>
<td>0.254***</td>
<td>0.105***</td>
<td>0.167***</td>
</tr>
<tr>
<td>Kendall’s tau_b</td>
<td>0.208***</td>
<td>0.084***</td>
<td>0.136***</td>
</tr>
</tbody>
</table>

*** Significant at the 0.01 level (2-tailed)

Table 2. Nonparametric correlation between Reward Points and No. of Answers

Our results show that in both the Computer category and the Entertainment category, Reward Points are significantly correlated with No. of Answers. The strength of correlation is higher in the Computer category than that in the Entertainment category. Therefore, H1 is accepted.

People are not only pragmatic but also expressive of feelings, values, and self-identities (Baudura 1986, Schlenker 1985, Shamir 1991). In knowledge sharing communities, members could increase self-worth, personal identification with the organization, self-respect, respect from others, and feelings of commitment by helping others. Bandura (1986) posited that self-evaluation based on competence and social acceptance is an important source of intrinsic motivation that drives engagement in activities for the sake of the activity itself, rather than for external rewards. Wasko and Faraj (2000) also demonstrated that members may contribute knowledge within the collective because solving problems is challenging or fun, and because helping others is enjoyable. Therefore, we may expect that enjoyment of helping is one of the main motivations for members’ contribution to the collective.

In knowledge sharing platforms, some questions do not provide Reward Points to answerers. If such questions are still answered by other members, we may expect that the answerers reply such questions not for Reward Points, but for the joy they achieved during the helping process. Therefore, we propose the hypothesis 2:

H2: Some member share their knowledge because of the enjoyment of helping others.

Proceedings ECIS 2009
Overall, the Reward Points given out are very low. 47.86% of questions do not have any Reward Points. 92.8% of questions have Reward Points less than 30. Only 3.3% of the questions have Reward Points higher than 100, the highest being 200. Those questions having zero Reward Points have on average 1.91 answers, and only 17.8% of them are not answered. These results seem to indicate that some members are not motivated by extrinsic awards, but possibly by the enjoyment of helping others, thus confirming H2.

Although exchanges in electronic social networks occur through weak ties between strangers, there is evidence of reciprocal supportiveness (Wellman and Gulia 1999). In online communities, members tend to show generalized reciprocity. Scholars believe that in online communities with strong sharing social norms, members may offer their help either because that they may have been helped by others in the past or they may expect that some one in the community would help them if they have a question in the future. If reciprocal supportiveness is one motivation for members’ knowledge sharing behavior, a member’s helping tendency should be valued by others in the community. In other words, a member who used to help others should be more likely to get help from others compared with free riders. Therefore, we propose the following hypothesis

H3: Members who shared their knowledge more frequently in the past will get more responses when they post their questions.

We use the following regression function to check the correlation between members’ past answering behavior and the No. of answers they received when they propose their own questions. The result are reported in Table 3

\[ \text{No. of Answers} = \text{Reward Points} + \text{Asker’s No. of Answers in History} \]

<table>
<thead>
<tr>
<th></th>
<th>Constant</th>
<th>Reward Points</th>
<th>No. of Answers in History</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole Sample</td>
<td>2.289(0.048)</td>
<td>0.035(0.002)</td>
<td>0.00001(0.000)</td>
</tr>
<tr>
<td></td>
<td>t = 47.417***</td>
<td>t = 21.076***</td>
<td>t = 0.206</td>
</tr>
<tr>
<td></td>
<td>R Square = 0.139</td>
<td>F = 222.152***</td>
<td></td>
</tr>
<tr>
<td>Computer Category</td>
<td>1.884(0.059)</td>
<td>0.04(0.002)</td>
<td>0.00001(0.000)</td>
</tr>
<tr>
<td></td>
<td>t = 31.735***</td>
<td>t = 22.390***</td>
<td>t = -0.802</td>
</tr>
<tr>
<td></td>
<td>R Square = 0.267</td>
<td>F = 250.793***</td>
<td></td>
</tr>
<tr>
<td>Entertainment</td>
<td>2.674(0.075)</td>
<td>0.028(0.003)</td>
<td>0.00001(0.000)</td>
</tr>
<tr>
<td></td>
<td>t = 35.865***</td>
<td>t = 9.555***</td>
<td>t = 1.031</td>
</tr>
<tr>
<td></td>
<td>R Square = 0.063</td>
<td>F = 46.152***</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Regression Results for No. of Answers against Asker’s No., of Answers in History

This regression results show that No. of Answers is correlated with Reward Point, which of course confirms H1. However, Askers’ No. of Answers in History is not significantly correlated with No. of Answers a question gets. The result is quite important: members’ helping behavior does not pay off in terms of getting more answers when they need help in the future. No. of Answers in History indicates how many times a user helped others and contributed his/her knowledge. When this user calls for help and posts questions, others do not seem to value the asker’s past behavior. That means, the conclusion that people help others due to reciprocal supportiveness motivation cannot be drawn in the Baidu Knows sample data, thus rejecting H3.

When members ask for help in online communities, they usually send out their Reward Points to attract others to answer their questions. However, around half of the questions do not send out Reward Points and the askers can be viewed as free riders. Such phenomenon begs the question who tends to value more others’ knowledge sharing behavior?

Individual centrality concept defines that individuals who are centrally embedded in a collective have a high probability of direct ties with other members. They have a richer social capital. Such members can
exert more influence than others. They are more likely to cooperate with others and comply with group norms and expectations (Rogers and Kincaid 1981).

Members who frequently shared their knowledge and have high Best Answer Ratio are core members of the community. Members with high Best Answer Ratio are more connected to other members and have developed the habit of cooperation. They understand the efforts of sharing knowledge through their past experience. When they call for help from others, they tend to follow the sharing group norm and value the efforts of others. Therefore, we propose the following hypotheses:

H4: Members who shared their knowledge more frequently before will set high Reward Points when they post their own questions.

H5: Members who have high Best Answer Ratio will set high Reward Points when they post their own questions.

To check H4 and H5, we use nonparametric tests to check the correlation between Reward Points and No. of Answers in History, and Best Answer Ratio. The results are reported in Table 4 and Table 5 respectively.

<table>
<thead>
<tr>
<th></th>
<th>Computer</th>
<th>Entertainment</th>
<th>Both Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson’s r</td>
<td>0.013</td>
<td>0.021*</td>
<td>0.006</td>
</tr>
<tr>
<td>Spearman’s rho</td>
<td>0.064**</td>
<td>0.047*</td>
<td>0.000</td>
</tr>
<tr>
<td>Kendall’s tau_b</td>
<td>0.051**</td>
<td>0.003</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 4. Correlation between Reward Points and No. of Answers in History

<table>
<thead>
<tr>
<th></th>
<th>Computer</th>
<th>Entertainment</th>
<th>Both Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson’s r</td>
<td>0.119***</td>
<td>0.059*</td>
<td>0.089***</td>
</tr>
<tr>
<td>Spearman’s rho</td>
<td>0.073***</td>
<td>0.023</td>
<td>0.026</td>
</tr>
<tr>
<td>Kendall’s tau_b</td>
<td>0.059***</td>
<td>0.19*</td>
<td>0.020</td>
</tr>
</tbody>
</table>

Table 5. Correlation between Reward Points and Best Answer Ratio

The results show that Reward Points of questions in the Computer category are correlated with askers’ past behavior. In that category, members who answer more questions, or Members whose answers are frequently adopted by others as the best answer tend to set high Reward Points for their question. Therefore, H4 and H5 are accepted in the Computer category, while the same can not be said about the Entertainment category.

4. FREE RIDER BEHAVIOR

To investigate free rider behavior, we classify members into askers and answerers. Askers are defined as those members in the sample who have asked at least one question. Answerers are defined as those members who have answered at least one question. Table 6 shows the cumulative percentage of members’ past No. of Answers in the Computers category. According to their past behavior, 38.4% of askers never answered questions. 63.1% of askers answered less than 10 questions before. In contrast, only 1.3% of answerers never answered question before and answered their first question in our sample. Also, only 9.3% of answerers answered less than 10 questions before. Further more, 87.4% askers answered less than 100 questions, while only 43.6% answerers answered less than 100 questions.

<table>
<thead>
<tr>
<th>Past No. of Answers in Computer Category</th>
<th>Askers’ Percentage</th>
<th>Cumulative Percentage</th>
<th>Answerers’ Percentage</th>
<th>Cumulative Percentage</th>
</tr>
</thead>
</table>

Proceedings ECIS 2009
In the Entertainment category, the results are quite similar. As shown in Table 7, 33.8% of the askers are free riders, while only 1.6% of the answerers are free riders.

<table>
<thead>
<tr>
<th>Past No. of Answers in Entertainment Category</th>
<th>Askers’ Cumulative Percent</th>
<th>Answerers’ Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>33.8%</td>
<td>1.6%</td>
</tr>
<tr>
<td>Less than 10</td>
<td>59.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td>Less than 100</td>
<td>87.4%</td>
<td>54.8%</td>
</tr>
<tr>
<td>Less than 1000</td>
<td>98.8%</td>
<td>87.5%</td>
</tr>
</tbody>
</table>

Table 7. The Cumulative Distribution of Past No. of Answers in Entertainment category

Such results clearly illustrate that askers show less willingness to help others compared with answerers. In both categories, askers show much less willingness to help than answerers. Around 35% to 40% askers in both categories are free-riders, as they never answered questions before. Around 60% askers in both categories answered less than 10 questions before. Although such members show more selfishness in contributing knowledge than others, it appears that they are actually the key factor to sustain the online community. They are active knowledge seekers and propose around 60% of questions. They value the answers from others and actively rate which one is the best. Because of the existence of such members, online community is able to sustain and prosper.

The free rider phenomenon may also be viewed from a box-plot of user behavior. We collect past behavior of askers and answerers and examine how many times each member have shared their knowledge before. The box-plot of the Computer category [Figure 3] and the Entertainment category [Figure 4] clearly show that answerers usually shared knowledge more frequently than askers. Moreover, answerers have much more outliers compared with askers, indicating that more answerers are extremely generous with respect to sharing their knowledge.

Figure 3. The Boxplot of the Computer Category
Besides the free rider behavior associated with askers and answerers, many other non-registered free riders are not included in our sample. Baidu Knows is an open community, that attracts users who did not register at Baidu Knows, but still browse the questions and answers in seeking knowledge. Baidu Knows also gets routed traffic from the regular Baidu search engine. Since this type of free rider behavior does not leave any trace in the community, it is difficult to investigate their behavior and assess their population size. We did surveys in two undergraduate and three postgraduate classes in a well-established university in China. All the students in these classes had visited Baidu Knows and browsed questions and answers. However, only one tenth of them were registered members in Baidu Knows. Such simple survey results indicate that the proportion of free riders in the knowledge sharing communities is much larger than that reflected in our sample.

5. QUALITY OF KNOWLEDGE SHARING

Quality of shared knowledge is one of the key factors determining whether the community will be sustainable over the long run. Those members who share high quality knowledge are the most valuable resources of the community. Who will share high quality knowledge? What are the characteristics of the members who share high quality knowledge? These are the questions we address in this section.

The self-verification concept, rooted in cognitive dissonance theory, suggests that people are more satisfied and likely to participate in a relationship when their salient identities are confirmed by others in a group (Swann 1983, Swann et al. 1987). Members who shared high quality knowledge obviously have higher Best Answer Ratio. Their sharing behavior is acknowledged by others. They will have a relatively high self verification than others and have higher motivation to help others in the future. Therefore, we suggest the following hypotheses:

H6: Members whose ratios of best answer are high will share knowledge more in the community.
H7: Members whose ratios of best answer are high will set high Reward Points for others.
H8: Answerers have higher ratios of best answer compared with askers.

To investigate who are sharing high quality knowledge in the community, we use the following regression function to examine the relationship between quality of shared knowledge and members’ characteristics.

<table>
<thead>
<tr>
<th>Y=Best answer ratio</th>
<th>Computer Category</th>
<th>Entertainment</th>
<th>Whole Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²=0.098 F=173.50***</td>
<td>R²=0.098 F=197.38***</td>
<td>R²=0.086 F=242.9***</td>
<td></td>
</tr>
<tr>
<td>No. of Answers 0.001(0.000) t=7.922***</td>
<td>0.001(0.000) t=15.163***</td>
<td>0.001(0.000) t=14.857***</td>
<td></td>
</tr>
<tr>
<td>Reward Points sent out by certain ID 0.001(0.000) t=3.484***</td>
<td>0.005(0.001) t=8.255***</td>
<td>0.002(0.000) t=6.726***</td>
<td></td>
</tr>
<tr>
<td>Asker/Answerer (Asker=0; Answerer=1) 8.666(0.436) t=19.869***</td>
<td>7.110(0.470) t=15.144***</td>
<td>7.877(0.323) t=24.357***</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. The Boxplot of the Entertainment Category
Table 8: The correlation between Best Answer Ratio and Reward Points

The regression results show some interesting findings: first, members answered more questions shared high quality knowledge as well. Second, members who shared high quality knowledge sent out more Reward Points, which means that they value others’ sharing behavior more. Third, compared with askers, answerers’ shared knowledge has high quality compared with askers’.

Such results indicate that those answerers who frequently answer questions are the most valuable resources in the community. They not only answered most questions but answered with higher quality. Also, we may envision a positive circulation formed in the community. Answerers who share high quality knowledge will be acknowledged by others and have high Best Answer Ratio. In return, those members are motivated due to the acknowledgment by other members, and thus will answer more questions. Further more, since they understand the efforts involved in sharing knowledge, they will send out high Reward Points to others when they propose their questions.

6. CONCLUDING REMARKS

This paper uses empirical data that we collected from a spider engine to investigate user profile, knowledge sharing motivations and quality of sharing knowledge in online communities. We found two groups of people are extremely important in driving the growth of these communities. First, the top contributors unselfishly answer other questions, contributing their own knowledge. Among these people, there are approximately 85% are pure contributors, seldom asking any questions. The second group of users are those who seldom answer questions, but always ask questions. Although their behavior seems selfish at the surface, they are actually the key factors driving the growth of online knowledge sharing communities. It is precisely because of these users’ questions that the online communities see reasons of existence.

Majority of free riders appear to be those non-registered members outside our sample. In order to spur further growth, organizers of online knowledge sharing communities may think of ways to have these free riders join the community.

We also investigated motivations of contributing knowledge. Although Reward Points and reciprocal-supportiveness are clear factors explaining knowledge sharing behavior, there are also many members who do not care about rewards. Our conjecture is that these members derive satisfaction from helping others as a motivation of contributing knowledge.

Unlike previous results, we found that members’ past behavior does not seem to be acknowledged by others. That is to say, those who have a history of helping others, do not seem to get above-average help from others when they need help. This may be because knowledge sharing communities are simply too large with too many members, and the bonding relationship seems to exist between members and the community itself, as opposed to between members. When users decide whether to answer a question, he may be more concerned about his relationship with the community, instead of the ID identity of the author of the question.

We also found that the value of shared knowledge sees a positive correlation with past behavior. Those members who historically contributed a lot tend to award members higher who answer their questions. These people are also the core members driving the growth of online communities.

Our study confirms the intrinsic and extrinsic motivations of knowledge contribution based on real member usage data, supplementing the survey methods that have been traditionally used in this type of studies. In future research, we plan to incorporate data from the Yahoo! Answers, to investigate different user behavior under the different reward mechanisms. We conjecture that there might be also different cultural factors in driving knowledge sharing behavior.
References


Use of Blogs for Collaboration in Organizations

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0737.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business value of IT, Commitment-trust theory, Knowledge Management Systems, knowledge transfer</td>
</tr>
</tbody>
</table>
USE OF BLOGS FOR COLLABORATION IN ORGANIZATIONS

Sundaravej, Thanaporn, University of Missouri – St. Louis, One University Boulevard, St. Louis, MO 63121, USA, sundaravejf@umsl.edu

Cheng, Jiesi, University of Arizona, 1130 E. Helen St., Tucson, Arizona 85721, USA, jiesi.cheng.js@gmail.com

Mirchandani, Dinesh, University of Missouri – St. Louis, One University Boulevard, St. Louis, MO 63121, USA, mirchandanid@umsl.edu

Abstract

Blogs are increasingly being used as collaboration tools in organizations. However the quality and accuracy of the posted messages in them are still causes for concern. This study proposes a research model to investigate the factors affecting trust in posted messages and the influence of such trust on knowledge use. It suggests that users’ intention to use trusted messages can act as a moderator variable in the relationship between trust and use of the collaboration tool.

Keywords: Blogs, Collaboration, Trust, Knowledge Use
1 INTRODUCTION

Organizations are increasingly adopting online collaboration tools as platforms where data, information, and knowledge about a particular topic can be shared among employees (Gordon, 2006). Blogs are relatively new forms of organizational online communication tools that are being increasingly used to promote internal collaboration as well as enhance knowledge sharing and use in organizations. However, the quality of messages shared on them is claimed to be average because some visitors may contribute useful information, while others may post irrelevant responses (Wen and Wen, 2006). The accuracy of posted message is also a critical concern (Cherry, 2003; Denning et al., 2005) because the messages may contain subjective opinions of the posters. The quality and accuracy of exchanged information has been shown in prior research to have a significant effect on user understanding and perceived effectiveness of communication (Daft and Lengel, 1986). In fact, the quality and accuracy of content on a Web site is a major factor influencing users’ trust towards the site (Mithas et al., 2007).

Therefore this research seeks to investigate what makes people trust the messages posted on Blogs and the effect this trust has on knowledge use. Understanding these relationships can be useful in redesigning the tools to achieve their maximum effective usage as well as justifying their adoption in organizations as a means to increase collaboration among employees.

The term Blog (i.e., Weblog) refers to an automated, personalized, and community-supported Web page, kept in reverse chronological order (Du and Wagner, 2007). It allows individuals to submit and share text, images, or video with others through the Internet or even a mobile phone (Bamford et al., 2006; Cemerlang et al., 2006). Common features of Blogs include searching, posting, and commenting on its messages. They differ from forums, message boards, newsgroups, and Wikis in that only the Blog owner is permitted to post entries, whereas readers are restricted to only comment on the posted entries (Gordon, 2006). The additional comments are appended to each posted message (Torrey et al., 2007). In terms of patterns to use, a Blog usually links to other Blogs in the same category or interest area.

Empirical research on Blogs has been limited. Although an initial study suggested that the primary purpose of using a Blog was as a personal journal, and its use for collaboration was fairly rare (Herring et al., 2004), more recent studies have shown the positive impact of Blogs on collaboration among users. Du and Wagner (2007) empirically examined the relationship between usage of Blogs for online learning and student learning performance. Their results suggest that effective use of Blogs promotes the constructivist models of learning by supporting both cognitive and social knowledge construction. Lin et al. (2006) studied the implementation of a Blog system in an international distance education course. Their findings suggest that Blogs are effective tools to document the student’s learning, to share experience and knowledge, and to have direct interaction with long distance peers.

2 THEORETICAL DEVELOPMENT

Several researchers have investigated the association of social networks with trust. Strength of ties is an aspect of social networks that correlates to a degree of trust. It can be measured by various criteria such as the length of time two actors spend together, the emotional intensity between two actors, the degree of trust and acquaintance, the frequency of interaction, or the time of relationship (Kristiansen, 2004). Strong ties bind actors who likely to be connected (Grannovetter, 1983). They exist for people with acquaintance such as family members, close friends, co-workers, etc. This study focuses on the strength of ties between the poster and reader of messages on corporate Blogs.

Many studies have shown that strong ties are important channels in transferring useful knowledge (Brass and Burkhardt, 1993; Droege and Anderson, 2003). Relationships between actors with strong ties, as a result, are beneficial. Levin and Cross (2004) claimed that when trust exists, people are more willing to give useful knowledge and also more willing to listen to and absorb others’ knowledge.
McGrath et al. (2003) found that software executives who were looking for advice tended to request people they knew well mostly from relationships formed through previous work settings, educational settings, or mutual contacts. In the current study, we measure the message reader’s perception about the relationship between the reader and the poster. It can be expected that messages posted on an online collaboration tool by an actor with strong ties to the reader will increase the reader’s trust on the message. Thus we hypothesize:

**H1: Greater strength of ties with message posters increases readers’ trust on posted messages.**

Reputation has been studied by many researchers in different perspectives. Barnett et al. (2006) defined corporate reputation as observers’ collective judgments of a corporation based on assessments of the financial, social, and environmental impacts attributed to the corporation over time. Helm (2005) proved that corporate reputation involved a variety of different measures and it correlated to stakeholders’ loyalty towards the company. Johnson et al. (1993) analyzed a relation between executive’s reputation and firm financial performance. They defined executive’s reputation as the expected value of true managerial ability conditional on all previous productivity observations.

Studies on reputation and trust have recently become important in information systems research, especially in the context of electronic markets. Reputation helps to build trust and cooperation among loosely connected and geographically dispersed economic agents (Dellarocas, 2006).

Reputation in this study refers to overall quality or character of the message poster as seen or judged by people in organization. In other words, reputation of the message poster is perceived as recognition by other people of some characteristics or abilities. According to the prior studies, it can be expected that messages on an online collaboration tool posted by a person with high reputation in an organization will increase the reader’s trust on the message. Thus:

**H2: Reputation of message posters increases readers’ trust on posted messages.**

Power has been researched for many years in multiple disciplines. An interesting study on interpersonal power was conducted by Garrison and Pate (1977). They measured power in different contexts, including the co-worker context. A co-worker is defined as someone a person has communicated with on a regular basis in the present job. The results of the study showed that perceptions of interpersonal power vary across the communication context. In the current study, power is conceived as relationships between two or more actors in which the behavior of one is affected by the behavior of the other (Jasperson et al., 2002). In other words, power is a perceived or attributed characteristic possessed by a social actor with a definite potential or capacity which may be used to influence or even control the behavior of others in some manner within some contexts (Brill, 1992). A person who holds more power can draw on rules and uses techniques of production and discipline through social relation (Backhouse et al., 2006). Regarding trust in online collaboration, it can be expected that messages posted on an online collaboration tool by a person who asserts more power will increase the reader’s trust on the message. Hence we hypothesize:

**H3: Power of message posters increases readers’ trust on posted messages.**

Apart from the characteristics of the message poster as mentioned above, features and elements used in posted messages are believed to affect trust and eventually the knowledge use of an online collaboration tool. In recent years, several studies have shown that Web site features and elements such as graphics, colors, links, tabular forms, frame location, background, navigation, organization, uniformity, etc. influence the beliefs and behaviors of Web customers (Song and Zahedi, 2005), user efficiency and frustration level (Pratt et al., 2004), and perceptions of customer service, information privacy, and product quality (Resnick and Montania, 2003). Interestingly, Everard and Galletta (2006) studied how Web site presentation flaws affect perceived site quality, trust, and intention to purchase from an online store. The flaws included poor aesthetics, quality, and reliability of a Website. Each of these flaws was found to prevent users from using the system in a meaningful manner.

Aesthetic aspects of a Web site should possess the same priority as the functionality aspects in Web design (Thorlacius, 2002). The organization of a Web site was found to have the greatest impact on the users’ experience visiting the site (Lynch and Horton, 2002). Several prior studies suggested that
customers are more likely to enjoy orderly Web sites due to ease of use and navigation (Song and Zahedi, 2005). In contrast, the poor style of Web sites was found to be inversely related to the users’ level of perceived quality, trust, and intentions to purchase from an online store (Everard and Galletta, 2006). The elements of a Web site’s poor style are graphical and visual elements such as background, font size, columns, words, line spacing, etc that interfere with reading. As a result, in this study, it is assumed that a poor style of posted messages on the online collaboration tool such as an unusual background color, font style, font size, words, line spacing, or frame layout, will reduce the reader’s trust on the message. Thus:

**H4: Poor style of posted messages decreases readers’ trust on posted messages.**

Incompleteness of a Web site is another factor that discourages confidence and trust of visitors (Everard and Galletta, 2006). Incompleteness refers to a Web site that includes obsolete content or links, non-loading pictures, or missing information. Information provided on a Web site should be accurate, detailed, up-to-date, and relevant (Li et al., 2002). A Website with high information quality provides valuable and helpful knowledge, leading to a larger audience (Chiu et al., 2005). On the contrary, incomplete sites can be risky and lead to negative impression of visitors. In online transactions, information incompleteness makes it difficult for buyers to assess sellers’ true characteristics and product quality, resulting in uncertainty perceptions about the transaction (Pavlou et al., 2007). For the current study, incompleteness of posted message on the online collaboration tool is likely to cause readers to become disappointed and not return. It can be expected that incompleteness of a posted message on the online collaboration tool will reduce the reader’s trust on the message. Thus:

**H5: Incompleteness of posted messages decreases readers’ trust on posted messages.**

A flaw that affects perceived site quality, trust, and intention to purchase from an online store includes typographical, grammatical, and factual errors (Everard and Galletta, 2006). Similar to incompleteness of Web sites, such errors distract users and make them suspect a generally poor quality of the system. Spelling errors may interfere with the user’s perception of the reliability of the information provided on the Web site (Liu and Ginther, 2001). Credibility in online communication is significant. It makes a competitive difference in an organization’s success (Booher, 2004). An experimental study on the effects of support conforming use of spelling and grammar in Computer-Mediated Communication (CMC) system found an increase of message persuasiveness (Wilson, 2005). In the current study, it is expected that language errors in a posted message on an online collaboration tool will reduce the reader’s trust on the message. Hence we hypothesize:

**H6: Errors of posted messages decrease readers’ trust on posted messages.**

Several studies have demonstrated that trust is a significant factor for organizational success. An organization that has a high level of trust environment promotes the exchange of valuable ideas between workers that will, in turn, leads to better firm performance (Collins and Smith, 2006). When trust among employees is low, individuals will be careful about exchanging information and ideas with one another. Trust has been shown to be a major factor in facilitating successful knowledge sharing (Kumar and Thondikulam, 2006) and R&D collaboration performance (Hurmelinna et al., 2005). Several prior studies, additionally, show that trust influences the intention to use or usage of information systems (IS) (Son et al., 2005; Kankanahalli et al., 2005; Salam et al., 2005).

**H7: Trust on posted messages increases readers’ use of the Blog tool.**

In the current study, intention to use trusted knowledge is defined as a degree of which a knowledge worker believes he or she has incorporated procedures for the capture and use of knowledge of various types into decision-making activities, routine, and otherwise and intends to use it (Kulkarni et al., 2006). Based on the findings of prior research, it can be expected that greater intention to use trusted knowledge will lead to a stronger relationship between trust and use of the tool whereas lower intention to use the knowledge will lead to a weaker relationship between trust and use of the tool.

**H8: Intention to use knowledge moderates the relationship between trust and knowledge use.**

The Research Model is shown in Figure 1.
3 RESEARCH METHOD

The study will utilize a Web-based field survey to collect data from organizations that are using Blogs for knowledge sharing among co-workers. Given that prior research on Blogs has largely relied on anecdotal information and not explored their use for organizational collaboration, the current research will aid further understanding on the subject.

The research questionnaire will include questions about (1) the collaboration tool used in the company; (2) trust in posted messages; (3) knowledge use; and (4) demographics. The questionnaire items have been derived from several prior studies in different disciplines such as psychology, IS, and management (Kulkarni et. al., 2006; Everard and Galletta, 2005; McKnight et. al., 2002; Levin and Cross, 2004; Evald et. al., 2006; Caruana, 2006; Garrison and Pate, 1977) with minor modifications.
made to fit the study’s context. Trust and knowledge use are measured using 5-point Likert scales. The questionnaire instructions, items, and sources can be found in Appendix A.

Ten organizations in China and Thailand that are currently using Blogs for collaboration will be contacted to participate in the study due to their accessibility to the researchers. A pilot study has been conducted with two subjects from different companies one based in China and the other in Thailand to assess the duration of a typical survey and to ensure that the questions would be correctly interpreted by the subjects. The first participant in the pilot recommended a translation of the questionnaire into Chinese to accommodate Chinese speakers, reduce the survey duration, and increase the response rate. The second participant recommended minor changes to the instrument which have been duly incorporated. The questionnaire has been translated to Chinese following the procedure recommended by Brislin (1986). One of the current study’s authors whose native language is Chinese translated the English questionnaire into Chinese and this translation was verified for accuracy by two other native Chinese speakers. Next, three different native Chinese speakers conversant with the English language, independently translated the revised Chinese version back to English. After the back-translation, the study’s authors compared the three back-translated English questionnaires with the original English version. No significant differences were found. Therefore, the revised Chinese version questionnaire will be posted online along with the English version for data collection. The data gathered will be analyzed using PLS-Graph 3.0 to test construct validity and the research hypotheses. Controls will be employed for the organization’s size, industry and national culture.

4 EXPECTED CONTRIBUTIONS AND LIMITATIONS

We believe that our research will be among the first studies to use quantitative methods to measure trust and knowledge use in organizational online collaboration tools such as Blogs. Because Blogs are not universally used in organizations, we believe our research could have implications for both early adopters and potential adopters. The results of the research could help early adopters better understand the factors that influence trust and knowledge use. Therefore, they could improve productivity within their organization by facilitating greater trust and knowledge use. The result will also help potential adopters understand the benefits of using Blogs, and how to implement them to maximize trust and knowledge use. Limitations in the current design are that the study utilizes a field survey with perceptual measures and single respondents. In-depth case studies could provide a richer context given that respondents may not accurately remember their social relations or simply provide socially acceptable answers to the questionnaires. However, given the limited research on the subject we believe our findings will inform future research that could utilize qualitative methods, objective measures and multiple respondents to overcome any response bias issues.
References


Appendix A

How Often Do You Use Your Organization’s Blog tool:
___ Hours per Day  ___ Times per Week

To what extent do you agree with the following statements related to messages posted on your corporate collaborative environment?

(1=Strongly disagree, 2=Disagree, 3=Neither agree/nor disagree, 4=Agree, 5=Strongly agree)

<table>
<thead>
<tr>
<th>Construct</th>
<th>items</th>
<th>source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention to use knowledge</td>
<td>a1 If I trust the posted message, I will apply the shared knowledge to my work.</td>
<td>Kulkarni, Ravindran, and Freeze 2006</td>
</tr>
<tr>
<td></td>
<td>a2 If I trust the posted message, I will share the knowledge to coworkers.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a3 If I trust the posted message, I will refer to the shared knowledge in future work.</td>
<td></td>
</tr>
<tr>
<td>Trust (on posters)</td>
<td>b1 I usually trust people until they give me a reason not to trust them.</td>
<td>Everard and Galletta, 2005</td>
</tr>
<tr>
<td></td>
<td>b2 When I first meet someone, I generally accept that he/she is telling the truth, even though I do not necessarily believe him/her.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b3 My typical approach is to trust new acquaintances until they prove I should not trust them.</td>
<td></td>
</tr>
<tr>
<td>Trust (on posted messages)</td>
<td>b4 I would trust a posted message if I feel it gives information about the particular topic very well.</td>
<td>McKnight et al., 2002</td>
</tr>
<tr>
<td></td>
<td>b5 I would trust a posted message if I feel assured that it is very knowledgeable about the particular topic.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b6 I would trust a posted message if I have confidence in it.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b7 I would trust a posted message if I believe it is trustworthy.</td>
<td></td>
</tr>
</tbody>
</table>

To what extent would you trust a message posted on your corporate collaborative environment (1= Very little extent, 2= Little extent, 3= Neutral, 4= Great extent, 5= Very great extent).

<table>
<thead>
<tr>
<th>Construct</th>
<th>items</th>
<th>source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strength of Ties</td>
<td>if you feel awkward talking to the message poster about a non-work related problem.</td>
<td>Levin and Cross, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>c2</strong></td>
<td>if you know the message poster well outside the work related areas.</td>
<td>Evald, Klyver, and Svendsen, 2006</td>
</tr>
<tr>
<td><strong>c3</strong></td>
<td>if you have had a long positive relationship with the message poster.</td>
<td></td>
</tr>
<tr>
<td><strong>c4</strong></td>
<td>if you have frequent interaction with the message poster.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>d1</strong></td>
<td>if the message poster can be relied upon to be sincere.</td>
<td>Stewart, 2006</td>
</tr>
<tr>
<td><strong>d2</strong></td>
<td>if promises and commitments made by the message poster are likely to be kept.</td>
<td></td>
</tr>
<tr>
<td><strong>d3</strong></td>
<td>if the message poster is trustworthy.</td>
<td></td>
</tr>
<tr>
<td><strong>d4</strong></td>
<td>if the message poster is well known in your team, department, or organization.</td>
<td>Caruana et al., 2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>e1</strong></td>
<td>if the message poster has a great deal of influence over your behavior.</td>
<td>Garrison and Pate, 1977</td>
</tr>
<tr>
<td><strong>e2</strong></td>
<td>if the message poster is a powerful person.</td>
<td></td>
</tr>
<tr>
<td><strong>e3</strong></td>
<td>If you see the message poster as a leader.</td>
<td></td>
</tr>
<tr>
<td><strong>e4</strong></td>
<td>if you find the message poster to be a very persuasive person.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>f1</strong></td>
<td>if the posted message looks rather unattractive.</td>
<td>Everard and Galletta, 2005/2006</td>
</tr>
<tr>
<td><strong>f2</strong></td>
<td>if the posted message is hard to read.</td>
<td></td>
</tr>
<tr>
<td><strong>f3</strong></td>
<td>if the posted message is rather messy.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>g1</strong></td>
<td>if some pictures, graphs, or figures are not visible.</td>
<td>Everard and Galletta, 2005</td>
</tr>
<tr>
<td><strong>g2</strong></td>
<td>if some parts of posted messages are unfinished.</td>
<td></td>
</tr>
<tr>
<td><strong>g3</strong></td>
<td>if the posted message has some missing information.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>h1</strong></td>
<td>if there are typographical errors in the posted message.</td>
<td>Everard and Galletta, 2005</td>
</tr>
<tr>
<td><strong>h2</strong></td>
<td>if some words are misspelled in the posted message.</td>
<td></td>
</tr>
<tr>
<td><strong>h3</strong></td>
<td>if there are spelling errors in the posted message.</td>
<td></td>
</tr>
</tbody>
</table>
PROBLEM SOLVING PATTERNS IN DESIGN SCIENCE RESEARCH – LEARNING FROM ENGINEERING

Gericke, Anke, Institute of Information Management, University of St. Gallen, Müller-Friedberg-Strasse 8, 9000 St. Gallen, Switzerland, anke.gericke@unisg.ch

Abstract

Within information systems the design science research (DSR) paradigm aims at the development of useful artifacts, e.g. models or methods, with which relevant IS problems can be solved. In analogy to the engineering discipline construction processes have been proposed for DSR. Although different phases of such construction processes are explicated in several articles, contributions are missing that propose patterns/principles that support the constructor during the different phases of the construction process. Vaishnavi and Kuechler (2007) address this issue by proposing DSR patterns. Their contribution is a substantial one; however, it does not include comparable pattern approaches from the engineering discipline for the foundation of the proposed patterns. Bearing in mind that DSR has its roots in engineering, it is important to analyze so called problem solving patterns from engineering and to compare them to the DSR patterns. Using this as a basis, it is our research goal to examine whether it is possible to expand the existing DSR patterns to include patterns from engineering. As a result, 14 additional DSR patterns are proposed which originate from engineering, have not been discussed so far, but promise to be useful for DSR in information systems.

Keywords: Design Research, Design Science, Research Methodology
1 INTRODUCTION

1.1 Problem Statement

The information systems (IS) discipline differentiates two main research paradigms: behavioral research and design science research (DSR) (Hevner et al. 2004, p. 76). In contrast to behavioral research that focuses on the development of theories, DSR is a problem solving paradigm which does not only have its roots in the sciences of the artificial (Simon 1996) but also in the engineering discipline (Hevner et al. 2004, p. 76). The goal of the DSR discipline is the development of useful artifacts with which IS-related problems can be solved (March & Smith 1995, p. 253). Within DSR the artifact types of March & Smith (1995, p. 256 ff.), i.e. constructs, models, methods and instantiations, have been established as artifacts of the DSR discipline (e.g. cf. Hevner et al. 2004, Vahidov 2006, vom Brocke & Buddendiek 2006). Lately, design theories have been discussed as DSR artifacts as well (e.g. cf. Kuechler & Vaishnavi 2008, Venable 2006b).

For the development of such artifacts, construction processes have been proposed in analogy to the engineering discipline (e.g. cf. Pahl et al. 2007, p. 53). In recent years, the construction process developed by March & Smith (1995) has achieved wide acceptance (e.g. cf. Cao et al. 2006, Hevner et al. 2004, Venable 2006a). This construction process consists of a “build” (develop artifact) and an “evaluate” phase (evaluate artifact) (March & Smith 1995, p. 258 ff.). In addition, articles have been published (e.g. cf. Peffers et al. 2006, Rossi & Sein 2003, vom Brocke & Buddendiek 2006) that detail these phases, e.g. by explicitly defining an “identify a need” phase (Rossi & Sein 2003) which has to be conducted prior to the development of the artifact. Besides, Hevner et al. (2004, p. 82 ff.) propose seven DSR guidelines that assist researchers and reviewers to understand the requirements for effective DSR. Moreover, the literature analysis shows that some articles also put their focus on solution patterns, i.e. patterns that represent parts of the result of the construction process, but do not support the construction process itself. To give an example, Schermann et al. (2007) propose three patterns that form a design theory for IT service data management systems (result of the construction process).

It is doubtless that the identified contributions are very useful. However, there are hardly any contributions that contain patterns/principles guiding the constructor within the different phases of the construction process in order to solve the research problem through the development of a DSR artifact. Within our literature analysis we only identified the contribution of Vaishnavi & Kuechler (2007) who propose DSR patterns that support the constructor in each phase of the construction process, e.g. the “build” or the “evaluate” phase. In addition to the construction process further phases of the whole research process are supported as well. To give an example, Vaishnavi & Kuechler (2007) identified patterns for the “conclusion” phase, which assist researchers in writing up and publishing their results. Next to patterns that are assigned to a certain phase of the construction/research process, they also proposed so called meta patterns, such as “Brain Storming” or “Stimulating Creativity”, that support more than one or even all construction/research phases. Although the contribution of Vaishnavi & Kuechler (2007) is a substantial one, the underlying foundation of the identified patterns has not been made visible in their publication. Based on a statement of Vaishnavi, both authors have engineering backgrounds but did not use concrete engineering approaches for the foundation of the DSR patterns. Hence, for the proposition of their patterns DSR literature was used without including comparable approaches from engineering. Due to the fact that DSR has its roots in engineering – as e.g. stated by Hevner et al. (2004, p. 76) – it is important to analyze so called problem solving patterns from engineering and to compare them to the DSR patterns of Vaishnavi & Kuechler (2007). The identification of analogous patterns provides a solid foundation and improves the validity of the corresponding DSR

1 E-mail conversation with Vijay K. Vaishnavi between 1 Oct 2008 and 29 Oct 2008.
patterns as they are backed up by their root discipline. This comparison is following the line of arguments developed in (Gericke 2009). Based thereupon, in this paper it is our research goal to examine whether it is possible to expand the existing DSR patterns to include patterns from engineering.

1.2 Research Methodology

In the paper at hand, an argumentative analysis is used as research method to address the proposed research goal. In order to develop new DSR patterns based on patterns from the engineering discipline convincing arguments will be derived from literature/relevant research and presented in a logical as well as comprehensible order. Due to space limitations of this paper we focus on the “build” phase.

Before new DSR patterns can be developed, the DSR patterns of Vaishnavi & Kuechler (2007) and pattern approaches from the engineering discipline need to be analyzed. Therefore, a structuring approach already successfully used in both disciplines is used (for engineering e.g. cf. (Günzler & Vilbig 2003, Jarke et al. 2003), for method engineering (a sub-discipline of DSR) e.g. cf. (Brinkkemper 1996)): Following this approach the object or matter under consideration is structured regarding a product view and a process view. Applying this structuring approach to problem solving patterns implies, that patterns referring to the product, i.e. the result of a construction process, (“product view”) and patterns supporting the construction process itself (“process view”) can be differentiated.

Following the described research goal and research methodology we structured our paper as follows: In the second section we describe and analyze problem solving approaches/patterns from the engineering discipline. Thereafter, we introduce and analyze the DSR patterns of Vaishnavi and Kuechler (2007) in detail. In section 4 the results of the comparison of problem solving patterns from engineering and the DSR discipline are presented. Using this as a basis, we attempt to expand the DSR patterns by problem solving patterns from engineering. The paper closes with a summary and an outlook.

2 PROBLEM SOLVING PATTERNS IN ENGINEERING

In engineering it was already recognized in the beginning of the last century that the trial-and-error method is not the most efficient way to develop a problem solution (Altschuller 1986, p. 12, Orloff 2006, p. 34). Instead problem solving approaches should be used in order to reduce the number of errors and to solve problems more efficiently (Altschuller 2005, p. 36, Creţu 2007, p. 7, Teufelsdorfer & Conrad 1998, p. 14). Realizing this, Altschuller developed the TRIZ approach, which is the Russian acronym for the “Theory of Inventive Problem Solving” (TIPS) (e.g. cf. Altschuller 2005, Altschuller 2006, Altschuller & Shulyak 2002, Orloff 2006). This approach is well established – in academia as well as in industry (Altschuller 2005, p. 15 ff., Herb et al. 1998, p. 18). Many other approaches addressing inventive problem solving in engineering, such as WOIS (German acronym for contradiction-oriented innovation strategy), the PI concept (Concept of Problem-Oriented Invention) or SIT (Systematic Inventive Thinking), are based thereupon (Pannenbäcker 2007, Teufelsdorfer & Conrad 1998, p. 10). Beside TRIZ, other problem solving approaches have been developed as well (e.g. cf. Hürlimann 1981, Kelley 2003). However, they are either single contributions and/or did not achieve wider acceptance. That is why we restrict our focus to the TRIZ approach for the analysis at hand.

“TRIZ is a comprehensive, systematically organized invention knowledge and creative thinking methodology” (Creţu 2007, p. 8). The methods and concepts belonging to TRIZ can be divided into four independent groups (Gimpel et al. 2000, p. 7, Löbmann 2002): (1) Analysis (2) Knowledge (3) Analogy, and (4) Vision. The analysis group contains methods that are used to analyze problems and to overcome mental blocks. The second group refers to knowledge bases, e.g. a scientific effects database. The third group deals with analogies by containing different solution principles/patterns, e.g. 40 innovation principles used to overcome technical contradictions. Finally, the vision group describes development trends and contains e.g. the S-curves of evolution.
The 40 innovation principles belonging to the analogy group are one of the most famous concepts of TRIZ (Rietsch 2007, p. 14). They are patterns which support the constructor in solving technical problems efficiently and effectively (Chen & Lin 2008, p. 14, Rietsch 2007, p. 14 f.). The 40 innovation principles are empirically well founded, because they were derived from more than 40,000 reviewed and analyzed patents of inventions (Altschuller 2005, p. 137). Knowing this, the question arises whether these principles can be transferred to DSR since the DSR discipline cannot build on a comparable empirical basis which can be used to derive such principles/patterns. Though, next to their use in the engineering discipline these principles have already been transferred to other disciplines using conclusions by analogy. Examples can be found for management (Ruchti & Livotov 2001), marketing (Pustogow 2007), human resource management (Müller 2006), etc. Following this argumentation we transfer the innovation principles of TRIZ to the DSR discipline using conclusions by analogy. Regarding the transfer of the 40 innovation principles to other disciplines, Zobel (1991, p. 111) recognizes that the principles are not coequal, but that some principles are more universal than others. Following him such universal principles can be transferred to other disciplines whereas the remaining specific principles are of a rather technical nature and are less suitable to be transferred to other disciplines (Bannert & Warschat 2007, p. 64, Zobel 1991, p. 114 f.). That is why only 22 universal innovation principles will be considered in the following (Zobel 1991, p. 114 f.).

To give an overview of all 40 innovation principles, Table 1 contains their names and gives an exemplary description for every principle. A number (E1, E2, etc.) is assigned as well in order to ease recognition in the remainder of the paper. The last two columns refer to the analysis of the patterns which is presented at the end of this section. The 18 specific principles are shaded in grey and will not be considered in the remainder of this paper.

<table>
<thead>
<tr>
<th>No.</th>
<th>PSP¹</th>
<th>Exemplary Descriptions</th>
<th>Prod.²</th>
<th>Proc.³</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>Segmentation</td>
<td>Divide an object into independent parts.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>E2</td>
<td>Extraction</td>
<td>Extract the “disturbing” part or property from an object or extract only the necessary part or property from an object.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>E3</td>
<td>Local Quality</td>
<td>Transition from homogeneous to heterogeneous structure of an object. Different parts of an object should carry out different functions. Each part of an object should be placed under conditions that are most favorable for its operation.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>E4</td>
<td>Asymmetry</td>
<td>Change the shape of an object from symmetrical to asymmetrical.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>E5</td>
<td>Merging/Consolidation</td>
<td>Bring closer together (or merge) identical or similar objects, assemble identical or similar parts to perform parallel operations. Make operations contiguous or parallel; bring them together in time.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>E6</td>
<td>Universality</td>
<td>Make a part or object perform multiple functions; eliminate the need for other parts.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>E7</td>
<td>Nesting</td>
<td>Place one object inside another; place each object, in turn, inside the other.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>E8</td>
<td>Counterweight</td>
<td>To compensate for the weight of an object, merge it with other objects that provide lift or make it interact with the environment (e.g. use aerodynamic, hydrodynamic, and other forces).</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>E9</td>
<td>Prior Counteraction</td>
<td>If it will be necessary to do an action with both harmful and useful effects, this action should be replaced with counteractions to control harmful effects. Create beforehand stresses in an object that will oppose known undesirable working stresses later on.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>E10</td>
<td>Prior Action</td>
<td>Perform, before it is needed, the required change of an object (either fully or partially). Pre-arrange objects such that they can come into action from the most convenient place and without losing time for their delivery.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>E11</td>
<td>Cushion in Advance</td>
<td>Prepare emergency means beforehand to compensate for the relatively low reliability of an object.</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

2 Zobel (1991, p. 114 f.) originally characterized 23 of the 40 innovation principles as universal. Due to the fact that the principle E18 “Mechanical Vibration” seems to be very specific for the engineering discipline we assign it to the category of specific principles and do not consider it any further.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>E12</td>
<td>Equipotentiality</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E13</td>
<td>Do it in Reverse</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E14</td>
<td>Spheroidality, Curvature</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E15</td>
<td>Dynamics</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E16</td>
<td>Partial or Excessive Action</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E17</td>
<td>Another Dimension</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E18</td>
<td>Mechanical Vibration</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E19</td>
<td>Periodic Action</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E20</td>
<td>Continuity of Useful Action</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E21</td>
<td>Rushing Through</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E22</td>
<td>Convert Harm Into Benefit</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E23</td>
<td>Feedback</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E24</td>
<td>Intermediary</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E25</td>
<td>Self-Service</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E26</td>
<td>Copying</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E27</td>
<td>Cheap Short-Living Objects</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E28</td>
<td>Mechanics Substitution</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E29</td>
<td>Pneumatics/ Hydraulics</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E30</td>
<td>Flexible Shells/ Thin Films</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E31</td>
<td>Porous Materials</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E32</td>
<td>Color Changes</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E33</td>
<td>Homogeneity</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E34</td>
<td>Discarding/ Recovering</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E35</td>
<td>Parameter Changes</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E36</td>
<td>Phase Transition</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E37</td>
<td>Thermal Expansion</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E38</td>
<td>Strong Oxidants</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>E39</td>
<td>Inert Atmosphere</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Problem Solving Patterns in Engineering – The TRIZ Approach

<table>
<thead>
<tr>
<th>No.</th>
<th>PSP¹</th>
<th>Exemplary Descriptions</th>
<th>Prod.²</th>
<th>Proc.³</th>
</tr>
</thead>
<tbody>
<tr>
<td>E40</td>
<td>Composite Materials</td>
<td>Change from uniform to composite (multiple) materials.</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

¹ Problem Solving Pattern, ² Product View, ³ Process View

Table 1. Problem Solving Patterns in Engineering – The TRIZ Approach

After presenting the 40 innovation principles of TRIZ, we analyze them regarding the proposed structuring of “product view” and “process view”. Studying the descriptions of the principles, e.g. E1 – “Divide an object into independent parts.” or E2 – “Extract the ‘disturbing’ part from an object …” (see Table 1), it becomes obvious that the innovation principles basically refer to the object, i.e. the result/product, of the construction process. Hence, the innovation principles of the TRIZ approach are patterns possessing a product view, as they refer to the product/artifact which will be developed in the construction process (see fourth column in Table 1). However, looking from a process perspective it can be realized that the innovation principles can also be interpreted as actions that have to be conducted in a construction process. That is why all innovation principles are also marked with an “X” in the “process view” column in Table 1.

3 PROBLEM SOLVING PATTERNS IN THE DESIGN SCIENCE RESEARCH DISCIPLINE

In DSR Vaishnavi & Kuechler (2007) were the first proposing problem solving patterns. Amongst patterns for other phases of the construction process, they identified 27 patterns that support the constructor in the “build” phase of the construction process. Each of these patterns is characterized by a name and described in various dimensions: (A) For each pattern the intended field of use is described. (B) This is followed by a short statement about the context and applicability of the pattern and (C) a description about how to use it. (D) This goes together with a short explanation about the consequences of the usage of the pattern. Most pattern descriptions are accompanied by (E) examples and (F) lists of sources/references. (G) Sometimes referrals to related patterns are stated as well.

To make an example, the pattern “Approaches for Building Theory” (DSR2) will be explained according to the above listed dimensions (Vaishnavi & Kuechler 2007, p. 122 f.). Due to space limitations (F) sources and references are omitted and the description of (E) examples is reduced to one example; referrals to (G) related patterns are not available for this pattern. This pattern is (A) intended to encourage researchers to obtain a general understanding of the different approaches for developing theories. This research pattern can be used after identifying and developing the research problem ((B) context and applicability). In order to develop theories the researcher can choose from four different general approaches: (1) hypothetical/deductive, (2) prototyping (hermeneutical/inductive), (3) case-based and (4) historical, whereas each general approach is described in more detail in the book of Vaishnavi & Kuechler (2007) ((C) description). As a (D) consequence of the use of this pattern, the researcher should assess the suitability of the taken general approach based on the research problem and research area. Maybe corrective actions are necessary in such a way that approaches have to be combined or completely new approaches have to be taken into consideration. As one (E) example the research of Chen (1976) is referenced as he used the hypothetical/deductive approach to build theory.

Table 2 contains an overview of the 27 pattern of Vaishnavi & Kuechler (2007) supporting the “build” phase of the construction process. Each pattern is listed by its name and a short description referring to the intent of the pattern. Analogous to the innovation principles of TRIZ each pattern is accompanied by a number (DSR1, DSR2, etc.) and two columns indicating the product or process view of a pattern.

<table>
<thead>
<tr>
<th>No.</th>
<th>PSP¹</th>
<th>Description (Intent)</th>
<th>Prod.²</th>
<th>Proc.³</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR1</td>
<td>Theory Development</td>
<td>Explicitly state the theory that underlies the problem solution.</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

¹ Problem Solving Pattern, ² Product View, ³ Process View
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR2</td>
<td>Approaches for Building Theory</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR3</td>
<td>Hermeneutical and Inductive Approach</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR4</td>
<td>Incremental Theory Development</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR5</td>
<td>Problem Space Tools and Techniques</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR6</td>
<td>Research Community Tools and Techniques</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR7</td>
<td>Empirical Refinement</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR8</td>
<td>Easy Solution First</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR9</td>
<td>Elegant Design</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR10</td>
<td>Divide and Conquer with Balancing</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR11</td>
<td>Hierarchical Design</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR12</td>
<td>Building Blocks</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR13</td>
<td>Sketching Solution</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR14</td>
<td>Emerging Tasks</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR15</td>
<td>Modeling Existing Solutions</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR16</td>
<td>Combining Partial Solutions</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR17</td>
<td>Static and Dynamic Parts</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR18</td>
<td>Simulation &amp; Exploration</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR19</td>
<td>Interdisciplinary Solution Extraposition</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR20</td>
<td>Different Perspectives</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR21</td>
<td>General Solution Principle</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR22</td>
<td>Abstracting Concepts</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR23</td>
<td>Using Surrogates</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR24</td>
<td>Using Human Roles</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR25</td>
<td>Integrating Techniques</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR26</td>
<td>Technological Approach Exemplars</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DSR27</td>
<td>Means-End-Analysis</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

1 Problem Solving Pattern, 2 Product View, 3 Process View

Table 2. Problem Solving Patterns in the “Build” Phase in DSR (Vaishnavi & Kuechler 2007)

In analogy to the innovation principles of the TRIZ approach, we analyze the DSR patterns of Vaishnavi & Kuechler (2007) regarding the proposed structuring of “product view” and “process view” as well. Studying the first pattern DSR1 “Theory Development” (see Table 2), it becomes obvious that this is a pattern containing a product view because it suggests that next to the solution, i.e. a construct, model, method or instantiation, another product, i.e. the underlying design theory, should be
explicated. Equivalent to the TRIZ principles this DSR pattern can also be regarded as a pattern containing a process view, because it can also be interpreted as an action that has to be conducted in a construction process. Studying the second pattern DSR2 “Approaches for Building Theory” reveals that this pattern only possesses as process view as it tries to support the researcher in obtaining a general understanding of the different approaches for developing theories throughout the construction process. In contrast to the process view of the pattern DSR1 at which the action is directly conducted on the later solution/product of the construction process, the process view of DSR pattern 2 refers to actions that a researcher conducts within a construction process but that are only indirectly related to the later result. The complete analysis of all DSR patterns shows that all DSR patterns contain a process view that is directly or indirectly related to the final solution/result and guides the researcher through the construction process (see Table 2). Furthermore, 13 of these 27 DSR patterns possess a product view as well. Irrespective of being directly or indirectly related to the result of the construction process, patterns that contain a process view will be in the focus of our research. Due to space limitations the product view of the above listed patterns is not compared or expanded any further.

### 4 VERIFICATION AND EXPANSION OF DESIGN SCIENCE RESEARCH PATTERNS

#### 4.1 Design Science Research Patterns Verified by the Engineering Discipline

So far the DSR patterns of Vaishnavi and Kuechler (2007) that support the build phase have been based on DSR literature. In order to back them up by their root discipline, they were compared to the 22 universal TRIZ innovation principles (cf. Gericke 2009). Thereby the focus was put on DSR patterns that refer to the process view (see Table 2). Based on the comparison of DSR patterns and the universal innovation principles of the TRIZ approach in (Gericke 2009), Table 3 shows the results of that comparison: Nine DSR patterns could be verified by TRIZ innovation principles of the engineering discipline, which builds the roots of DSR. In Table 3 the first two columns present DSR patterns. The next two columns contain the corresponding TRIZ principles. Finally in the fifth column an explanation about the comparison of the DSR and the TRIZ pattern is given.

<table>
<thead>
<tr>
<th>No.</th>
<th>PSP DSR</th>
<th>No.</th>
<th>PSP TRIZ</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR5</td>
<td>Problem Space Tools/Techniques</td>
<td>E10</td>
<td>Prior Action</td>
<td>The identification of problem space tools and techniques can be interpreted as a preliminary action in the construction process.</td>
</tr>
<tr>
<td>DSR6</td>
<td>Research Community Tools/Techniques</td>
<td>E10</td>
<td>Prior Action</td>
<td>The identification of tools and techniques that the relevant research community uses for solving similar problems can be interpreted as a preliminary action in the construction process.</td>
</tr>
<tr>
<td>DSR7</td>
<td>Empirical Refinement</td>
<td>E23</td>
<td>Feedback</td>
<td>The results of an empirical observation which was conducted on an artifact developed beforehand can be interpreted as feedback and results in the refinement of the artifact.</td>
</tr>
<tr>
<td>DSR10</td>
<td>Divide &amp; Conquer with Balancing</td>
<td>E1</td>
<td>Segmentation</td>
<td>Dividing a problem into smaller problems of identical size is (partly) equivalent to dividing an object into parts (segmentation).</td>
</tr>
<tr>
<td>DSR11</td>
<td>Hierarchical Design</td>
<td>E1</td>
<td>Segmentation</td>
<td>Dividing a complex system into a hierarchy of sub-systems is (partly) equivalent to dividing an object into parts (segmentation).</td>
</tr>
<tr>
<td>DSR12</td>
<td>Building Blocks</td>
<td>E1</td>
<td>Segmentation</td>
<td>Dividing a problem into smaller problems of unequal size is (partly) equivalent to dividing an object into parts (segmentation).</td>
</tr>
<tr>
<td>DSR20</td>
<td>Different Perspectives</td>
<td>E17</td>
<td>Another Dimension</td>
<td>The TRIZ pattern “Another Dimension” is equivalent to the Vaishnavi/Kuechler pattern “Different Perspectives”.</td>
</tr>
</tbody>
</table>

Following a relaxed understanding of “product view”, the number of patterns that possess a product view as well can be reduced by three.
Comparison of Problem Solving Patterns in DSR and Engineering (TRIZ)

<table>
<thead>
<tr>
<th>No.</th>
<th>PSP DSR</th>
<th>No.</th>
<th>PSP TRIZ</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR21</td>
<td>General Solution Principle</td>
<td>E6</td>
<td>Universality</td>
<td>An object that performs multiple functions (universality) is comparable to the “General Solution Principle” aiming at the development of a general solution for a class of problems.</td>
</tr>
<tr>
<td>DSR23</td>
<td>Using Surrogates</td>
<td>E26</td>
<td>Copying</td>
<td>The use of simpler and inexpensive copies can be compared to the DSR pattern “use of surrogates”.</td>
</tr>
<tr>
<td>DSR23</td>
<td>Using Surrogates</td>
<td>E27</td>
<td>Cheap Short-Living Objects</td>
<td>The replacement of an expensive object by cheap short-living objects is equivalent to the “use of surrogates”.</td>
</tr>
<tr>
<td>DSR23</td>
<td>Using Surrogates</td>
<td>E28</td>
<td>Mechanical Substitution</td>
<td>The replacement of mechanical means can also be interpreted as a special case of the DSR pattern “use of surrogates”.</td>
</tr>
</tbody>
</table>

1 Number of the DSR Pattern (Vaishnavi & Kuechler 2007), 2 Problem Solving Pattern in DSR (Vaishnavi & Kuechler 2007), 3 Number of the Problem Solving Pattern of the TRIZ Approach, 4 Problem Solving Pattern of the TRIZ Approach

Table 3. Comparison of Problem Solving Patterns in DSR and Engineering (TRIZ) (cf. Gericke 2009)

4.2 Transfer of TRIZ Patterns to the Design Science Research Approach

Based on the comparison of the patterns, we try to transfer the remaining 14 universal innovation principles of TRIZ to DSR. Studying the first remaining TRIZ principle E2 “Extraction”, disturbing parts or only the necessary part of an object should be extracted. Transferred to DSR this could be interpreted in such a way that the researcher has to concentrate on solvable parts of the research problem whereas unsolvable parts are (temporarily) not considered. Table 4 presents the results of the transfer of TRIZ patterns. In the first two columns of the table the remaining TRIZ principles are presented. The next two columns of the table contain the new DSR patterns (new number and new name of the pattern) that result from the transfer of the TRIZ principles. Finally in the fifth column an explanation is given on how the TRIZ principle can be used in DSR.

Transfer of TRIZ Patterns to the DSR Discipline

<table>
<thead>
<tr>
<th>No.</th>
<th>PSP TRIZ</th>
<th>No.</th>
<th>new PSP DSR</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2</td>
<td>Extraction</td>
<td>DSR28</td>
<td>Focused Artifact Construction</td>
<td>Concentrate on the construction process and eliminate (temporarily) unsolvable parts of the problem.</td>
</tr>
<tr>
<td>E3</td>
<td>Local Quality</td>
<td>DSR29</td>
<td>Construction Process Adaptation</td>
<td>Consider contingency aspects in the development phase of an artifact construction process. For example, adapt the construction process to the culture of the research team.</td>
</tr>
<tr>
<td>E5</td>
<td>Merging/Consol.</td>
<td>DSR30</td>
<td>Multiple Tasks</td>
<td>Bring together the input of multiple researchers on one research problem.</td>
</tr>
<tr>
<td>E9</td>
<td>Prior Counteraction</td>
<td>DSR31</td>
<td>Side Effect Evaluation</td>
<td>Before evaluating/using a constructed artifact, explore possible negative effects of its use and propose counteractions that have to be conducted prior or parallel to the use of the artifact.</td>
</tr>
<tr>
<td>E11</td>
<td>Cushion in Advance</td>
<td>DSR32</td>
<td>Rough Solution First</td>
<td>Iteratively develop an artifact to have a rough solution as soon as possible. In the remaining time improve and refine your solution step by step.</td>
</tr>
<tr>
<td>E12</td>
<td>Equipotentiality</td>
<td>DSR33</td>
<td>Reduce Research Efforts</td>
<td>Reduce your research efforts within the “build” phase by falling back on existing (parts) of solutions stored in construction catalogues, such as method repositories for the construction of methods used in the field of method engineering within the DSR discipline.</td>
</tr>
<tr>
<td>E13</td>
<td>Do It in Reverse</td>
<td>DSR34</td>
<td>Unconventional Approach</td>
<td>Do something other than expected within the “build” phase.</td>
</tr>
<tr>
<td>E15</td>
<td>Dynamics</td>
<td>DSR35</td>
<td>Loose Coupling</td>
<td>Use “loose coupling” as a design paradigm.</td>
</tr>
<tr>
<td>E16</td>
<td>Partial or Excessive Action</td>
<td>DSR36</td>
<td>Partial or Excessive Action</td>
<td>If 100 percent of an artifact is hard to achieve using a given method then, by using “slightly less” or “slightly more” of the same method, the problem may be considerably easier to solve.</td>
</tr>
<tr>
<td>E20</td>
<td>Continuity of Useful Action</td>
<td>DSR37</td>
<td>Continuous Construction Process</td>
<td>Try to continuously work on the solution of the research problem. Avoid long breaks in order to stay familiar with the problem, your ideas and the planned research procedure.</td>
</tr>
</tbody>
</table>
Transfer of TRIZ Patterns to the DSR Discipline

<table>
<thead>
<tr>
<th>No.</th>
<th>PSP TRIZ</th>
<th>No.</th>
<th>new PSP DSR</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>E21</td>
<td>Rushing</td>
<td>DSR³⁸</td>
<td>Idea Tracking</td>
<td>If you spontaneously got an idea regarding the solution of a given problem, immediately pursue this idea.</td>
</tr>
<tr>
<td>E22</td>
<td>Convert Harm into Benefit</td>
<td>DSR³⁹</td>
<td>Provocation</td>
<td>Provoke your research team with wrong assumptions in order to improve idea generation regarding the solution of the problem.</td>
</tr>
<tr>
<td>E24</td>
<td>Intermediary</td>
<td>DSR⁴⁰</td>
<td>Intermediary</td>
<td>Call a mentor or consultant in your construction process that can support you in different activities.</td>
</tr>
<tr>
<td>E25</td>
<td>Self-Service</td>
<td>DSR⁴¹</td>
<td>Re-Use Ideas</td>
<td>Document all ideas during the construction process, even if you dismiss them, to take them up in future research projects.</td>
</tr>
</tbody>
</table>

Table 4. Transfer of TRIZ Principles to DSR

The attempt to expand the existing DSR patterns by the remaining universal TRIZ innovation principles from the engineering discipline successfully resulted in 14 new DSR patterns. Hence, all 22 universal TRIZ innovation principles could either be used to back up existing DSR patterns or to serve as a basis for the derivation of new DSR patterns.

5 SUMMARY AND FURTHER RESEARCH

Analyzing the body of literature of DSR it became obvious that there are patterns/principles missing that guide a constructor within the different phases of the construction process. To address this issue, we took up the DSR patterns of Vaishnavi & Kuechler (2007). Due to the fact that their foundation is limited to DSR literature and comparable concepts from the engineering discipline, which form the roots of DSR, have not been used for their foundation, we analyzed problem solving patterns from the engineering discipline. TRIZ, which is the most established problem solving approach in this discipline, was chosen and the included 22 universal innovation principles presented and analyzed in detail.

To support the analysis of both, the problem solving patterns from TRIZ and the DSR patterns of Vaishnavi & Kuechler (2007), a structuring approach which differentiates between patterns referring to a product perspective (the result of the construction process) and patterns referring to a process perspective (the construction process itself) was used. All DSR patterns possess a process character, whereas only some of them contain a product view as well. In a next step, the DSR patterns referring to the process perspective were compared to the universal innovation principles of TRIZ trying to verify the DSR patterns. In doing so, nine DSR patterns could be verified (cf. Gericke 2009). Using this as a basis, we attempted to expand the DSR patterns of Vaishnavi & Kuechler (2007). The remaining 14 universal innovation principles of TRIZ were additionally transferred to the DSR discipline, i.e. the existing DSR patterns of Vaishnavi & Kuechler (2007) could be successfully expanded to include further patterns from the engineering discipline.

In further research works these new patterns should be evaluated. On the one hand, evidence for these new patterns should be adduced by analyzing existing DSR literature and trying to retrieve these patterns from former research. On the other hand, these patterns can be used in future construction processes and the extent to which the efficiency of the construction process was improved by their use should be evaluated. Furthermore it would be helpful and improve usability of the patterns if the new identified patterns would be described as detailed as the existing DSR patterns, e.g. by describing intent, context and applicability, consequences etc. of the patterns. In addition, further research could address the identification of further DSR patterns. First, it is possible to try to transfer the 18 specific innovation principles of TRIZ to the DSR discipline. Second, the DSR patterns referring to a product perspective could be expanded by including the TRIZ innovation principles. Finally, next to problem solving methodologies from the engineering discipline such approaches from other disciplines, e.g. psychology, could be included in DSR as well.

Proceedings ECIS 2009
Acknowledgment

I would like to thank Vijay K. Vaishnavi who supported me during the preparation of this paper by answering questions regarding the understanding of some patterns and the foundation of the DSR patterns. I am grateful to Tobias Bucher and Joachim Schelp for their support during different stages of the paper development including several valuable discussions. Many thanks go to the anonymous paper reviewers for their valuable feedback.

References


# Action Research and Design Science Research – Seemingly similar but decisively dissimilar

<table>
<thead>
<tr>
<th><strong>Journal:</strong></th>
<th><em>17th European Conference on Information Systems</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manuscript ID:</strong></td>
<td>ECIS2009-0424.R1</td>
</tr>
<tr>
<td><strong>Submission Type:</strong></td>
<td>Research Paper</td>
</tr>
<tr>
<td><strong>Keyword:</strong></td>
<td>Action research, Design Science, Research methodology, IS Philosphy</td>
</tr>
</tbody>
</table>
ACTION RESEARCH AND DESIGN SCIENCE RESEARCH –
SEEMINGLY SIMILAR BUT DECISIVELY DISSIMILAR

Iivari, Juhani, University of Oulu, P.O. Box 30000, SF-90014 Oulu, Finland,
juhani.iivari@oulu.fi

Venable, John, School of Information Systems, Curtin University of Technology, GPO Box
U1987, Perth WA 6845, Australia, j.venable@curtin.edu.au

Abstract

Prior research has identified the similarity of Action Research (AR) and Design Science Research
(DSR). This paper analyses AR and DSR from several perspectives, including paradigmatic
assumptions of ontology, epistemology, methodology, and ethics, their research interests, and
activities. We identify that often AR does not share the paradigmatic assumptions and the research
interests of DSR, that some activities in DSR are always mutually exclusive from AR, and that there
may be no, little, or significant (but not total) overlaps between AR and DSR. Thus we judge that AR
and DSR are decisively dissimilar. We further identify several key problems with combining AR and
DSR based on the ethical requirement of researchers to identify and manage risks to research
stakeholders. Management of such risks is done by careful disclosure, identifying research limitations
or by choosing alternative methods than AR for accomplishing DSR.

Keywords: Action Research, Design Science Research, Research Paradigm, Ontology, Epistemology,
Research Methodology, Ethics.
1 INTRODUCTION

The recent attention to Design Science Research has led many commentators to consider the relationship between Action Research (AR) and Design Science Research (DSR) (e.g. Burstein and Gregor, 1999; Cole et al., 2005; Järvinen, 2007). Järvinen especially makes a strong case for DSR being similar to AR. He analyzed and aligned the matching activities and characteristics of both AR and DSR. He found that AR and DSR have five main activities in a research cycle, which match. He also identified seven characteristics of AR and six characteristics of DSR and shows how they match. Overall, he found AR and DSR to be very similar, and as a consequence suggests that AR as a qualitative research method has a wrong “home” and should be more closely associated with DSR.

We agree that AR and DSR share many features, but at the same time we view Järvinen’s conclusions as overly hasty. First, if we accept that, as DSR has been implicitly practiced in engineering for centuries and in Computer Science and Software Engineering for decades, it is hard to interpret that all this DSR has actually been AR. In our opinion, doing so would extend the concept of AR much too far (we will discuss the concepts of AR and DSR in the next section).

Second, in our opinion, Järvinen’s analyses of both AR and DSR are based on somewhat biased conceptions of AR and DSR. In the case of AR, he sees action researchers’ intent as to plan and take action in order to change a part of reality. His focus is on the practical problem solving interest side of AR, de-emphasizing its research interest side (McKay and Marshall 2001). This focus may bias his interpretation of AR toward one that aligns better with DSR. In the case of his interpretation of DSR, Järvinen relies heavily on van Aken (2004). Although van Aken (2004) distinguishes improvement problems and construction problems, his major interest lies in improvement problems rather in construction of complex artifacts. Solution concepts (van Aken 2004) – whether general or specific – to improvement problems often are not complex artifacts, although they may be implemented in a complex and messy organizational environment. As a consequence, building artifacts does not have as central a role in addressing improvement problems as it does in addressing construction problems. Järvinen’s focus on van Aken’s (2004) emphasis on improvement problems in DSR biases his interpretation of DSR toward one that aligns better with AR than DSR would align more generally.

The purpose of the present paper is to continue the discussion about the relationship between AR and DSR. We will argue that although AR and DSR are seemingly similar, they are quite different phenomena. This conclusion is based on two analyses. First in Section 3, we analyze the paradigmatic – ontological, epistemological, methodological and ethical – roots and assumptions of AR and DSR, finding that they are somewhat different, but not incompatible. For our second analysis, Section 4 will analyze the research interests and activities of AR and DSR, showing how, while they are again compatible and may significantly overlap, they are decisively different. Based on these analyses and some further literature, Section 5 will discuss some ethical problems and more pragmatic reasons to keep AR and DSR as separate, especially when the focus of DSR is building and testing innovative artifacts. These all revolve around the fact that when building innovative, cutting-edge artifacts, possible failure is always present. That may make it difficult and also risky to combine AR and DSR.

2 WHAT IS ACTION RESEARCH AND WHAT IS DESIGN SCIENCE RESEARCH?

In this section we delineate the conceptions of AR and DSR adopted in the present paper.

2.1 Action Research

Action Research has been widely discussed in the IS literature (e.g. Davison et al. 2004; Kock 2004; Lau 1999; McKay and Marshall 2001). The purpose here is not to repeat that discussion, but to summarize the interpretation of AR adopted in the present paper.
A widely adopted definition of AR by Rapoport (1970) characterizes it as follows: “Action research aims to contribute both to the practical concerns of people in an immediate problematic situation and to the goals of social science by joint collaboration within a mutually acceptable ethical framework”. According to this definition AR has a dual goal of contributing to practice and research at the same time. The definition also assumes that there is a concrete client involved. As a consequence, AR is highly context dependent while attempting to address the specific client’s concerns.

Action Research is not a homogenous whole, but includes a number of versions (Baskerville and Wood-Harper 1998; Chandler and Torbert 2003). This paper focuses on Canonical Action Research (CAR) (Davison 2004), which is based on five principles:

1. The principle of the researcher-client agreement
2. The principle of the cyclical process model
3. The principle of theory
4. The principle of change through action, and
5. The principle of learning through reflection

The definition of AR by Rapoport (1970) emphasizes the first principle. The cyclical model of Susman and Evered (1978) – diagnosing, action planning, action taking, evaluating, and specifying learning – is the most well-known example of the second principle. McKay and Marshall (2001) summarize other cyclical AR models and propose a dual cycle model in which the cycle of research interest and the cycle of problem solving interest proceed in parallel while closely interacting with each other. For the third principle Davison et al. (2004) note that a CAR project may begin with theory-free action learning, but suggest that action researchers need theory to guide and focus their activities. We agree that qualitative research project may start relatively theory-free, but not totally. However, “theory” is an ambiguous concept (Sutton and Staw 1995) with multiple meanings (Gregor 2006). Most often it refers to theories for predicting and explaining, which are rare in IS research. Therefore, in line with McKay and Marshall (2001), we would speak about theoretical frameworks rather than about theories.

The fourth principle, the principle of change through action, attempts to ensure that diagnosing the problem, action planning, action taking and evaluating are appropriately done so that one can expect an improvement in the client’s problematic situation (Davison et al. 2004). The final principle, the principle of learning through reflection, attempts to ensure that both researchers and the client examine what they have learned in an explicit, systematic and critical manner (Davison et al. 2004).

2.2 Design Science Research

To our knowledge, there is no widely accepted definition of Design Science Research. Vaishnavi and Kuechler (undated) characterize “Design Research” as “yet another "lens" or set of analytical techniques and perspectives (complementing the Positivist and Interpretive perspectives) for performing research in IS. Design research involves the analysis of the use and performance of designed artifacts to understand, explain and very frequently to improve on the behavior of aspects of Information Systems.” Hevner et al. (2004) note that the design science paradigm seeks to extend the boundaries of human and organizational capabilities by creating new and innovative artifacts, including constructs, models, methods, and instantiations (also as in March and Smith, 1995). Venable (2006) identified “solution technology invention” as the core of DSR. We are ready to define DSR as a research activity that invents or builds new, innovative artifacts for solving problems or achieving improvements, i.e. DSR creates new means for achieving some general (unsituated) goal, as its major research contributions. Such new and innovative artifacts create new reality, rather than explaining existing reality or helping to make sense of it.

When compared with AR, an essential difference is that DSR assumes neither any specific client nor joint collaboration between researchers and the client. One should note, however, that typically the developed artifact aims at addressing a class of problems (Walls et al. 1992) in a way that it is useful in addressing specific problems of a specific client. As a consequence, one could argue that DSR has would-be-clients who comprise “The set of all members of the generalised class of all people or
organizations who could potentially be motivated to solve instances of the generalised class of problem(s)” (Venable 2009) addressed by the DSR outcome (artifact).

Some researchers have advocated a ‘soft’ version of DSR. Venable and Travis (1999) and Venable (2006) argued that more qualitative methods, including AR, could be used for evaluation of DSR outcomes. Venable (2006) further noted that AR could be used across the entire DSR lifecycle (on the naturalistic evaluation side). Baskerville et al (2007) proposed the development of a Soft DSR approach to better accommodate risks in DSR, including poor problem understanding and the potential for evaluation to give rise to type I or type II errors. However, these ‘soft’ approaches to DSR still assume a general, unsituated goal or problem to be solved rather than the situated one inherent in AR.

More fundamentally, equating AR to DSR suffers from the category problem: AR is a research method while DSR is more a research orientation, within which one can use different research methods (including AR). See the various research methods identified for evaluating innovative solutions as part of DSR in Hevner et al (2004) or Venable (2006), the latter of which includes AR.

3 PARADIGMATIC ASSUMPTIONS OF AR AND DSR

One possibility to scrutinize similarities and differences between AR and DSR is to analyze their paradigmatic assumptions. The section introduces first the paradigmatic framework to be applied. After it AR and DSR are analyzed in turn using the framework and finally the results are summarized.

3.1 The paradigmatic framework

Adopting Burrell and Morgan (1978), Iivari (1991) analyzed paradigmatic assumptions of a number of IS development approaches. The analytical framework included ontology, epistemology, methodology and ethics of research. In the following we will apply this framework.

Burrell and Morgan (1978) distinguish two extremes in the case of ontology: realism vs. nominalism. Realism “postulates that the social world external to individual cognition is a real world made up of hard, tangible and relatively immutable structures. [...] For the realist, the social world exists independently of an individual's appreciation of it [...] the social world has an existence which is as hard and concrete as the natural world” (p. 4). Nominalism, on the other hand, “revolves around the assumption that the social world external to individual cognition is made of nothing more than names, concepts and labels which are used to structure reality [...]” (p. 4).

Similarly, Burrell and Morgan (1978) identity two epistemological positions: positivism vs. anti-positivism. Positivism seeks “to explain and predict what happens in the social world by searching for regularities, causal relationships between its constituent elements”, whereas anti-positivism maintains that the social world “can only be understood from the point of view of the individuals who are directly involved in the activities which are to be studied. Anti-positivists reject the standpoint of the ‘observer’, which characterizes positivist epistemology, as a valid vantage point for understanding human activities. They maintain that one can only 'understand' by occupying the frame of reference of the participant in action. One has to understand from the inside rather than the outside.” (p. 5).

In the case of methodology, Iivari (1991) distinguishes three categories: nomothetic methods, idiographic methods and constructive methods. The last category complements Burrell and Morgan’s distinction between nomothetic methods and idiographic methods. Focusing specifically on IS development, Iivari (1991) distinguished conceptual development and technical development in the context of constructive methods explaining that “Conceptual development as a category of constructive research methods refers to the development of various models and frameworks which do not describe any existing reality but rather help to create a new one, and which do not necessarily have any "physical" realization (e.g. IS development "methodologies"). Technical development produces as its outputs "physical" artifacts, the adjective "physical" being interpreted here broadly to include
executable software (e.g. CASE environments)” (Iivari 1991, p. 257). These characterizations attempted to capture specific research methods used in the development of the above artifacts.

In the case of ethics, Iivari (1991) distinguishes three ethical positions: means-end, interpretive, and critical. In the first case, research aims at providing means knowledge for achieving given ends (goals), without questioning the legitimacy of the ends. The interpretive stance questions the realism of the idea of human and organizational action as goal-directed action. Goal statements follow rather than precede action (cf. March and Olsen, 1976). They are reconstructed retrospectively to give meaning to action. According to Chua (1986), the aim of an “interpretivist scientist is to enrich people’s understanding of their action” and “how social order is produced and reproduced” (p. 615), and a critical scientist sees that research has “a critical imperative: the identification and removal of domination and ideological practice” (p. 622). As a consequence, critical research subjects goals (ends) to a critical analysis.

In the following we apply the above paradigmatic framework to analyze similarities and differences between AR and DSR.

3.2 Paradigmatic assumptions of Action Research

To our knowledge, the early article of Oquist (1978) and the recent article of Cassell and Johnson (2006) are the most serious attempts to analyze the philosophical assumptions of AR. Oquist (1978) analyzed the kind of knowledge AR produces and its relation to different schools of philosophy of science (Empiricism, Logical Positivism, Structuralism, Pragmatism and Dialectical Materialism). Although his main focus lies in the epistemology, his analyses also cover ontological and ethical assumptions. He concludes that AR is compatible with the assumptions of Pragmatism and Dialectical Materialism, but incompatible with those of Empiricism, Logical Positivism and Structuralism.

Cassell and Johnson (2006) argue that the diversity of AR is due to different philosophical assumptions adopted. More specifically, they discuss philosophical assumptions underlying five different approaches to AR, which they label “experimental”, “inductive”, “participatory action research”, “participatory”, and “deconstructive”. “Experimental” AR covers Lewinian and Tavistockian traditions of AR with realistic ontology and positivistic epistemology. Cassell and Johnson (2006) explain “inductive” AR as an approach in which “theory … is generated from the data (…) of thick descriptions of the patterns of subjective meanings that organizational actors use to make sense of their worlds, rather than entailing the testing of hypotheses deduced from a priori theory that causally explains what has been observed by the action researchers” (p. 793). They refer to grounded theory as an example of the inductive approach and discuss “action science” (Argyris et al. 1985) as an example of the “inductive” approach. “Participatory action research” in Cassell and Johnson (2006) refers to an AR approach in an organizational or corporate setting in which the researcher usually works in a consultancy role serving the corporate elite, while “participatory research”, typically informed by ‘critical theory’, “requires that those individuals and groups whose perspectives are ordinarily silenced in organizations must be given voice through action research” (p. 798). Finally, “deconstructive” AR is an emerging approach that is based on postmodernism. It assumes a plurality of incommensurable voices, implying that “any discursive closure, whether grounded on democratic consensus or otherwise, implies the arbitrary dominance of a particular discourse which serves to silence alternative possible voices” (p. 804). Quite interestingly, none of the five approaches discussed by Cassell and Johnson (2006) directly corresponds to CAR, but in our view the “inductive” approach is closest to it.

Based on Oquist (1978) and Cassell and Johnson (2006) and our own reading of AR, we conclude that the ontology of CAR is anti-realist. It does not see (social) reality as “hard, tangible and relatively immutable structures” (Burrell and Morgan 1978), but as socially constructed that can be changed. At the same time we do not interpret that AR necessarily adopts the extreme ontological position of nominalism. Therefore we characterize the ontological position of CAR as anti-realist.
Similarly, the epistemological position of CAR is mainly anti-positivist. Although an action researcher may search for regularities and causal relationships in the social world or may apply such when interpreting the world, CAR clearly recognizes the limits of such regularities. Each CAR project is unique and the case can only be understood from the point of view of the individuals who are directly involved in the activities which are to be studied. As a consequence, in accordance to Iivari (1991), we interpret that AR as a research method is clearly idiographic.

In the case of ethics, we interpret that CAR is mainly means-ends oriented, but can also be interpretive. The major reason is that CAR is typically oriented in solving or alleviating a specific client’s problems, as evidenced by the 5-step AR cycle proposed by Susman and Evered (1978). CAR can also be interpretive if evaluation focuses on rich understanding of meanings attached to the executed action and its intended and unintended consequences. To some extent differing from Oquist (1978), we do not see CAR as representing a critical approach in the sense that it would challenge existing power structures or structures of domination and would suggest radical changes in that meaning. The major reason is that CAR represents joint collaboration between researchers and the specific client, within a mutually acceptable ethical framework. It is extremely unlikely that the most powerful representatives of the client (its management) would be ready for such radical changes.

Assuming that CAR is closest to the “inductive” approach discussed by Cassell and Johnson (2006), our assessment differs from theirs in that they conclude that the “inductive” approach ultimately reflects positivistic epistemology. The positivistic epistemology manifests in the presupposition of neutral observational language, and in the privileged position of the researcher who is assumed to access empirically data from the independent reality. We agree with Cassell and Johnson (2006) in the sense that the positivistic epistemology is so deep-rooted in our conception of research and science that often times when one scratches the surface of seemingly anti-positivistic approach, positivistic ideas can be found beneath the surface. One should note, however, that we do not see positivism and anti-positivism as mutually exclusive, but a research approach may – possibly unconsciously - include ideas from both philosophical positions. In our view CAR implies a clear move towards anti-positivistic epistemology although it may include some positivistic ideas.

3.3 Paradigmatic assumptions of Design Science Research

The DSR literature only rarely discusses its paradigmatic assumptions. Iivari (2007) discusses DSR’s ontology based on Popper’s 3-world ontology, epistemology in terms of types of knowledge associated with DSR, distinctive constructive research methodology, and ethics, but does not focus on its assumptions specifically. Vaishnavi and Kuechler (undated) briefly discuss DSR’s ontology, epistemology, methodology and axiology (= ethics), contrasting DSR with positivism and interpretivism. Unlike Vaishnavi and Kuechler (undated), the present paper does not see DSR as a third “paradigm”, but as based on more or less “positivistic” or “interpretivist” assumptions. This is in line with Niehaves (2007) who has a detailed analysis of the epistemological assumptions of DSR, claiming that “interpretive” (non-positivistic) epistemology is highly relevant in the context of DSR.

We conclude that DSR may adopt a realistic or anti-realistic ontological position. DSR especially in engineering exemplifies the realistic position, but we do not see realism inherent to DSR. Markus et al. (2002) exemplifies this point – in our view it clearly adopts a non-realistic ontological position.

DSR may also adopt both positivistic and anti-positivistic epistemology. Much of DSR in engineering is based on positivistic epistemology. Although March and Smith (1995) and Hevner et al. (2004) have a positivistic epistemological bias (Niehaves 2007), we agree with Niehaves (2007) that the anti-positivistic epistemology is also relevant in DSR, especially in evaluation of developed artifacts. For example, Baskerville et al (2007) propose “Soft Design Science” and Venable (2006) and Venable and Travis (1999) identify interpretive methods as appropriate for naturalistic, in situ evaluation of DSR outputs. On the other hand, since DSR aims at producing general solution concepts (van Aken 2004) or meta-artifacts (Walls et al. 1991; Iivari 2003), which are more widely applicable than in a specific organizational context, DSR aims at general knowledge typical to the positivistic epistemology.
Methodologically, DSR may apply both nomothetic and idiographic methods, as the variety of design evaluation methods proposed by Hevner et al. (2004) implies. As explained above, Iivari (1991) suggested constructive research as a third category of methods. Since it focuses on building IT artifacts, it can be considered a distinctive category of methods in the case of DSR.

With respect to ethics, much (if not all) of DSR is means-end-oriented. The artifact to be developed is assumed to have some purpose (Hevner et al. 2004; Gregor and Jones 2007). However, Venable (2009) claims that Hevner et al. (2004) consider profit (utility) maximization as the ultimate goal of DSR and that, overall, they privilege managers as a stakeholder group on business organizations. In AR, one can also use interpretive evaluation studies that aim at a rich understanding of the meanings attached to the constructed artifact and its intended and unintended consequences in practice. However, DSR may more easily have a critical orientation. DSR may aim at developing new artifacts that challenge existing power structures of domination. The Scandinavian trade-unionist IS development approach, is an example this orientation (see Iivari et al. 1998).

3.4 Summary of the analyses

Table 1 summarizes the above analysis. It suggests that DSR is more varied in its paradigmatic assumptions than AR. Paradigmatically, AR can be considered a special case of DSR, yet keeping in mind that DSR by definition includes building new innovative artifacts (unlike AR). As a consequence, there is not necessarily any paradigmatic incommensurability problems between AR and DSR, so they can be applied together, especially if DSR adopts paradigmatic assumptions that are compatible with those of AR. We also believe that if DSR activities of building an artifact and its evaluation are separate, one can easily apply AR in the evaluation (Venable, 2006, Johnstone and Venable, 2008), even if the building of the artifact has been based on paradigmatic assumptions that differ from those of AR. For example, one may have engineered an artifact based on a positivistic epistemology reflecting a realistic ontology (e.g. laws of physics). The evaluation of the very same artifact may follow an anti-positivistic epistemology and an anti-realistic ontology. In fact, we believe that idiographic research methods reflecting an anti-positivistic epistemology are often justified when attempting to understand organizational or individual appropriation, usage and consequences of any technology in real life.

<table>
<thead>
<tr>
<th>Paradigmatic dimension</th>
<th>Action Research</th>
<th>Design Science Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ontology</td>
<td>Anti-realism</td>
<td>Reality or anti-realism</td>
</tr>
<tr>
<td>Epistemology</td>
<td>Mainly anti-positivism</td>
<td>Mainly positivism, but also anti-positivism especially in evaluation</td>
</tr>
<tr>
<td>Methodology</td>
<td>Idiographic</td>
<td>Constructive (building)</td>
</tr>
<tr>
<td>Ethics</td>
<td>Means-end</td>
<td>Means-end</td>
</tr>
<tr>
<td></td>
<td>Possibly interpretive</td>
<td>Possibly interpretive</td>
</tr>
<tr>
<td></td>
<td>Unlikely critical</td>
<td>Possibly critical</td>
</tr>
</tbody>
</table>

Table 1. Summary of the paradigmatic assumptions of Action Research and Design Science Research.

4 COMBINING AR AND DSR

As noted in Section 2 considering our conceptions of AR and DSR, by definition the research interest of DSR is to construct new and innovative ways to solve a class or classes of problems, thus creating new reality. AR does not necessarily share any such purpose. Much of AR is conducted to understand existing reality, such as the complex workings of organisational situations and human behaviour. These research interests are decidedly different.
We now turn our analysis to the activities of AR and DSR. To facilitate our analysis, we conceptualize of two broad classes of problems and solutions: purely technical problems and innovations and socio-technical problems and innovations. Furthermore, we assume that any solutions or innovations to purely technical problems do not have direct and complex implications for the socio-technical systems within which they will be embedded (other than monetary costs, which are always taken into consideration in any innovation). Therefore, there is little or no interest in such problems and innovations in the context of AR, although there can be considerable interest in the context of DSR. Rather, AR is exclusively interested in innovations with significant impacts on socio-technical systems and the human context. This distinction between purely technical and socio-technical problems and innovations is reflected in the potential activities or AR and DSR, which we will now analyse.

We base our analysis of AR and DSR activities on the model of DSR activities presented in Venable (2006), which extended the multi-methodological framework of Nunamaker (1990) and Venable and Travis (1999) to differentiate between Naturalistic and Artificial evaluation in DSR, as in Figure 1.

**Figure 1. Framework and Context for DSR (Venable, 2006).**

In Figure 1, AR is shown as one of several means of conducting naturalistic evaluation of a new and innovative “solution technology” (Venable, 2006). DSR can also (or alternatively) conduct artificial evaluation. Artificial evaluation may be all that is required for purely technical artifacts. Artificial evaluation is not part of or relevant to AR, which is a decisive difference between AR and DSR.

Using this framework to analyze the overlapping activities between AR and DSR, we can identify three different cases: (1) completely non-overlapping (in three different ways), (2) slightly overlapping, and (3) significantly overlapping. Note that we do not call the third case “fully overlapping”, because artificial evaluation will never be a part of AR, as described in the previous paragraph. We will consider each of these three cases in turn.

In the first case, AR and DSR activities will be completely non-overlapping. Here AR is not concerned at all with the DSR research interest of building and evaluating innovative artifacts. One such situation is where the problem solving interest of the client will be addressed without any technology or through the application of existing solutions without innovation. In the latter case AR may include design, but it is normal design practice rather than DSR. Here, in so far as the action researcher is concerned with
the efficiency or effectiveness of the solution technology, it is concerned with the client’s problem solving interest. This problem solving interest merely provides the context and opportunity to research other aspects of socio-technical systems, such as organisational or human behaviour.

A second situation in which there is no overlap (which we will call case 1b) is where the DSR is concerned with solving a purely technical problem. In this case, evaluation is primarily artificial and there is no need for naturalistic evaluation (applying AR, for instance).

A third situation in which there is no overlap (which we will call case 1c) is where the DSR is concerned with solving a socio-technical problem, but not arising in an AR context and where AR is not used for naturalistic evaluation. Evaluation may be conducted artificially and/or using other research methods for naturalistic evaluation, such as case studies or surveys. Note that such evaluation could complement, precede, or follow separate AR evaluations (see case 2 below).

The second case concerns the use of AR to evaluate DSR. In this case, the action researcher does not develop a new, innovative artifact or solution technology, but has the express purpose of evaluating an existing solution technology that is still of research interest. In this case, the research interest (or a major part of it) is the evaluation and a planned research outcome is statements about the efficiency, effectiveness, or efficacy of artifacts that have been proposed and developed by other researchers or practitioners. In this case, there is only slight overlap, which is limited to the naturalistic evaluation activities in Figure 1.

The third case (significant overlap) is where the action researcher actually is also conducting DSR, in that he/she is inventing a new, innovative artifacts or solution technology to better address the client’s problem solving interest (a socio-technical problem). In this case, the research interest includes the development and evaluation of the solution technology. In the framework shown in Figure 1, this would include the theory building, solution technology invention, and naturalistic evaluation activities.

Table 2 summarizes these three cases.

<table>
<thead>
<tr>
<th>Case</th>
<th>AR Interest</th>
<th>DSR Interest</th>
<th>DSR Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a. No overlap</td>
<td>Understanding reality in an organizational context</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>1b. No overlap</td>
<td>None</td>
<td>Solving a purely technical problem by developing and evaluating a new solution technology</td>
<td>Theory building, Solution technology invention, and Artificial evaluation</td>
</tr>
<tr>
<td>1c. No overlap</td>
<td>None</td>
<td>Solving a socio-technical problem in a non-AR context by developing a new solution technology, but evaluating it by means other than AR</td>
<td>Theory building, Solution technology invention, and Artificial and/or naturalistic evaluation</td>
</tr>
<tr>
<td>2. Slight overlap</td>
<td>Evaluating an existing solution technology in an organizational context</td>
<td>Evaluation of a solution technology developed separately</td>
<td>Naturalistic evaluation only</td>
</tr>
<tr>
<td>3. Significant overlap</td>
<td>Solving a socio-technical problem by developing and a new solution technology and evaluating it in an organizational context</td>
<td>Solving a socio-technical problem by developing and a new solution technology and evaluating it in an organizational context</td>
<td>Theory building, Solution technology invention, and Naturalistic evaluation</td>
</tr>
</tbody>
</table>

Table 2. Overlaps in activities between AR and DSR.

5 PRACTICAL PROBLEMS OF COMBINING AR AND DSR

Although AR and DSR are in principle compatible with each other, they may be difficult to combine for more practical reasons. In the following we shall discuss these difficulties in more detail.
As pointed out above, AR assumes collaboration between researchers and the client within a mutually acceptable ethical framework. This ethical framework may not be self-evident in an AR project, but if the AR is part of a larger DSR effort, it may be particularly difficult to achieve such agreement since the major parties, researchers vs. clients, may have widely opposing goals. DSR by definition attempts to construct new and innovative artifacts or solution technologies. As a consequence, it operates at the edge of the existing technology. Such cutting edge technology is not usually robust. Petrovski (1982) claims that failure is central to understanding engineering, and that lessons learned from failures can advance engineering knowledge more than all the successful cases. Similarly, DSR may push a new technology to its limit until it collapses. However, in AR the client’s interest is usually that the technology employed is proven and robust and does not disturb the performance of their work. Having a technology fail would be unlikely to be appropriate or agreeable to the client in an AR context.

Weedman (1998, 2008) describes such a conflict of interest in the Sequoia 2000 project. It was essentially a combined DSR and AR project, although she does not refer to AR in her two papers. The project comprised computer scientists who acted as researchers and earth scientists as users. She describes the conflict of interest when the technology was repeatedly pushed to the limit by computer scientists until it broke, with consequent costs to the users when the system crashed and the users’ work came to a halt.

In conclusion, DSR may apply a conscious “learning from failures” strategy when developing new, innovative artifacts. Although the goal of this “learning from failures” strategy is to develop more robust technology that would be safer for the potential users of the developed technology or artifact, one should note that new technology is always failure prone and risky. In such a case, the research interest of the researcher would conflict severely with the problem solving interest of the client. AR in the DSR context implies that potentially failure prone technology is applied in the real-life context. Depending on the nature of technology and situation, it is more or less risky to the client. The potential of risk to the clients or the public have led DSR in engineering and medicine to prioritize laboratory testing in the initial evaluation (testing) of the developed artifacts or technology. This initial testing is done clearly separated from potential clients, partially to reduce risks.

In terms of the framework from Venable (2006) in Figure 1 above, where there are significant risks, the evaluation of the developed artifacts (e.g. as in case 2 above) should be done artificially, if possible, before attempting to evaluate naturalistically, using AR, for example. One can identify two opposing aspects in AR as far as risks to the client are concerned. On one hand, the involvement of the researcher in choosing the appropriate solution technology and the provision of on-going guidance for the using client organization during the AR study may mitigate the risk to the client. On the other hand, the AR researcher is often an advocate of the solution technology, and therefore may be prepared to push the technology too far from the client’s viewpoint in order to make it work. To mitigate this the client should always have the right to quit, if it does not want to continue the AR project.

When AR overlaps significantly with DSR (as in case 3 in the previous section), the solution technology invention is “on the fly”, which is inherently more risky than careful and thorough development before moving on to a separate evaluation. Thus, in our view, the mixed invention and evaluation of a new solution technology combining AR and DSR should only be used in situations where the risk to the client (and other stakeholders) is acceptably low. Also, the mutually agreeable ethical framework for the research should identify and make clear the risk of developing and trying out something new and take steps to mitigate that risk.

Finally, as identified in Venable (2009) and Pries-Heje et al (2008), there are risks in DSR that new or innovative solutions, once published and made available to be adopted by the public, do place these other stakeholders at risk. It is necessarily the case that AR is conducted in a narrow context, which implies that a new solution technology is untried in other contexts. Steps therefore need to be taken to clarify the extent of the evaluation in the AR context and to decide and state whether risks are yet
acceptable for public consumption of new technologies. Alternatively, other research methods than AR might be more appropriate for conducting the DSR.

6 CONCLUSIONS

There has been a significant and ongoing theoretical discussion in the IS field as to the nature and compatibility of AR and DSR, some of which has claimed that the two are similar, if not identical. In this paper we have analyzed the similarities and differences between AR and DSR based on their respective paradigmatic roots and assumptions, their research interests, the overlaps in their activities, and potential problems in combining the two approaches. Our conclusion is that, while the two approaches are compatible (and their synergistic use warrants further development), their paradigmatic assumptions, research interests, and activities may differ dramatically, depending on the purpose of research. More fundamentally, their levels or categories as research orientation (DSR) or method (AR) are different. Thus, we find that their similarities are actually quite superficial and the two things are decisively different.

Considering the practical implications of this analysis, we further identify that the employment of AR in the conduct of DSR needs to be done with care, especially where there is the potential for significant risk to the client or other stakeholders. Steps need to be taken where new technologies are being evaluated with AR to ensure that the evaluation is done carefully and that the completeness of the evaluation is judged and reported with a view to its limitations and remaining risk. This is especially the case where the new technology is developed “on the fly” as part of an AR project.

Our further recommendation is that clear consideration of risks due to the untried nature of the technology must be part of an ethical framework to be agreed at the beginning of a joint AR and DSR project. This is not just for improving the likelihood of successful conduct of the research project, but is the ethical responsibility of the researcher. Such agreement must be reached with the full, informed, and uncoerced consent of the client. Steps should also be taken, where possible, to reduce the risk to the client by evaluating the technology beforehand using other evaluation methods.

References


Järvinen, P. (2007). Action research is similar to design science, Quality & Quantity, 41, 37-54.


<table>
<thead>
<tr>
<th><strong>Journal</strong></th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manuscript ID</strong></td>
<td>ECIS2009-0676.R1</td>
</tr>
<tr>
<td><strong>Submission Type</strong></td>
<td>Research Paper</td>
</tr>
<tr>
<td><strong>Keyword</strong></td>
<td>Design research, Design Science, IS Philosophy, Research methodology</td>
</tr>
</tbody>
</table>
SCIENTIFIC PROGRESS OF DESIGN RESEARCH ARTEFACTS

Aier, Stephan, University of St. Gallen, Institute of Information Management, Müller-Friedberg-Strasse 8, CH-9000 St. Gallen, Switzerland, stephan.aier@unisg.ch
Fischer, Christian, University of St. Gallen, Institute of Information Management, Müller-Friedberg-Strasse 8, CH-9000 St. Gallen, Switzerland, christian.fischer@unisg.ch

Abstract

In parallel to widely accepted behavioural research, Design Research (DR) has emerged in Information Systems. Nonetheless, the debate about the scientificity of DR is still ongoing. In the course of this debate, the role of scientific progress has hardly been discussed. But, doubtlessly, scientific progress is regarded as one of the main aims of science; science can even be defined by scientific progress. Philosophy of science has therefore developed a variety of concepts for scientific progress mostly adapted to explanatory and/or predictive theories. Nonetheless, the output of DR differs from those theories; therefore, concepts developed cannot be applied to DR without further ado. In this paper, we propose a first concept for scientific progress of DR artefacts. Because of the complexity of the field we firstly restrict to the research question: What is scientific progress of DR artefacts? Progress is commonly defined as a transition from step A to B whereby B is “better” than A. The aim of our research is to identify criteria for concretizing what is “better” in the context of DR and to define criteria of progressive DR artefacts. We thereby identified the following five criteria: utility, internal consistency, external consistency, scope, and efficiency.

Keywords: Design Science, Design Research, Design Theory, Scientific Progress.
1 THE IMPORTANCE OF SCIENTIFIC PROGRESS TO DESIGN SCIENCE

In parallel to widely accepted positivistic research approaches, design research (DR) has emerged in Information Systems (IS) Research (Hevner et al. 2004) and in other management sciences, e. g. organisational science (van Aken 2004; Romme 2003; Romme & Endenburg 2006).1 In ISR, a variety of articles describing and reflecting the construction and evaluation of concrete DR artefacts have been published.2 In parallel to the construction of concrete artefacts, authors also reflected the nature of design oriented research as well as the construction and evaluation of DR artefacts in general.3

Although many advances have been done in Design Science (DS) in IS, the debate on the scientificity of DR is still ongoing. In the course of this debate, the role of scientific progress has hardly been discussed. However, scientific progress is regarded as one of the main aims of science; science can even be defined by scientific progress.4 Philosophy of science has developed a variety of concepts for scientific progress. But, in most cases, they were adapted to theories in natural sciences, later also to theories in other disciplines (e. g. social sciences). Nonetheless, the output of DR differs from explanatory and predictive theories (Gregor 2006; Romme 2003). Therefore, concepts developed in philosophy of science referring to “traditional theories”5 cannot be applied to DR artefacts without further ado.

The aim of this paper is to transfer results of the discussion about scientific progress led in Philosophy of Science to the field of DS in IS. This is doubtless a complex enterprise. Yet, there has not been conducted much research on scientific progress in DR; consequently, this paper can only be a first approach. At the beginning of such a research, it is necessary to advance stepwise. The aim of this paper is to develop a set of criteria of progress of DR artefacts. This research aim concerns the ontological understanding of DR artefacts. Gregor and Jones (2007) make a seminal contribution by synthesising the ontological understanding of design research artefacts in a broad literature review. We therefore take their work as a starting point for our investigations.

In order to transfer results from philosophy referring to traditional theories to DR artefacts, we firstly trace the complex discussion of scientific progress in philosophy of science briefly. Secondly, we describe the anatomy of DR theories in order to understand differences to traditional theories. Thirdly, we transfer concepts from philosophy of science referring to traditional theories to DR artefacts and identify criteria of progress of DR artefacts. Finally, we summarise the criteria identified and show the necessity of further research.

---

1 Different authors in management sciences and information systems have a different understanding of what they mean by “design”. Some authors emphasise on aesthetical aspects and on creativity (Ehn & Malmberg 1998; Stolterman 2008) whereas others emphasise on utility and an engineering-like construction (van Aken 2004; Hevner et al. 2004; Romme 2003). In this paper, we clearly emphasise on the utility of the artifact and an engineering-like construction process.

2 An overview of the articles published at the Design Science Research in Information Systems Conference (DESRIST) is summarised by Bucher & Winter (2008, table 1 and table 2).

3 According to Winter (2008, p. 471), in this article, we call the concrete construction of design research artefacts “Design Research”, the reflection on the nature of design research artefacts and on their construction and evaluation in general “Design Science”.

4 Kuhn (1970, p. 162) asks the rhetorical question: “Does a field make progress because it is a science, or is it a science because it makes progress?”

5 In the following, we call descriptive, explanatory, and predictive theories “traditional” in order to demarcate them from design theories (Gregor 2006).
Progressivity of science has always been discussed in philosophy of science. Early works on the subject were already published in the 19th century, e.g. by William Whewell (Losee 2004, pp. 7-27). As always in philosophy, the discussion of scientific progress is multifarious. However, in the entire discourse there is one consensus: Scientific progress is a normative concept. A step from stage A to B constitutes progress if B is better than A in respect to some standards or criteria (Bird 2008). Object of discussion is the nature of such standards or criteria.

It is impossible to entirely summarise that broad discussion in this short article. Instead, we restrict ourselves to three major schools in philosophy of science whose concept of scientific progress influenced other philosophers of science significantly (Kaiser 1991, p. 12): the Logical Empiricism by the philosophers of the Vienna Circle, the Critical Rationalism by Popper, and the descriptive approach to philosophy of science by Kuhn. Finally, we propose how to synthesise some of the concepts presented by referring to Scientific Structuralism.

2.1 Logical Empiricism

The emergence of the Logical Empiricism is often said to be the beginning of modern philosophy of science. The concept of science of the Logical Empiricism was developed by the philosophers of the Vienna Circle since the 1920s. The seminal contributions of Logical Empiricism to philosophy of science in general, and to the subject of scientific progress in particular, were a starting point for the discussion by many philosophers later-on. Thereby, the Deductive-Nomological model for scientific explanations by Hempel and Oppenheim (1948) is central for the conceptualisation of science by Logical Empiricism. A scientific explanation consists of an explanans and an explanandum logically deduced from the explanans. The explanandum consists of at least one statement of antecedent conditions and at least one general law. The explanandum is a description of the empirical phenomenon to be explained, cf. Figure 1 (Hempel & Oppenheim 1948, p. 138).

The scientific explanation by Hempel and Oppenheim can explain concrete phenomena, but also general laws. For instance, the general law L2 “All copper expands when heated.” (explanandum) can be explained by the more general law L1 “All metals expand when heated.” in conjunction with the antecedent condition “Copper is a metal.” (explanans). In the conception of logical empiricism, law L1 is progressive in comparison to L2 because it does not only explain all phenomena L2 can explain, but also phenomena that cannot be explained by L2, e.g. “Tin is a metal”, “This is tin.”, “Therefore, this expands when heated.” (explanandum).

6 The argumentation of subsections 2.1 and 2.2 refers to Craig Dilworth’s introductory book “Scientific Progress” (Dilworth 2007).
7 An example for explaining a concrete phenomenon is: “All copper expands when heated.” (general law), “This is copper.” (statement of antecedent condition), “Therefore, this expands when heated.” (explanandum).
tin expands when heated.” (Dilworth 2007, p. 21). Hempel (1962, p. 100 f.) gives the following example from history of science: “[T]he uniformity expressed by Galileo’s law for free fall can be explained by deduction from the general laws of mechanics and Newton’s law of gravitation, in conjunction with statements specifying the mass and radius of the earth.”

The concept of scientific progress by the Logical Empiricism is based on exactly that deduction of one less comprehensive theory from another more comprehensive one: the more comprehensive theory is then progressive compared to the less comprehensive one. This concept is based on the assumption of a continuity of scientific knowledge, such that proceeding theories logically entail their predecessors.

2.2 Critical Rationalism

The Popperian concept of scientific progress fundamentally differs from that of Logical Empiricism in exactly this point. Popper does not emphasise on continuity, but on contradiction. In response to the Logical Empiricism, he states: “[F]rom a logical point of view, Newton’s theory, strictly speaking, contradicts both Galileo’s and Kepler’s. […] For this reason it is impossible to derive Newton’s theory from either Galileo’s or Kepler’s or both, whether by deduction or induction” (Popper 1973, p. 198).

Whereas Logical Positivism is able to distinguish between more and less progressive theories, but not to explain conflicts of theories Popper’s concept can explain theory conflict; but, in a first step, it fails to distinguish more progressive theories from less progressive ones (Dilworth 2007, p. 26). As a concept of scientific progress is central for each concept of science, Popper claims two criteria that make a theory more progressive in comparison to another one: their content and their verisimilitude.

Popper’s “study of the content of a theory […] was based on the simple and obvious idea that the informative content of the conjunction, $ab$, […] will always be greater than, or at least equal to, that of any of its components. Let $a$ be the statement ‘It will rain on Friday’; $b$ the statement ‘It will be fine on Saturday’; and $ab$ the statement ‘It will rain on Friday and it will be fine on Saturday’: it is then obvious that the informative content of this last statement, the conjunction $ab$, will exceed that of its component $a$ and also that of its component $b$” (Popper 1963, pp. 217-218). For instance, according to Popper, Kepler’s or Galileo’s theory is logically less strong and testable compared to Newton’s one; therefore, Newton’s theory has a richer content and is progressive in relation to Kepler’s or Galileo’s one. Moreover, the content of a theory is inverse to its probability. Obviously, the probability of $ab$ is less than the probability of $a$ or the probability of $b$. Consequently, a less probable theory has a greater content and is therefore ceteris paribus (c. p.) progressive (Popper 1963, p. 218).

The second Popperian criterion for scientific progress, the verisimilitude of a theory, can be described by its nearness to the truth. According to Popper, truth of a theory cannot be attained, but rather approached. A theory $a$ is closer to truth than a theory $b$ if $a$ has more true consequences and less false ones than $b$ (Bird 2008).

2.3 Kuhn’s historical view

Whereas Logical Empiricism and Critical Rationalism have a prescriptive concept of science, i. e. they state how science should be, Kuhn develops a descriptive approach by analysing history of science, i. e. he describes how science factually is. Especially Feyerabend, seizing many suggestions by Kuhn, strongly criticises Popper’s emphasis on falsification and argues that falsification has nearly never been the reason for rejecting a theory (Feyerabend 1989, pp. 250-259).

Instead, Kuhn’s (1970) concept of Scientific Revolution is said to better describe science as it is. According to that concept, each scientific discipline starts in a pre-paradigmatic phase. There, its members even do not agree on fundamental aspects, e. g. they do not speak a common language. When such different fundaments are harmonised usually a phase of Normal Science begins. Normal science is based on a paradigm supported by the research community. Kuhn (1970, pp. 35-42) describes science in such a phase as “puzzle solving”: Normal science is often very efficient, but it does not pro-
duce revolutionary outcomes. Nevertheless, at some moment, scientists begin to question their common fundamentals, e.g. core theories that are suddenly unable to explain newly detected phenomena. A crisis begins and leads to a scientific revolution. New theories are developed and tested until an appropriate candidate for explaining the newly detected phenomena is agreed on. Then again, a period of normal science begins until a new crisis appears.

After a scientific revolution, it might happen that the new paradigm cannot explain every phenomenon as good as the antecedent one (Kuhn 1970, pp. 99-100). Such a partial regress was called “Kuhn-loss” later-on (Bird 2008). Because of this Kuhn-loss, Kuhn’s understanding of scientific progress is not cumulative – as opposed to Logical Positivism and Critical Rationalism. Kuhn moreover denies any concept of nearness to truth. In his point of view, science develops rather evolutionarily, i.e. science develops like an organism that continuously adapts itself to its environment (Kuhn 1970, pp. 170-173).

Kuhn’s and Feyerabend’s conception of scientific progress is strongly marked by their incommensurability thesis. In everyday speech, “incommensurable” means not to have a common standard of measurement. Kuhn uses this term by referring to theories. His view is that a theory is always evaluated by comparing it to a paradigm. Therefore; the standards of theory evaluation are not permanent because paradigms change; they are not theory-independent because they involve a comparison to a (paradigm) theory; and they are not based on rules because they involve perceived relations of similarity (Bird 2008). Consequently, it is very difficult – if not impossible – to compare two theories.

Kuhn differentiates three types of incommensurability (Bird 2008): methodological incommensurability, observational incommensurability, and semantic incommensurability. Firstly, methodological incommensurability is caused by the fact that theories are always evaluated in reference to a paradigm. Therefore, theories having evolved in two different paradigms are not evaluated on the same basis. Secondly, observational incommensurability is caused by theory-dependency of observation. One basis for Kuhn’s argumentation is the gestalt-switch theory developed by Hanson (1958). Hanson uses the example of a picture called bird-antelope (Hanson 1958, p. 87). Usually, at a first view, the observer of that picture sees a bird. But, when somebody tells the observer that it can also be seen as an antelope, the observer usually begins to see an antelope, too. Hanson concludes our perception to be strongly influenced by our previous knowledge. Kuhn (1970, pp. 113-115) transfers Hanson’s discovery to theories and states that all observation is always conducted in the light of the theories of the accepted paradigm. Therefore, observation depends on theory. Consequently, observations conducted against the background of two paradigms cannot be easily compared with each other. This is called observational incommensurability. Lastly, semantic incommensurability is based on a change of the meaning of theoretical terms when a paradigm changes (Bird 2008). For example, the meaning of the word “mass” has slightly changed when Einstein formulated the theory of relativity. In the Newtonian mechanics, mass is a constant concept – e.g. a skyrocket on earth and on moon were considered to have the same mass. Einstein, in contrast, denies mass to be constant by stating that mass and energy are equivalent to each other ($E=mc^2$). Such slight changes in the meaning of theoretical terms also make comparisons of two theories difficult.

Despite all difficulties revealed by Kuhn, he identifies five characteristics “of a good theory” (Kuhn 1977, p. 321):

“First, a theory should be accurate: within its domain, that is, consequences deducible from a theory should be in demonstrated agreement with the results of existing experiments and observations. Second, a theory should be consistent, not only internally or with itself, but also with other currently accepted theories applicable to related aspects of nature. Third, it should have broad scope: in particular, a theory's consequences should extend far beyond the particular observations, laws, or subtheories it was initially designed to explain. Fourth, and closely related, it should be simple, bringing order to phenomena that in its absence would be individually isolated and, as a set, confused. Fifth – a somewhat less standard item, but one of special importance to actual scientific decisions – a theory should be fruitful of new research findings: it should, that is, disclose new phe-
nomena or previously unnoted relationships among those already known” (Kuhn 1977, pp. 321-322).

Kuhn (1977, p. 322) later-on explains that accuracy is important, but would be insufficient as a single criterion. Accuracy comprises explanatory and predictive power.

Nevertheless, two critical remarks referring to these criteria are made by Kuhn himself. Firstly, they are individually imprecise: individuals might differ about their application to concrete cases. Secondly, “when deployed together, they repeatedly prove to conflict with one another” (Kuhn 1977, p. 322).

2.4 Criteria for Scientific Progress

The three important approaches in philosophy of science presented in the preceding sections can hardly be harmonized. Philosophers of science often call the three positions the “Bermuda Triangle of Philosophy of Science”. If one tries to take a position inside that triangle, i.e. to harmonize aspects of at least two of the three positions, he will be caught by a vortex he cannot escape from any more (Kaiser 1991, p. 12). Being aware of this problem, we do not even try to harmonize the three positions - this task should be undertaken by philosophy.

Nonetheless, in order to be able to transfer the results of the discussion in philosophy of science to IS design theories, it is necessary to summarize the three approaches. We propose a set of criteria relevant for the discussion on scientific progress. The summary is illustrated in Table 1.

<table>
<thead>
<tr>
<th>Criterion: Theory T1 is ceteris paribus progressive relative to theory T2 if …</th>
<th>Logical Empiricism</th>
<th>Critical Rationalism</th>
<th>Kuhn</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 … T2 is logically deducible from T1</td>
<td>X</td>
<td>cf. C2</td>
<td>cf. C6</td>
</tr>
<tr>
<td>C2 … T1 has a greater content than T2</td>
<td>cf. C1</td>
<td>X</td>
<td>cf. C4, C6</td>
</tr>
<tr>
<td>C3 … T1 is closer to truth than T2</td>
<td>X</td>
<td></td>
<td>cf. C4</td>
</tr>
<tr>
<td>C4 … T1 is more accurate than T2, i.e. T1 has a greater explanatory power than T2 or T1 has a greater predictive power than T2</td>
<td>cf. C2, C3</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>C5 … T1 is more consistent with itself than T2 or T1 is more consistent with other theories than T2</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>C6 … T1 has a broader scope than T2</td>
<td>cf. C1</td>
<td>cf. C2</td>
<td>X</td>
</tr>
<tr>
<td>C7 … T1 is simpler than T2</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>C8 … T1 is more fruitful of new research findings than T2</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Table 1. Criteria for Scientific Progress in Philosophy of Science

Some of the criteria listed in Table 1 are redundant. The criterion C1 stemming from logical empiricism, i.e. the logical deducibility, is strongly based on an idea of continuity in science that is denied by the two other authors. Nonetheless, if a theory T2 is logically deducible from a theory T1, the scope of theory T1 is broader than the scope (C6) of T2 (or at least as broad as it). Moreover, the special notion of “content” formed by Popper (1973) implies that a theory with a higher content is more accurate than another with a lesser one. Therefore, there is also a relationship between content and accuracy.

The closeness to truth (C3) has some similarities to the accuracy of the theory. Kuhn avoids the term of truth in order to be independent of strong ontological and epistemological positions. The term “truth”, at least if following the correspondence theory of truth, necessitates a position of a scientific realist. Kuhn does not explicitly take such a realist position. Instead, he emphasizes the explanatory
and predictive power of theories and thereby adopts an instrumentalist view. Closeness to truth for a scientific realist is accuracy for a scientific instrumentalist.

We can conclude that the criteria mentioned by Kuhn (1977, pp. 321-322) are – as he states himself – exhaustive in the sense that they cover the criteria we extracted from Logical Empiricism and from Critical Rationalism. Consequently, we will use the following criteria for our further argumentation: accuracy, internal consistency, external consistency, scope, simplicity, and fruitfulness of new research findings.

3 THE NATURE OF DESIGN RESEARCH THEORY

In order to be able to transfer concepts of scientific progress developed for theories in natural sciences to IS DS, the differences between the outcomes of natural sciences and those of design research should be explained. Many publications in DS dealt with epistemological aspects of DR, e.g. with epistemological aspects of the construction and evaluation of design research. Compared to research about epistemological aspects, ontological aspects of design research in IS have been neglected. Nonetheless, in recent years, some authors emphasised on ontological aspects of DR. A seminal article on the “Anatomy of a Design Theory” was published by Gregor and Jones (2007) in which they described core components of an IS design theory.

3.1 On the output of DR: Design Research Artefacts or Design Theories?

Gregor and Jones (2007) write about theories as main outcome of DR. Up to now, there is no consensus whether DR should theorise (e.g. Hevner et al. 2004; Kuechler & Vaishnavi 2008a; Kuechler & Vaishnavi 2008b; Venable 2006; Walls & Widmeyer & El Sawy 1992) or not (e.g. March & Smith 1995). The proponents argue that “[d]eveloping theory is what we are meant to do as academic researchers […]. Theories are practical because they allow knowledge to be accumulated in a systematic manner and this accumulated knowledge enlightens professional practice” (Gregor 2006, p. 613).

Most authors understand design theory as a “utility theory” (Venable 2006) prescribing “how a design process can be carried out in a way which is both effective and feasible” (Walls & Widmeyer & El Sawy 1992). A design theory is a mean to communicate, justify, and develop knowledge in DS cumulatively (Gregor & Jones 2007).

In our paper, we clearly argue for theorising in DR. A synthesis of design knowledge in theories helps to compare results and makes it easier to proof scientific progress, as we will see later-on. Nonetheless, Frank’s (2006) approach is very helpful because it allows us to take a second perspective on design research.

3.2 Components of a Design Theory according to Gregor and Jones (2007)

Gregor and Jones (2007) detail the anatomy of a design theory. Their publication is thereby oriented at a paper published by Walls et al. (1992). The core elements of a design theory identified by Gregor and Jones (2007, p. 322) are listed in Table 2 and explained in the following:

(1) **Purpose and scope** are the characteristic elements of a design theory. Whereas the scope is also relevant for “traditional theories” the purpose is specific to design theories as they are “utility theories” (Venable 2006).

(2) **Constructs** represent the entities of interest in the design theory. Frank (2006) also emphasises on the language the artefact is represented in; he thereby differentiates formal, semi-formal, and natural languages. Constructs are also elements of traditional theories as Dubin (1978) states.
### Core components

1) **Purpose and scope** (the *causa finalis*)
   - "What the system is for," the set of meta-requirements or goals that specifies the type of artefact to which the theory applies and in conjunction also defines the scope, or boundaries, of the theory.

2) **Constructs** (the *causa materialis*)
   - Representations of the entities of interest in the theory.

3) **Principle of form and function** (the *causa formalis*)
   - The abstract “blueprint” or architecture that describes an IS artefact, either product or method/intervention.

4) **Artefact mutability**
   - The changes in state of the artefact anticipated in the theory, that is, what degree of artefact change is encompassed by the theory.

5) **Testable propositions**
   - Truth statements about the design theory.

6) **Justificatory knowledge**
   - The underlying knowledge or theory from the natural or social or design sciences that gives a basis and explanation for the design (kernel theories).

### Additional components

7) **Principles of implementation** (the *causa efficiens*)
   - A description of processes for implementing the theory (either product or method) in specific contexts.

8) **Expository instantiation**
   - A physical implementation of the artefact that can assist in representing the theory both as an expository device and for purposes of testing.

---

**Table 2:** Components of an IS Design Theory (Gregor & Jones 2007, p. 322)

(3) **Principles of form and function** constitute the architecture of an IS artefact. Principles of form and function also exist for traditional theories: Dubin (1978) refers to “laws of interaction” in reference to traditional theories.

(4) The **artefact mutability** refers to changes in the state of the artefact anticipated in the theory, e.g. suggestions for improving the approach in further works. For traditional theories, there is no equivalent to artefact mutability.

(5) **Testable propositions** are truth statements about the design theory. They are a central element of traditional theories; Dubin (Dubin 1978) names them simply “propositions”.

(6) **Justificatory knowledge** is the underlying knowledge or theory from natural or social sciences. Frank (2006) differentiates a variety of justification methods depending on different underlying concepts of truth: correspondence theory of truth, coherence theory of truth, and consensus theory of truth. For each of them, he mentions at least one validation method: for correspondence theory of truth experiment, field study, or case study; for coherence theory literature analysis; and for consensus theory of truth (virtual) discourse according to Habermas (1984). Dubin (1978) does not mention an equivalent element for traditional theories. According to Gregor and Jones (2007), this is due to the fact that Dubin, being realist, explicitly does not emphasize on explanation.

(7) **Principles of implementation** describe the process for implementing the theory in a specific context. This includes situational adaptation of artefacts (cf. Bucher et al. 2007). For traditional theories, Dubin (1978) does not mention an equivalent.

(8) **Expository instantiations** of the theory are physical implementations of the artefact. Often, they are used for testing the artefact. As well as principles of implementation, Gregor and Jones consider them as optional. Frank (Frank 2006) mentions “prototyping” as one method for justifying the artefact. In traditional theories, hypotheses and empirical indicators are very similar to an expository instantiation (Dubin 1978).
4 CRITERIA OF SCIENTIFIC PROGRESS OF DESIGN THEORIES

After having shortly summarised aspects of scientific progress from philosophy of science in section 2 and structural components of DR artefacts and DR theories respectively in section 3, we now transfer the criteria originally created for explanatory and predictive theories to DS. We thereby take, as a starting point, the five criteria for scientific progress by Kuhn (1977, pp. 322-323) which are, as explained in section 2.4, under some presuppositions exhaustive: accuracy, consistency, scope, simplicity, and fruitfulness of new research findings.

**Accuracy** of a theory in natural sciences is composed by its explanatory and predictive power. Theories in DS do not aim at explaining or predicting, but at being useful. Design Theories are therefore also called utility theories (Venable 2006).

Nonetheless, utility is mentioned as a component of DR [theory] neither by Frank nor by Gregor and Jones. They both mention the purpose of an artefact. The usefulness of an artefact can be defined as its ability to fulfil its purpose if the purpose itself is useful, i.e. relevant. The testable propositions claimed by Gregor and Jones operationalise the usefulness of the artefact. If the testable propositions are exhaustive, the accuracy of a DR theory corresponds to the truth of its testable propositions.

By **consistency**, Kuhn (1977, pp. 321-322) means both internal and external consistency.

The **internal consistency** of a design theory involves nearly all its components. One necessary internal consistency has already been mentioned: the consistency between testable propositions and the purpose of the artefact. The purpose of the artefact should be its leading property because the identity of an artefact is strongly linked to its purpose. The only component of a design theory which cannot be consistent to its purpose is its scope – both concepts are orthogonal one to another. Moreover, the testable propositions should be consistent to all other components of a design theory. A special role in determining the consistency of an artefact is played by its principles of form and function, i.e. its architecture. They give an overview of the artefact and facilitate internal consistency checks.

By **external consistency**, Kuhn means consistency with “other currently accepted theories applicable to related aspects of nature” (Kuhn 1977, p. 322). A design theory should therefore be consistent to both other accepted IS design theories and behavioural theories in IS. External consistency should be shown by a literature review. For external consistency; Gregor and Jones (2007) identify the design theory component of justificatory knowledge; Frank (2006) mentions adequacy according to the correspondence theory of truth.

The broader the **scope** of a theory the better is the theory. The scope of a theory should not only be mentioned, but it should also be justified. The internal consistency of a theory necessitates that testable propositions are formulated and tested such that the whole scope of the theory is covered.

**Simplicity** is often claimed as a quality factor of theories. Already, William of Ockham said that entities must not be multiplied beyond necessity. Albert Einstein stated to make everything as simple as possible, but not simpler. Simplicity is related to the explanatory power of theories. A complex explanation might be correct, but also might be that complex that it cannot be understood any more. Both, simplicity and explanatory power lead to understanding. In DS theory, understanding is not the primary goal, but utility. The utility of a DR artefact or of a design theory is only slightly affected when it is complex, but strongly when it is expensive, i.e. inefficient. Therefore, the efficiency of a utility theory is the equivalent to the simplicity of an explanatory theory. Simplicity, however, is still important: An efficient design theory implies its simplicity. As the understanding of the theory is time consuming for the practitioner who wants to apply it, an unnecessarily complex theory is inefficient compared to a simpler one.

The efficiency of a design theory should mainly be shown by formulating and testing propositions that show the efficiency of the artefact. The simplicity of a design theory (in the narrow sense of the word...
simplicity) mainly refers to the constructs, i.e. the language used, and to the principle of form and function of the artefact, i.e. its architecture.

The **fruitfulness of a design theory for further research** is a criterion that is only mentioned by Kuhn, i.e. neither by Logical Empiricism nor by Critical Rationalism. The fruitfulness for further research is – as Kuhn says himself – compared to the other criteria “a less standard item” (Kuhn 1977, p. 322). Fruitfulness for further research can only be determined in a historical analysis of a theory. Therefore, it is difficult to relate fruitfulness for further research to actual components of a design theory.

In summary, progress is a normative concept and can be defined as a transition from stage $A$ to stage $B$ whereby $B$ is better than $A$ in respect to certain criteria (Niiniluoto 2007). In order to identify such criteria for DS, we took five criteria for theories in natural sciences formulated by Kuhn (1977, pp. 322-323) and transferred them to DS. Four of these five criteria could be meaningfully transferred to DS and put in relation to components of a design theory. Some of the criteria changed in meaning. The final quality criteria for a “good” design theory are utility, internal consistency, external consistency, scope, and efficiency.

### 5 FINAL REFLECTION AND OUTLOOK

Scientific progress is a core aim of science. The subject therefore has been treated by philosophy of science for many years. The evolution from Logical Empiricism whose concept is based upon the relatively naïve assumption of continuity in science over the concept of scientific progress by Critical Rationalism to the historical, i.e. descriptive, view by Kuhn shows not only the complexity and diversity of the subject, but makes us also understand that the question of scientific progress is not even finally solved in philosophy of science.

In this paper, five criteria by Kuhn (1977, pp. 322-323) are applied to the structure of design theory. In summary, a good design theory is useful, internally and externally consistent, it has a broad scope, and its artefact is efficient. A theory $T_1$ that is better in at least one of these criteria and not worse in another criterion compared to a theory $T_2$ is progressive in reference to $T_2$.

The discussion of the problems related to scientific progress formulated by Kuhn (1970), i.e. the methodological, observational, and semantic incommensurability, have shown that scientific progress is a very complex concept. We can learn from Kuhn that it is important to consider the paradigm in which a theory is formulated. A theory can only be evaluated towards such a paradigm. Therefore, it is necessary to know actual paradigms. In order to detect such paradigms, rigorous historical research is necessary (Mason & McKenney & Copeland 1997). It is arguable whether, in our discipline, we are by now in a phase of normal science or still in a pre-paradigmatic phase. A common language is a fundamental requirement for a discipline in a phase of normal science. Such a common language does not broadly exist. Research about different language communities clearly shows that (Schelp & Winter 2008; Schelp & Winter 2009). Such research is very fruitful for IS because it allows for a better understanding of our research outcomes. Difficulties arising from different paradigms in reference to the discussion of scientific progress in IS DS could not be sufficiently discussed in this article. More research in this field is necessary.

Moreover, the subject of scientific progress was only treated on an ontological level. We proposed a criteria-based definition of scientific progress in DS. Besides research focussing on paradigms IS, at least three more research questions concerning scientific progress in design science follow: Firstly, an epistemological question can be posed: How can we detect scientific progress in DS? The criteria developed are a good basis for answering such a question. Moreover, at least two praxeological questions follow: How can we ameliorate DR processes in order to simplify the detection of scientific progress; are new guidelines in addition to those proposed by Hevner et al. (2004) helpful or even necessary? Lastly, we can ask how to construct artefacts that are progressive.
References


Kuechler, B. and Vaishnavi, V. (2008b) Theory Development in Design Science Research: Anatomy of a Research Project, Third Conference on Design Science Research in Information Systems and Technology (Eds, Vaishnavi, V. and Baskerville, R.) Georgia State University, Atlanta, Georgia, USA.


GAME-THEORETIC ANALYSIS OF PAY-AS-BID MECHANISMS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0001.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Marketplaces, Auctions, Supply chain management, IT Strategy</td>
</tr>
</tbody>
</table>
GAME-THEORETIC ANALYSIS OF PAY-AS-BID MECHANISMS

Meinl, Thomas, Universität Karlsruhe (TH), Englerstr. 14, 76133 Karlsruhe, Germany, meinl@iism.uni-karlsruhe.de
Stößer, Jochen, Universität Karlsruhe (TH), Englerstr. 14, 76133 Karlsruhe, Germany, stosser@iism.uni-karlsruhe.de
Neumann, Dirk, Albert-Ludwigs-Universität Freiburg, Platz der Alten Synagoge, 79085 Freiburg, Germany, dirk.neumann@vwl.uni-freiburg.de

Abstract

Enterprises are facing a challenging dilemma. In order to be able to accommodate peak loads on their IT systems, they must maintain large computing clusters, which lie idle most of the time. At the same time, IT departments are under constant pressure to cut down on hard- and software expenses. Grid technology offers a promising way out of this dilemma by allowing the dynamic sharing both within enterprises as well as across organizational boundaries. This sharing approach, however, requires proper economic incentives. This paper is concerned with the determination of dynamic market-based prices. Due to their simplicity, so-called pay-as-bid mechanisms have become popular. This paper is novel as we provide an in-depth analysis of two such pay-as-bid mechanisms – Proportional Share and a discriminatory pay-as-bid mechanism – for the case of three users, thus extending previous work by Sanghavi and Hajek (2004) and Stößer et al. (2008). This analysis is important as we show that the nice results for two users cannot be retained once three or more users are present. Even worse, we show that these results can even be reversed if we move to games with more than two players.

Keywords: Pay-as-bid Mechanisms, Game-theoretic Analysis, Distributed Resource Allocation.
1 INTRODUCTION

Increasingly complex applications require massive amounts of processing power while exhibiting fluctuating utilization patterns. One example for such applications are CAx applications such as Computer Aided Engineering and Computer Aided Design or so-called “digital factory applications”, e.g. for production planning. As a result, companies pile up resources to accommodate few peak loads on their systems. In a meta-study, Carr (2005) reports that data centers are only using between 10-35% of their available processing power, leading to tremendous inefficiencies. This development is exacerbated by intra-organizational boundaries. Even business units and departments within one enterprise are unwilling to share idle resources with other units.

Grid technology offers a promising way out of this dilemma. Using virtualization technologies, physical resources can be dynamically assigned to applications according to these applications’ priority. Thus, if a unit experiences a peak load on its system, it can accommodate this peak load on remote resources, and analogously, in case an organizational unit temporarily has idle resources, it can host applications of other units. These applications can be forced out and moved to other resources as soon as the unit falls short of resources itself. In essence, Grid technology yields more efficient resource utilization as the system load is distributed across all organizational units. This sharing approach, however, requires proper economic incentives. Organizational units will only contribute their resources to the Grid if they get something in return. At the same time, units need to charge resource usage to avoid excessive usage. Both issues can be addressed by introducing dynamic prices. This paper is concerned with the determination of dynamic prices that are simple enough to be useful in Grids. Due to their simplicity, so-called pay-as-bid mechanisms have become popular – the user simply pays what he bid, and receives a share of the resource according to some specific allocation rule which relates this user’s bid to the other users’ bids.

This paper is novel as we provide an in-depth analysis of two such pay-as-bid mechanisms – Proportional Share and a discriminatory pay-as-bid mechanism – for the case of three users, thus extending previous work by Sanghavi and Hajek (2004) and Stößer et al. (2008). This analysis is important as we show that the nice results for two users cannot be retained once three or more users are present. Even worse, we show that these results can even be reversed if we move to games with more than two players.

This paper is structured as follows. In Section 2, we introduce a sample scenario which illustrates the business case for Grids. Section 3 discusses previous work on Grid market mechanisms. At the core of this paper, Section 4 provides a game-theoretic analysis of Proportional Share and the discriminatory pay-as-bid mechanism introduced by Sanghavi and Hajek (2004). In Section 5, we discuss our analytic results. Subsequently, Section 6 concludes the paper and points to future research directions.

2 MOTIVATIONAL SCENARIO

In the following, we showcase a scenario where market mechanisms create value in real business cases. The scenario is TXTDemand, an application by TXT e-solutions’ for forecasting demand and replenishment within a supply chain. The application combines CPU intense forecasting algorithms as well as algorithms for the analysis of historic data and current sales data with interactive revision tools. In a typical business scenario, a customer is running the application daily for defining demand and replenishment strategies. During the night batches process the previous days’ sales data to define initial forecast plans. For example, the replenishment module generates optimized replenishment plans, taking historic inventory levels, future demand forecasts and the type of the market into

\[http://www.txtgroup.com/\]
consideration. The plans are analyzed and refined by the user at the other day in the interactive mode. The rationale for user interaction in the planning process is that users have more in-depth knowledge about the context. For example, if a certain fair takes place for the first time, the replenishment plan generated from historical data are most likely ignoring this. The user can, hence, significantly improve the replenishment plan by conducting as-if scenario analyses, refining the automatically produced plan.

The calculations of the replenishment module are particularly resource demanding when used in real business due to the very large amounts of data surpassing several hundred millions of entries. Grid computing can help by distributing the calculations on several machines. By pooling resources of several providers, the costs for IT infrastructure can be significantly reduced. Since the calculations are so demanding, the total cost of ownership for one company alone would make the use of the forecasting module unprofitable. It is the pooling of resources in a Grid that amortizes the use of the replenishment module. As the pure sharing of resources without compensation is hampered by free-riding behavior, market mechanisms seem to work well, establishing the right incentives for the participants.

The market mechanisms need to cope with the peculiarities of the domain. We identify two main requirements for the market mechanism. In case of interactive applications, the requests for Grid resources depend on daily human interactions. Thus there can be unpredictable peaks of requests, occurring at any second. This requires from the mechanism to attain an allocation of resources in near real-time. Furthermore, from a technical viewpoint, avoiding starvation is an important objective. In scheduling theory, starvation denotes the fact that low-priority tasks are prevented from doing any progress because all resources are assigned to other higher-value tasks. Combining the economic and the technical viewpoint, it can be desirable to give “better” service to high-value applications but to also give at least “some” service to low-value applications.

3 RELATED WORK

A bulk of mechanisms has been proposed for Grid resource allocation (see Wolski et al. (2001) and Neumann et al. (2008) for surveys). In AuYoung et al. (2004), Bapna et al. (2006) and Schnizler et al. (2006), the scheduling problem in Grids is formalized as a combinatorial allocation problem. Bapna et al. (2006) and Stößer et al. (2007) present greedy heuristics to mitigate the resulting computational complexity. While these mechanisms allow for dependencies between multiple Grid resources (e.g. CPU and memory), they are based on strong informational assumptions, such as complete knowledge about the time constraints and resource requirements of applications. However, in practice the users themselves will only have fuzzy knowledge about this information, as can be seen when comparing the user estimates and actual runtimes and resource requirements in workload traces from supercomputers.

A fundamentally different approach is taken by mechanisms that continuously assign resource shares to applications. These shares are continuously updated as new user requests are submitted or tasks are completed, thus allowing for real-time allocations. With an allocation rule purely based on economic reasoning (e.g. the prominent Vickrey auction), all available resources would be given to one single user, with the highest valuation. However, this will lead to the starvation of other tasks with lower value. Consequently, considering the requirements of our scenario, in this work we will consider two different allocation mechanisms: Proportional Share (Chun and Culler 2000, Lai et al. 2004) and the discriminatory pay-as-bid mechanism by Sanghavi and Hajek (2004).

Proportional Share is well understood and has been implemented by Hewlett Packard in its Tycoon system (Lai et al.). If user $i$ reports a valuation of $w_i$, he will receive a fraction of the resources amounting to $\frac{w_i}{\sum_{j=1}^{n} w_j}$.

Proceedings ECIS 2009
The discriminatory pay-as-bid mechanism has been proposed by Sanghavi and Hajek (2004) for bandwidth allocation in computer networks. It works slightly different than Proportional Share, giving a discount to the high-value user, thus resulting in a lower unit price than the low-value users. The idea is to encourage the high-value user to bid close to his true valuation rather than to shade down his bid. We will denote the allocated share for each user \( i \) of \( n \) users for Proportional Share with \( \tau_{ps}^i = \frac{w_i}{\sum_{j=1}^{n} w_j} \) and the Sanghavi-Hajek allocation rule with \( \tau_{sh}^i = \frac{w_i}{w_{\text{max}}} \int_0^1 \prod_{j \neq i} \left( 1 - s \frac{w_j}{w_{\text{max}}} \right) ds \).

Sun Microsystems is considering to integrate the latter mechanism into their Sun Grid Engine (Stößer et al. 2008).

4 THE MODEL

One of the main goals of this work is to compare the performance of the two given allocation mechanisms and thus to conclude which of these may be more appropriate for use in typical Grid scenarios. Clearly our work is abstract and can be applied to other domains as well though in this work we restrict our attention to Grids as the mechanisms can be applied there reasonably.

In general the performance of mechanisms can be measured analytically, through simulations or with laboratory experiments. In this work though we will focus on analytical evaluation since it is well known that a common metric for measuring a mechanisms performance analytically is given by computing the performance ratio in its Nash equilibrium. Therefore we outline the setting used in our work and give a short introduction into Nash equilibria followed by the main analysis of the Nash equilibria of the two allocation mechanisms, Proportional Share and Sanghavi-Hajek. We will analyze under which conditions (multiple or unique) Nash equilibria exist for the case of more than two users. It is well known that the results from 2-player games cannot always be generalized to n-player games. Since this applies to our analysis as well, we will consider three players first, showing the tremendous challenges and difficulties as well as the rigorous conditions one encounters even for the simple case of one additional player.

4.1 The Setting

We assume there is an auction with one seller who offers a perfectly divisible good (i.e. there are no constraints in how many parts the good may be split up nor the size of these parts) and \( n \) buyers, each one offering a bid \( w_i \) and receiving a share through the utilized allocation mechanism \( \tau \). Let \( w = (w_1, ..., w_n) \) be the vector of bids given by the users. We also assume \( w_i \geq 0 \) for \( i = 1, ..., n \) with at least one \( w_i > 0 \), since a positive \( w_i \) represents the payment made by user \( i \). At this point we may already say, that we can exclude any Nash equilibria, where negative bids \( w_i \) are submitted, since this would mean that user \( i \) would receive a payment by the seller for his allocated part of the good and thus would turn into a seller, also leading to a contradiction to the assumption of a given non-negative payment vector. The vector \( x = (x_1, ..., x_n) \), \( x_i \in \mathbb{R}_+^0 \), \( \sum_{i=1}^{n} x_i = 1 \), denotes the actual allocation, where the shares \( x_i \) for each user are allocated according to the allocation mechanism \( \tau_{sh} \) for the Sanghavi-Hajek mechanism and \( \tau_{ps} \) for Proportional Share respectively, where each mechanism depends on the payment vector \( w \). Furthermore we assume in line with auction theory (Mas-Colell et al. 1995) that each user has a quasi-linear utility function \( u_i(x) = v_i x_i - c(x_i) \), with \( v_i \in \mathbb{R}_+^n \) denoting the valuation of user \( i \) for the good and \( c(x_i) = p_i x_i \) the linear price function, where \( p_i \) denotes user \( i \)’s unit price, i.e. the price user \( i \) would have to pay if she were awarded with the whole resources. As was shown in Johari and Tsitsiklis (2004), when evaluation the worst-case-performance ratio in our setting, it is enough to look at linear utility functions.

Additionally we use in our setting a so-called pay-as-bid pricing scheme, so that we have \( c(x_i) = p_i x_i = w_i, i = 1, ..., n \). Thus we can reformulate the utility functions to \( u_i(w) = v_i \tau_i(w) - w_i \).
4.2 Nash Equilibria

Since the main goal of this section is to analyze the possible Nash equilibria for the several given allocation mechanisms for two, three and \( n \) users, the concept and the idea behind this equilibrium needs to be clear. One of the main goals of mechanism design is to construct these mechanisms in a way that they are incentive compatible, i.e. each user bids his true valuation of the good and has no advantage of reporting false valuations, because it will not lead to any improvement for him. In this situation the concept of Nash equilibria is widely used (cf. Sanghavi and Hajek 2004).

**Definition 1. (Nash equilibrium)** A bid vector \( \mathbf{w}^{NE} \) is a *Nash equilibrium*, if \( u_i(\mathbf{w}_i^{NE}) \geq u_i(\mathbf{w}_i, \mathbf{w}_-^{NE}) \) for all \( i = 1, \ldots, n \), meaning that no user \( i \) can benefit by unilaterally deviating from his equilibrium bid \( \mathbf{w}_i^{NE} \).

For a given allocation mechanism there may exist no, one unique or several Nash equilibria. We will analyze under what conditions we may retrieve no equilibrium but by changing the realizations of the users bids and/or valuations slightly we obtain a unique or several Nash equilibria. Also we have not only to worry about multiple Nash equilibria but also acceptable (i.e. feasible) ones, meaning that we will not accept any solution resulting in a contradiction to our assumptions as already pointed out above. Having said this we will speak of a feasible Nash equilibrium if an equilibrium chosen out of a given set of several Nash equilibria fulfills all taken assumptions and leads to no contradiction at all. Therefore for any possible solution we have to check afterwards if the foregoing assumptions and conditions still hold.

Given the term of a Nash equilibrium we may define the performance ratio of an allocation mechanism:

**Definition 2. (Performance ratio)** Suppose a set of \( n \) users each having a utility function \( u_i \) and a provider selling the perfectly divisible good with a utility function \( u_p \). Let the bid vector \( \mathbf{w}^{NE} \) be a Nash equilibrium. Then the *performance ratio* of a given allocation mechanism \( \tau \) is given by

\[
\frac{u(\tau(\mathbf{w}^{NE}))}{u^*} = \frac{\sum_{i=1}^{n} u_i(\tau(\mathbf{w}^{NE})) + u_p(\tau(\mathbf{w}^{NE}))}{u^*}
\]

where \( u^* \) denotes the theoretical optimum giving the whole good to the user with the highest bid.

The subsequent analysis will be along the following lines: We will compute the resulting Nash equilibria for \( n \) users (\( n \in \{2, 3\} \)) for the Sanghavi-Hajek as well as the Proportional Share mechanism. We will show under which conditions these equilibria exist and to what outcome (utility function) they entail. For this we use the following, well known proposition:

**Proposition 3.** For a set of given users \( i, i = 1, \ldots, n \), the payment vector \( \mathbf{w}^{NE} \) is a Nash equilibrium, if and only if \( \frac{\partial u_i(w^{NE})}{\partial w_i^{NE}} = 0 \) for all \( i = 1, \ldots, n \).

Proposition 3 simply means that every agent maximizes his own utility in conjunction with all bids given by the other agents.

4.3 Two Users

Stößer et al. (2008) analyze the case with two requesters. We will briefly report the main results for the sake of completeness. We assume that the two users have the valuations \( v_1 \) and \( v_2 \) and without loss of generality user 2’s valuation is not less than user 1’s, i.e. \( v_1 \leq v_2 \).

**Lemma 4.** In the Nash equilibrium \( \mathbf{w}^{NE} \) of the pay-as-bid mechanism \( \tau^{sh} \), user 1 bids \( w_1^{NE} = \frac{v_1^2}{2v_2} \) and receives a share of \( \tau^{sh}_1(\mathbf{w}^{NE}) = \frac{v_1}{2v_2} \), whereas user 2 bids \( w_2^{NE} = \frac{v_1}{2} \), thus receiving \( \tau^{sh}_2(\mathbf{w}^{NE}) = 1 - \frac{v_1}{2v_2} \).
Inserting the above results in the utility functions \( u_i \) for every user \( i \), shows that the low-value user receives zero utility, while the high-value user has \( u_2(w^{NE}) = v_2 - v_1 \).

**Lemma 5.** In the Nash equilibrium \( w^{NE} \) of the pay-as-bid mechanism \( \tau^{PS} \), user 1 bids \( w_1^{NE} = \frac{v_1 v_2}{v_1 + v_2} \) and receives a share of \( \tau_1^{PS}(w^{NE}) = \frac{v_1}{v_1 + v_2} \), whereas user 2 bids \( w_2^{NE} = \frac{v_2}{v_1 + v_2} \), thus receiving \( \tau_2^{PS}(w^{NE}) = \frac{v_2}{v_1 + v_2} \).

Consequently, for two users there exists a unique Nash equilibrium without implying further restrictions on the valuations \( v_i \) of the users. Unfortunately, as we will show below, this does not hold any longer for three or more users.

### 4.4 Three Users

Now we will show our results for three users followed by additional examples demonstrating that our argumentation is not based on an empty set but that there exist scenarios allowing us to apply our analytic results. To illustrate our approach for each proposition we sketch the proofs which are straightforward but very in the appendix. We begin with the results of our analysis for the proportional Share mechanism for three users and subsequently move to the Sanghavi-Hajek mechanism.

#### 4.4.1 Proportional Share

In this section we will analyze whether there exist a unique Nash equilibrium when there are three users present. Our main result is captured by Proposition 6, which is quite powerful, as it shows the conditions for which feasible Nash equilibria exist.

**Proposition 6.** For three users with quasi-linear value functions there exists no unique Nash equilibrium for the Proportional Share allocation mechanism. Given the restrictions

\[
v_3 \geq \frac{v_1 v_2}{v_1 + v_2}, \quad v_2 \geq \sqrt{v_1 w_1} \quad \text{and} \quad v_1 \geq w_2 + w_3
\]

there exists a feasible equilibrium.

**Proof.** See Appendix.

Proposition 6 illustrates clearly that there does not exist a unique but two Nash equilibria, from which only one is feasible as it avoids payments from the seller to the user. The restrictions imply certain conditions (i.e. lower bounds) on the realizations of the valuations of the users but are not unlikely to happen in real-world scenarios, since they though being strong are nevertheless easy to check and fulfill. For the sake of simplicity we state them as a composition of constraints between the bids and the valuations, but since the bids depend unambiguously on the valuations only, more complex terms may be stated just using the users valuations. For more details we refer the reader to the appendix.

**Example 7.** We will give two simple examples, which illustrate this proposition. Assume there is a provider offering a certain amount of Terabytes of storage and three users having their own valuation and willing to pay a certain amount to achieve a share of the storage. The valuations are common knowledge and each user bids strategically trying to maximize his own utility.

- **a)** We assume the valuations to be \( v_1 = 1, v_2 = 2 \) and \( v_3 = 3 \). Using the results above we can compute the only payment vector \( w = \left( \frac{-12}{121}, \frac{60}{121}, \frac{84}{121} \right) \). The negative value of \( w_1 \) is the result of the third condition of the above proposition not being fulfilled, as \( w_2 + w_3 = \frac{144}{121} > 1 = v_1 \). Thus there exists no Nash equilibrium even for this simple case.

- **b)** Changing the valuations given in a) slightly, yields us a feasible outcome and Nash equilibrium: Assuming that we have the realization of the valuation vector \( v = (2, 3, 4) \) this
yields \( w = \left( \frac{24}{169}, \frac{120}{169}, \frac{168}{169} \right) \) and the following vector of utilities \( u = \left( \frac{2}{169}, \frac{75}{169}, \frac{196}{169} \right) \) thus amounting to a total revenue for the provider of \( \sum_{i=1}^{3} w_i = \frac{273}{169} \).

### 4.4.2 Sanghavi-Hajek Mechanism

In this section we analyze Nash equilibria for the Sanghavi-Hajek mechanism. These results differ very much from the ones we achieved for the Proportional Share mechanism.

**Proposition 8.** For three users, the Sanghavi-Hajek allocation mechanism combined with a pay-as-bid pricing scheme and quasi-linear utility functions, we have

a) A unique Nash equilibrium for the realization \( v_3 = \frac{1}{3} \frac{v_1 v_2 (3(v_1 + v_2) + 4 \sqrt{v_1 v_2})}{9(v_1^2 + v_2^2) - 14v_1 v_2} \) and the condition \( \frac{8}{9} v_1 \leq v_2 \leq \frac{9}{8} v_1 \), resulting in zero utility for the lower-bidding users and non-negative utility for the highest-bidding user.

b) Multiple Nash equilibria for \( v_3 > \frac{1}{3} \frac{v_1 v_2 (3(v_1 + v_2) + 4 \sqrt{v_1 v_2})}{9(v_1^2 + v_2^2) - 14v_1 v_2} \) and the realizations of the valuations of the three users being so that \( 2w_{3 \pm} \leq v_i \leq 3w_{3 \pm} \) for \( i = 1, 2 \) holds.

**Proof.** See Appendix.

As we learn from Proposition 8, for the Sanghavi-Hajek mechanism a unique Nash equilibrium may exist, but occurs under very restrictive assumptions on the users valuations. We can state immediately that in a real-world scenario the exact realization of \( v_3 \) as given above for a unique equilibrium is very unlikely to happen and tends to converge to zero. This is due to the fact that users usually will have expectations only about the valuations of the other users. Multiple Nash equilibria may also exist under less restrictive conditions though these conditions as seen in Proposition 8 depend on the highest given bid \( w_3 \). A strict dependency on the valuations of the users only can’t be given since \( v_3 \) is not determined but can be chosen freely above the limit.

**Example 9.** To clarify our results, we give two short examples assuming the same scenario as already given in Example 7.

a) Taking into account the conditions involved for finding a unique Nash equilibrium, we assume the realizations of the valuations of the lower-bidding users to be \( v_1 = v_2 = 1 \) and \( v_3 = 2 + \frac{3}{2} \sqrt{2} \). Thus the users may bid the strategy \( w_3 = \frac{4 + 3 \sqrt{2}}{4 + 3 \sqrt{2}} \) and \( w_1 = w_2 = \frac{3(2 + \sqrt{2}) (4 + 3 \sqrt{2})}{6 (3 + 2 \sqrt{2})} \) resulting in a utility vector \( u = (0, 0, \frac{1}{12(3 + 2 \sqrt{2})}) \).

b) If given the valuation vector \( v = (3, 3, 5) \) all necessary conditions are satisfied, resulting in a zero utility for the lower-bidding users and a positive utility for the highest bidder.

It is obvious that the conditions necessary for having a Nash equilibrium for both allocation mechanisms are different. The question which arises now is if the conditions given for one mechanism are stronger than the ones of the other meaning that if we have given one Nash equilibrium of one mechanism we would obtain automatically an equilibrium of the second allocation mechanism and thus allowing us to compare these mechanisms one to another. Fortunately we have the following result:

**Corollary 10.** Assume a unique Nash equilibrium \( w^{NE} \) for the Sanghavi-Hajek mechanism. Then \( w^{NE} \) is also a feasible equilibrium for Proportional Share.

**Proof.** See Appendix.

With this we may now compare the performance ratio of the two given allocation mechanisms assuming we have the conditions fulfilled for a unique Nash equilibrium for Sanghavi-Hajek. Since
the proof is a simple computation and straightforward it is omitted. We just state the final result of our paper:

**Corollary 11.** Assume a unique Nash equilibrium $w^{NE}$ for the Sanghavi-Hajek mechanism. Then the performance ratio for the Proportional Share mechanism exceeds the ratio of the Sanghavi-Hajek mechanism for all feasible realizations of $v_1$ and $v_2$.

This outcome is very surprising indeed as it is in opposition to the result obtained by Stößer et al. (2008) for two users and will be discussed in the following section.

**Corollary 12.** Assume a unique Nash equilibrium $w^{NE}$ for the Sanghavi-Hajek mechanism. Then the revenue for the Proportional Share mechanism is exceeded by the revenue of the Sanghavi-Hajek mechanism for all feasible realizations of $v_1$ and $v_2$.

This result coincides with the work for 2-players done by Stößer et al (2008).

## 5 DISCUSSION

Our goal is to compare market mechanisms which can be integrated into existing Grid schedulers to achieve efficient resource allocations and dynamic prices. As pointed out earlier, there are three possible techniques to do such a comparison: game-theoretic analysis, numerical experiments, and laboratory experiments. In this paper, we report the results of our game-theoretic approach.

Earlier work on Proportional Share and the discriminatory pay-as-bid mechanism for the two user case has shown that the Sanghavi and Hajek mechanism is superior to Proportional Share with respect to overall efficiency and, given sufficiently close valuations of the two users, also as regards provider’s revenue (Stößer et al. 2008). The aim of our analysis in this paper was to generalize these results for more than two users. To this end, we need to model the user behavior, i.e. how the users bid based on their true valuations. We chose to ground our analysis on the assumptions of quasi-linear utility functions and the prominent solution concept of Nash equilibria. We were able to derive conditions on the existence of such equilibria for both the Proportional Share and the Sanghavi and Hajek mechanism. However, the interpretation of these equilibria for the Sanghavi and Hajek mechanism gives rise to four issues:

- The two low-bidding users are pushed to zero utility while the high-bidding user only achieves a slightly positive utility, meaning that the provider reaps almost all the welfare generated by the mechanism. This corresponds to the results of Stößer et al. (2008) for the two user case. Consequently, the users are indifferent to not participating in the mechanism. One possibly remedy could be that the provider gives a small kickback to the low-bidding users to encourage them to participate. It can be an interesting question for future research to investigate how this impacts the users’ strategic considerations.

- The user with the lowest valuation ends up being the highest bidder. We explain this somewhat paradoxical result by the fact that in order to achieve a (unique or feasible) Nash equilibrium we had to set up some (rigorous) restrictions on the valuations of the users. Obviously this cannot be done in real-world scenarios and though this procedure is legitimate from a mathematical point of view, it may result in odd economic interpretations.

- The unique Nash equilibrium comprises some interesting cross-dependencies between the given bids. Assume two users bid according to their equilibrium strategy. If one of the low-value users deviates from his bid, this will not affect his own utility, but instead result in a decrease of the other users’ utilities.

- Interestingly, as shown in Corollary 11, the analytic results for the performance ratio reverse if we move from a scenario with two users to a scenario with three users. While with two users the Sanghavi-Hajek mechanism always generates higher efficiency than Proportional Share, with three users Proportional Share dominates the Sanghavi-Hajek mechanism.
In summary, our analysis showed that complexity increases tremendously already when adding only one more user. We need to introduce strong restrictions on the users’ valuations in order to obtain feasible Nash equilibria, let alone unique equilibria. While the analytic approach is attractive due to its elegance, this strikingly shows the limitations of the analytic approach, as these equilibria are unlikely to occur in real-world scenarios.

This is exacerbated by the fact that we assume a scenario with complete information. In practice, users will at most have a rough estimate about the valuations of other users. Including such Bayesian solution concepts into our model would surely drive complexity beyond analytic feasibility.

In conclusion, we need to rely on numerical simulations and/or laboratory experiments when trying to analyze these mechanisms in more complex and realistic settings.

6 CONCLUSION & FUTURE WORK

In this article, we have approached the concept of scheduling in Grid systems from an economic viewpoint. Due to the strategic nature which is inherent to these systems, market-based schedulers are deemed promising to increase the efficiency of such systems, while at the same time providing incentives to contribute resources to the Grid. So-called pay-as-bid mechanisms score with their ease-of-use. Furthermore, they impose only a very low communicational and computational burden on the scheduling process and consequently allows for real-time allocations.

In this paper, we analyzed the Proportional Share mechanism and the discriminatory pay-as-bid mechanism by Sanghavi and Hajek (2004), two prominent proxies of such pay-as-bid mechanisms, thus extending previous work by Sanghavi and Hajek (2004) and Stößler et al. (2008). In order to study the performance of these mechanisms, e.g. as regards efficiency and provider’s revenue, we modeled the strategic behavior of rational users using the prominent solution concept of Nash equilibria. The result of this paper is twofold. We show that for three users, Nash equilibria can only be obtained under strong assumptions. Moreover, in case we have a Nash equilibrium which holds for both mechanisms, Proportional Share achieves higher efficiency than the Sanghavi-Hajek mechanism, thus reversing the results for two users.

This shows the limitations of the analytic approach in more complex and realistic settings, and strengthens the case for alternative techniques. In the future, we will analyze the mechanisms by means of numerical experiments. This will allow us to model agents with learning capabilities. It will also be interesting to compare these simple pay-as-bid mechanisms to combinatorial approaches to Grid resource allocation, which have been proposed by several authors (cf. Neumann et al. 2008 for an overview). Moreover, our analysis is set within a real-world scenario, the Biz2Grid project which we introduced earlier. We thus plan a prototypical implementation of a market-based scheduler, which will serve as proof of our concept.

References


7 APPENDIX

Proof of Proposition 6. Applying the conditions given in Proposition 3 leads us to the following system of equations, where for the sake of simplicity we write $w_i$ instead of $w_i^{NE}$ for the components of the strategic chosen Nash equilibrium payment vector:

$$w_1 = \sqrt{v_1 (w_2 + w_3)} - (w_2 + w_3), \quad w_2 = \sqrt{v_2 (w_1 + w_3)} - (w_1 + w_3) \quad \text{and}$$

$$w_3 = \sqrt{v_3 (w_1 + w_2)} - (w_1 + w_2).$$

Solving this system leads us to a unique value of $w_3$ and two possible options of $w_1$ and $w_2$:

$$w_3 = \frac{2v_1 v_2 v_3 (v_1 v_3 + v_2 v_3 - v_1 v_2)}{(v_1 v_2 + v_1 v_3 + v_2 v_3)^2}, \quad w_{2\pm} = \frac{v_2 (v_1 v_2 + \sqrt{v_1 v_2 (v_1 v_2 + 4w_3 (v_1 + v_2))}) - 2v_1 w_3 (v_1 + v_2)}{2 (v_1 + v_2)^2} \quad \text{and}$$

$$w_{1\pm} = \sqrt{v_1 (w_{2\pm} + w_3)} - (w_{2\pm} + w_3).$$

Clearly we have multiple Nash equilibria. In order to achieve a unique equilibrium we have to imply one condition on the valuation of user 3, i.e. $v_1 v_2 (v_1 v_2 + 4w_3 (v_1 + v_2)) = 0$ yields the condition $v_3 = \frac{1}{3} v_1 v_2$. Furthermore, we need to check whether or not all payments $w_i$ made by the users are non-negative, which is not given by the above equations. A simple computation shows us, that $w_3 \geq 0$ is equivalent to $v_3 \geq \frac{v_1 v_2}{v_1 + v_2}$, $w_2 \geq 0$ is equivalent to $v_2 \geq \sqrt{v_1} w_3$ and $w_1 \geq 0$ is equivalent to $v_1 \geq w_2 + w_3$. Of course all of these conditions may be expressed by the valuations $v_i$, $i = 1, ..., 3$, but results in much more complex inequations. This leads us directly to the conclusion that there is no unique Nash equilibrium for three users, since the realization of $v_3 = \frac{1}{3} v_1 v_2$ yields to the negative bid $w_3 < 0$, which is not feasible. Finally, it is straightforward, that for any positive value of $w_3$ we obtain $w_{2\pm} \leq 0$, hence we may discard the second solution and focus on $w_{14}$ and $w_{2+}$ though we may not speak of a unique equilibrium but rather of a feasible equilibrium.
Proof of Proposition 8. In order to be able to apply the mechanism, we have to determine the user who submitted the highest bid. Without loss of generalization we may assume an ascending order for the payment vector, thus being $w_3 = w_{\text{max}}$ the highest bid. Similar to the discussion of the Proportional Share allocation mechanism we have the following equation system

$$w_1 = \frac{3w_3(v_2 - 2w_3)}{v_2}, \quad w_2 = \frac{3w_3(v_1 - 2w_3)}{v_1}$$

and

$$v_3 = \frac{3w_3}{3w_3(w_1 + w_2) - 4w_1w_2},$$

which leads to two possible solutions of $w_3$:

$$w_3\pm = \frac{1}{48v_3}(9v_3(v_1 + v_2) - v_1v_2 \pm \sqrt{81v_3^2(v_1^2 + v_2^2 - \frac{9}{14}v_1v_2) - 18v_1v_2v_3(v_1 + v_2) + v_1^2v_2^2}).$$

It is obvious that in order to obtain a unique Nash equilibrium, the sum contained in the root has to be zero, hence we need to demand a further condition on one of the valuations. Choosing that valuation to be $v_3$, it has to be that

$$81v_3^2(v_1^2 + v_2^2 - \frac{9}{14}v_1v_2) - 18v_1v_2v_3(v_1 + v_2) + v_1^2v_2^2 = 0$$

is valid for all valuations $v_1$ and $v_2$ resulting in a direct dependency of $v_3$ from the other two valuations. We receive two possible solutions for $v_{3\pm} = \frac{1}{3}v_1v_2(3(v_1 + v_2) \pm 4\sqrt{2v_1v_2^2})$ and therefore forcing $v_3$ to be on a cone-like structure.

As it is seen, the second solution gives us, even for very high valuations of the remaining two users, a relatively low value of the third valuation, though we assumed the third user to be the highest bidder. Since this scenario appears to be an extremely unlikely case, we may discard the second solution and focus on the first. Therefore we have a unique solution for $v_3$ and thus for $w_3$, both depending only on the valuations of the first two, lower-bidding users. The question is, if that is the one and only restriction even though it is a very tough condition and may rarely be achieved in real-world applications. Unfortunately the answer is that we need even more restrictions, this time on the realizations of the valuations of the lower-bidding users to ensure $0 \leq w_i \leq w_{\text{max}}$ for $i = 1, 2$, since otherwise we would have a contradiction to our primary assumptions. A simple computation shows us that the first inequality holds, if $2v_1 \geq v_2$ and $2v_2 \geq v_1$, while the second inequality holds if $\frac{8}{9}v_1 \leq v_2 \leq \frac{9}{8}v_1$. Since we need both inequalities to hold and the second condition already implies the first one, we may focus thereon, which leaves only a narrow margin for the possible choices of $v_1$ and $v_2$. At this point we have to remark, that for any given valuations of the lower-bidding users, $v_3$ will take on a value which is less than any of this valuations, though we already have chosen the highest possible value according to our calculations. The maximum for $v_3$ will be reached exactly if $v_2 = \frac{9}{8}v_1$ or vice versa, since the lower-bidding users are symmetric, i.e. it does not matter who bids or valuates the good more than the other one, as long as they do not bid higher than the third user.

Using the results above, we may compute a solution under the given conditions on the valuations $v_i$, $i = 1, ..., 3$. Since the components of the resulting payment vector $w$ are relatively complex, we just state here the final result of the utility functions: For any given valuations $v_i$, $i = 1, 2$, of the lower-bidding users their utility functions equal zero, i.e. $u_i = 0$ for $i = 1, 2$ while the utility of the higher-bidding user is besides being relatively low, non-negative for $\frac{8}{9}v_1 \leq v_2 \leq \frac{9}{8}v_1$.

Finally, omitting the restriction of a unique Nash equilibrium we may analyze the case of two equilibria. As is seen in Figure 1 a unique equilibrium is achieved for $v_3$ being on the cone-like structure. For $v_3$ not holding that requirement we can distinguish between two cases: First, $v_3$ being inside the structure and second, $v_3$ taking a value outside the cone. For the first case, it is easily seen, that we have no Nash equilibrium at all, because we have to deal with a square root of a non-positive
value. In the second case, we may discard the possibility of \( v_3 \) taking a value below the structure because we assumed the third user to be the highest bidder and thus this choice seems odd. So we proceed with the following condition

\[
\frac{1}{3} v_1 v_2 (3(v_1 + v_2) + 4\sqrt{2}\sqrt{v_1 v_2})
\]

\[
\geq \frac{9}{9v_1^2 + v_2^2} - 14v_1 v_2
\]

which gives rise to two Nash equilibria depending on the choice of \( w_{3\pm} \). To ensure a feasible solution, i.e. \( 0 \leq w_i \leq w_{max} \) for \( i = 1, 2 \), a subsequent analysis shows that it has to hold that \( 2w_{3\pm} \leq w_i \leq 3w_{3\pm} \) for \( i = 1, 2 \) resulting in a zero utility for these users as well. At this point we may remark, that \( w_{3\pm} \) is ascending and limited for all choices of \( v_3 \) and converges so that we have

\[
\lim_{v_3 \to \infty} w_{3\pm} = \frac{3(v_1 + v_2) \pm \sqrt{9(v_1^2 + v_2^2) - 14v_1 v_2}}{16}.
\]

We see, that the above condition on the valuations of the lower-bidding users is based upon the value of \( w_{3\pm} \). Therefore, instead of being able to check the conditions just on the given valuations of the three users, we have to compute \( w_{3\pm} \) first, followed by a decision, which solution of \( w_3 \) to use. Having in mind the conditions on \( v_i \), \( i = 1, 2 \), gives rise to the consideration choosing strategically \( w_{3+} \) and discard \( w_{3-} \) though one cannot eliminate the possibility of a feasible equilibrium with the last realization.

*Proof of Corollary 10.* We just have to show that the conditions on the valuations of the users which are being implied by the unique Nash equilibrium satisfy also the conditions necessary for a feasible equilibrium for the Proportional Share mechanism. This is straightforward. We may fix without loss of generalization the valuation of the first user \( v_1 \). Then \( v_2 \) has to satisfy \( \frac{8}{9} v_1 \leq v_2 \leq \frac{9}{8} v_1 \). A simple analysis shows us that \( v_3 \), depending on the other valuations, is ascending and thus we choose the lowest possible value for \( v_3 \) with \( v_2 = \frac{8}{9} v_1 \). This yields \( v_3 = v_2 \) and thus the first condition \( v_3 \geq \frac{v_1 v_2}{v_1 + v_2} \) for Proportional Share holds. To check the second condition \( v_2 \geq \sqrt{v_1 w_3} \) we confirm that \( w_3 \) is ascending as well and with \( v_1 \) fixed and the above choice of \( v_2 = \frac{9}{8} v_1 \) this results in \( w_3 = \frac{3}{8} v_1 \) and the second condition holds as well. Since \( w_2 \) depending on \( w_3 \) is also ascending we continue with the above choice of \( w_3 \) and have \( w_2 = \frac{9}{32} v_1 \) which makes the third condition hold for all possible realizations of the valuations of the users implied by the conditions of the unique Nash equilibrium by the Sanghavi-Hajek mechanism thus concluding our proof.
ALGORITHMIC TRADING ENGINES VERSUS HUMAN TRADERS - DO THEY BEHAVE DIFFERENT IN SECURITIES MARKETS?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0026.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Behaviour change, Empirical study, Marketplaces, Double auction</td>
</tr>
</tbody>
</table>
ALGORITHMIC TRADING ENGINES
VERSUS HUMAN TRADERS
– DO THEY BEHAVE DIFFERENTLY IN SECURITIES MARKETS?

Gsell, Markus, Goethe-University Frankfurt, Grüneburgplatz 1, RuW Box 69, D-60629 Frankfurt am Main, Germany, gsell@wiwi.uni-frankfurt.de
Gomber, Peter, Goethe-University Frankfurt, E-Finance Lab, RuW Box 69, D-60629 Frankfurt am Main, Germany, gomber@wiwi.uni-frankfurt.de

Abstract

After exchanges and alternative trading venues have introduced electronic execution mechanisms worldwide, the focus of the securities trading industry shifted to the use of fully electronic trading engines by banks, brokers and their institutional customers. These Algorithmic Trading engines enable order submissions without human intervention based on quantitative models applying historical and real-time market data. Although there is a widespread discussion on the pros and cons of Algorithmic Trading and on its impact on market volatility and market quality, little is known on how algorithms actually place their orders in the market and whether and in which respect this differs from other order submissions. Based on a dataset that – for the first time – includes a specific flag to enable the identification of orders submitted by Algorithmic Trading engines, the paper investigates the extent of Algorithmic Trading activity and specifically their order placement strategies in comparison to human traders in the Xetra trading system. It is shown that Algorithmic Trading has become a relevant part of overall market activity and that Algorithmic Trading engines fundamentally differ from human traders in their order submission, modification and deletion behavior as they exploit real-time market data and latest market movements.

Keywords: Electronic Markets, Algorithmic Trading, Order Submission, Securities Trading.
1 INTRODUCTION

IT has triggered a significant transformation in securities trading: The electronification of market venues in Europe, i.e. exchange trading systems like Xetra (Deutsche Börse), SETS (London Stock Exchange) or NSC (Euronext Paris) took place in the late 1990s and enabled market participants (banks, brokers as well as their institutional and retail customers) to access electronic order books via remote access without the need for physical presence on an exchange floor. Now, a second electronic revolution in securities trading is taking place (Preuss 2007): market participants along the value chain started an arms race by automating their trading processes, specifically by applying Algorithmic Trading. Definitions of Algorithmic Trading conceptualize it as the general “use of computer algorithms to manage the trading process” (Hendershott et al. 2008, p.1) or as the “computerized execution of financial instruments following pre-specified rules and guidelines” (Kissell & Malamut 2006, p.12). Gomber & Gsell (2006, p.541) define it as a technology that “emulates a broker’s core competence of slicing a big order into a multiplicity of smaller orders and of timing these orders to minimize market impact via electronic means”. These algorithms determine ex ante or continuously the optimum volume of the (next) order slice and its time of submission to the market based on mathematical models and considering historical and real-time market data.

For Algorithmic Trading engines, speed of execution, availability of real-time market data and minimum latency have become key success factors as already milliseconds can make a difference. As the speed of data communication is limited by the speed of light, the best option for minimizing latency is to get physically closer to the market. Market operators therefore offer co-location services where market participants can place their trading servers adjacent to the technical infrastructure of the market itself and thus ensure low latency (a latency measurement methodology has been proposed by Budimir & Schweickert (2007)). The downside of this development for market operators is that they have to cope with increasing demands for speed and growing amounts of data and message traffic, i.e. higher investments to upgrade their infrastructure especially for peak loads. The load on market operators’ systems is steadily increasing as “Algorithmic Trading is the fastest growing source of order flow” (Preuss 2007, p.154). However, there is only little academic work on how Algorithmic Trading engines schedule their trading strategies and adapt their behavior to current market movements and whether or to which extent it is different to the trading behavior of (human) traders.

Based on a unique dataset that encompasses all order book activity in the 30 most liquid shares traded on Xetra, the electronic trading system of the Frankfurt Stock Exchange, this research aims at demonstrating the manifest differences in the order submission and deletion behavior of Algorithmic Trading engines versus other order flow submitters. This is facilitated as the dataset enables to distinguish between orders submitted by Algorithmic Trading systems and orders submitted by humans. In particular, differences in the order submission strategies and order aggressiveness as well as concerning update/deletion strategies are disclosed. In the following section 2, related work on Algorithmic Trading is discussed. Section 3 provides a brief overview of the application domain, describes the available dataset as well as the applied methodology. Section 4 presents the results obtained while the final section concludes and gives an outlook on future research in this field.

2 RELATED WORK

Algorithmic Trading systems typically aim at achieving or beating a specified benchmark with their executions and may be distinguished by their underlying benchmark, their aggressiveness or trading style as well as their adaptation behavior (Kissell & Malamut 2006). The volume-weighted average price (VWAP), which is calculated as the ratio of the value traded and the volume traded within a specified time horizon, commonly serves as a benchmark for (automated) trading (Domowitz & Yegerman 2005). The universe of possible strategies has been narrowed down to the efficient frontier of optimal trading strategies (Almgren & Chriss 2000, Almgren & Lorenz 2007). There is evidence
that strategies that are adaptive to market developments, i.e. that can vary their aggressiveness, are superior to static strategies (Almgren & Chriss 2000). Furthermore, Algorithmic Trading systems must avoid to be detected as this leaked information could be exploited by other market participants (Brunnermeier & Pedersen 2005). Research on aggressiveness of orders in general is given by Ranaldo (2004) who investigated how (human) traders adapt their behavior to changes in the order book. Empirical research found the execution quality of algorithms to be inferior to executions handled by a broker. Nevertheless, this underperformance can be overcompensated by the fact that algorithms are offered at lower fees than human order handling (Domowitz & Yegerman 2005), as no (expensive) human traders are involved. Due to the increased cost consciousness among market participants, algorithms have become an attractive alternative.

Little research is available that deals with the behavior of Algorithmic Trading systems and their impact on the market itself. Hendershot et al. (2008) showed that Algorithmic Trading has a positive impact on liquidity, while Gsell (2008) found evidence that it has potential to lower market volatility. Datasets similar to the one used for the research at hand have been analyzed by Prix et al. (2007) and (2008). The former analyzes the lifetime distribution of cancelled orders and finds systematic patterns, while the latter investigates cancellation and re-insertion structures in the Xetra order flow. The main distinction to those datasets and the novelty of the dataset used for this research is the information whether an order event was submitted by an Algorithmic Trading system or a human trader.

3 APPLICATION DOMAIN AND METHODOLOGY

3.1 Xetra – The electronic trading system of Deutsche Börse AG

The Frankfurt Stock Exchange, operated by Deutsche Börse AG, has launched the fully-electronic exchange trading system Xetra in 1997. It offers a range of market models that address different asset classes as well as securities with differing liquidity. For high-liquid shares Xetra offers the market model continuous trading. Continuous trading starts after an opening call auction and can be interrupted by one or several intraday call auctions. The trading day ends with a closing call auction.

For securities in the DAX index that constitutes the dataset of this research the timing is as follows: At 08:50 the call phase of the opening auction starts. In the call phase, the order book is partially open, as only the indicative auction price and volume (the volume and price at which would be executed if the call phase would end instantaneously) or best bid and offer and their volumes are disseminated. Instantly after the end of the call phase, the auction price is determined according to the principle of most executable volume (Schwartz & Francioni 2004). All auction call phases feature a predefined length plus a random end of at most 30 seconds. This means, that for all DAX securities the price of the opening auction is determined between 09:00:00 and 09:00:30. After the opening auction, continuous trading starts and for each order immediately upon entry it is checked whether the new order is executable against orders on the other side of the order book. If no execution is possible or the order was not completely executed, the order is stored in the order book according to price-time priority. During continuous trading the order book is open, i.e. the limits, the accumulated volume per limit and the number of orders per limit are displayed. An intraday call auction interrupts continuous trading at 13:00. After the auction, continuous trading resumes until it is ended by the closing auction which starts at 17:30. To ensure price continuity, continuous trading may be contingently interrupted by volatility interruptions. In case the next potential price lies outside pre-defined price ranges, a volatility interruption stops continuous trading for an additional unscheduled call auction.

To submit their trading intentions, market participants use market or limit orders. Market orders are unlimited buy or sell orders, which are to be executed at the next price determined. Limit orders are buy or sell orders, which are to be executed at their specified limit or better. The buy limit order with the highest limit and the sell limit order with the lowest limit in the order book define the spread of the market. A buy (sell) limit order that is immediately executable due to a limit equal to or higher (lower) than the current best offer (bid) is also called “marketable limit order” as its result equals the result of
a market order, i.e. immediate execution. Therefore, market orders and limit orders that trigger immediate executions are called aggressive (or submitted by an aggressor), whilst limit orders that are not immediately executable and that are positioned in the order book are called non-aggressive (or submitted by a non-aggressor). An iceberg order is a hidden order type specified by a limit, an overall volume and a peak volume. The peak is the visible part of an iceberg and is introduced into the order book according to price-time priority. In continuous trading, as soon as the peak has been completely executed and hidden volume is still available, a new peak is entered into the book. For further information on order types and the market model for equity trading see Deutsche Börse AG (2008a).

3.2 The available dataset

The blue-chip index DAX comprises the 30 largest and most actively traded companies that are listed at the Frankfurt Stock Exchange. For these securities in 2007, 98% of the order book turnover on German exchanges was executed on the Xetra trading system (Deutsche Börse AG 2007). The dataset provided by Deutsche Börse AG encompasses all Xetra order book events for the DAX securities within the week from October 8th to 12th, 2007, comprising of 9,036,638 events in total and 593,857 trades with an overall value of € 33,019,307,823.5 6 in continuous trading and call auctions. For each single order book event a code is given that specifies the type of event that occurred. 46.6% of the events are order insertions, 0.9% are modifications, 34.6% are deletions. The number of modifications is rather low because in the Xetra trading system only a reduction of the order’s volume leads to a modification event retaining the time-stamp, as it does not affect price-time priority. All other changes to order parameters are mapped to a deletion event and a subsequent submission event of a ‘new’ (the modified) order applying a new timestamp. 11% of the events are full executions and 5.9% represent partial executions. The remaining 1% consist of other primarily technical events that are not relevant for the analysis. Each event is assigned a timestamp, identifiers for the affected securities and orders, characteristics of the order and event-specific fields, e.g. a price for an execution. The given timestamps have a precision of 1/100 second.

What makes the available dataset unique is an additional flag in the data that indicates whether the submitter of the order event has been an algorithm. As an order can only be modified by the submitter, all events corresponding to the same order will have the same flag, i.e. an order is either an Algorithmic Trading order or not. Deutsche Börse AG offers a special pricing model, the so-called Automated Trading Program (ATP), which charges a lower fee for automated trading. The exchange defines ATP transactions as “all transactions that have been generated by an electronic system of either the ATP member or the ATP member’s clients, whereby the electronic system has to determine two out of the three following order parameters: price (order type and/or order limit where applicable), timing (time of order entry) and quantity (quantity of the order in number of securities)” (Deutsche Börse AG 2008b, p.1). In this program, depending on the accumulated monthly ATP volume per ATP member, a marginal rebate of up to 60% of trading fees applies. To qualify for fee rebates offered by ATP, a member’s trading process has to fulfill some prerequisites. The thereby generated orders have to be channeled directly into the Xetra system without further manual intervention. To enable the application of lower trading fees the ATP orders furthermore must be submitted using a designated ATP Trader-ID. All order events that were submitted using such an ATP Trader-ID are tagged in the available data set. As the tag is anonymous it just gives the information whether this order event is an ATP event or not. It is not possible to pin down behavior to a specific ATP Trader-ID or to directly determine whether two different orders have been submitted by the same market participant or not. The requirements set by Deutsche Börse to qualify for ATP shall ensure that the users are machines, i.e. Algorithmic Traders. However, it cannot be ensured that vice versa all machines make actually use of the ATP fee rebate, i.e. an event not flagged as an ATP event (in the following: Non-ATP) may still have been submitted by an algorithm. Though, given that the fee rebates of ATP constitute a strong truth-telling incentive for market participants, one can assume the accuracy of the ATP flag to be high.
3.3 Methodology

Based on the succession of order events in the dataset, the state of each instrument’s order book has been reconstructed event-by-event for each single point of time in the observation period according to the matching rules effective on Xetra. The executions determined based on this order book reconstruction have been validated against the actual executions reported in the dataset to ensure the reconstruction’s accuracy. The reconstruction enables to investigate order events relative to the current order book situation, i.e. in continuous trading submissions or deletions of orders may be assessed in the context of the best bid and offer limits (the spread) prevailing at the time of the event.

Chi-square tests are applied to test whether the obtained results are stochastically dependent on the ATP and Non-ATP classification, i.e. whether there are significant differences in the unveiled behaviour.

4 RESULTS – STRATEGIES OF ALGORITHMIC TRADING ENGINES VERSUS HUMAN TRADERS

The analysis targets at comparing the trading strategies of Algorithmic Trading engines and human traders in several dimensions and to answer the following research questions in this context:

1. What is the overall extent of Algorithmic Trading activity relative to human traders’ activity?

2. Does Algorithmic Trading activity, i.e. do actual orders submitted by Algorithmic Trading engines, reflect their technical ability to monitor and exploit real-time market movements and market information when algorithms execute orders aggressively?

3. Does Algorithmic Trading activity, i.e. do actual orders submitted by Algorithmic Trading engines, reflect their technical ability to monitor and exploit real-time market movements and market information when algorithms position non-aggressive orders in the order book?

The focus for the analysis of the research questions is laid on the continuous trading phases as auctions with their call phases of up to ten minutes without any executions exhibit a different trading behavior. The volumes executed in auctions would distort the time-series analysis as they concentrate large execution volumes at a single point of time. Within the dataset 13.4% of overall executed shares (volume) are executed in call auctions, representing 14.0% of the total traded value (in €). These figures include call auctions triggered by volatility interruptions. Within the observation period, five volatility interruptions occurred in different securities executing a total of 46,068 shares representing less than 0.01% of the overall value traded. The analysis of Algorithmic Trading behavior in auctions will be subject to future research. Therefore, the results presented here refer to continuous trading that represents 97% of the overall trading time.

Research question 1: What is the overall extent of Algorithmic Trading activity relative to human traders’ activity?

The activity of ATP traders can be analyzed from two perspectives. On the one hand there is the sheer amount of events (traffic on the electronic trading system) that can be analyzed comparing ATP and Non-ATP events. On the other hand, the focus may be laid on the actual executions by algorithms, i.e. the trading activity, rather than on the mere technical events. Table 1 summarizes the events (see the first two columns for ATP and Non-ATP) and actual executions (see columns 3 to 6 for ATP and Non-ATP) occurring during continuous trading per security. It shows that Algorithmic Trading constitutes a relevant part of overall system traffic and also of actual trading activity. In the observation period, the overall share of ATP events is 52.3% and for only five out of the 30 securities the share of ATP events is below 50%. Fresenius Medical Care with 69.0% has the highest rate of ATP events while Volkswagen has the lowest rate (35.8%). Concerning the actual execution events, i.e. partial or complete executions, ATP has a share of 54.7% whereby Linde has the highest rate of ATP execution events with 65.0% while Deutsche Telekom has the lowest with 42.9%.
Adidas
Allianz
BASF
Bayer
BMW
Commerzbank
Continental
Daimler
Deutsche Bank
Deutsche Börse
Deutsche Post
Deutsche Postbank
Deutsche Telekom
E.ON
Fresenius Med. Care
Henkel
Hypo Real Estate
Infineon
Linde
Lufthansa
MAN
Merck
Metro
Münchner Rück
RWE
SAP
Siemens
ThyssenKrupp
TUI
Volkswagen
Total

Instrument

Proceedings ECIS 2009
52.6%
52.3%
59.3%
57.5%
50.7%
46.1%
58.5%
50.1%
53.4%
49.1%
52.2%
51.6%
51.6%
56.1%
69.0%
52.0%
50.4%
60.0%
62.6%
65.7%
51.2%
66.8%
51.7%
48.1%
41.5%
51.4%
55.4%
64.2%
52.0%
35.8%
52.3%

Share

Table 1.

91,795
292,196
172,465
172,201
106,861
126,099
114,421
258,604
267,273
133,699
102,557
59,799
113,233
265,464
107,384
76,442
123,036
98,330
134,642
135,855
172,761
78,830
77,147
212,972
155,455
243,279
278,179
158,403
65,474
229,196
4,624,052

Events
308,233,612
1,455,589,385
889,758,389
989,237,021
444,371,422
742,364,567
401,969,476
2,151,416,279
1,594,711,428
960,562,420
581,767,692
168,314,037
1,012,429,825
1,529,680,829
158,454,475
185,661,502
361,043,311
469,101,313
333,153,037
384,794,459
619,494,708
307,386,567
221,789,717
881,141,537
997,175,684
2,049,684,431
1,786,939,922
442,257,844
180,619,398
1,779,973,952
24,389,078,240

Value (€)
41.9%
48.9%
50.4%
42.8%
43.5%
32.9%
39.2%
46.0%
42.4%
43.2%
33.8%
41.6%
33.2%
56.0%
53.0%
35.8%
41.5%
40.5%
55.6%
51.5%
38.1%
43.9%
40.6%
50.9%
48.5%
37.9%
50.4%
46.9%
36.4%
36.5%
43.0%

Share
82,882
265,999
118,232
127,413
103,943
147,673
81,009
258,018
233,233
138,387
93,961
56,156
106,378
207,421
48,137
70,543
120,956
65,612
80,345
71,015
164,805
39,156
72,191
230,222
218,999
230,183
224,035
88,513
60,343
411,168
4,216,928

Events
47.4%
47.7%
40.7%
42.5%
49.3%
53.9%
41.5%
49.9%
46.6%
50.9%
47.8%
48.4%
48.4%
43.9%
31.0%
48.0%
49.6%
40.0%
37.4%
34.3%
48.8%
33.2%
48.3%
51.9%
58.5%
48.6%
44.6%
35.8%
48.0%
64.2%
47.7%

Share

ATP and Non-ATP activity during continuous trading

ATP
Execution
Share
Events
15,899 54.6%
43,208 59.6%
30,156 62.2%
37,073 56.4%
19,722 53.4%
27,922 45.9%
18,541 51.8%
52,263 55.8%
41,985 52.5%
31,833 55.5%
18,433 45.6%
10,870 53.6%
19,752 42.9%
40,173 62.4%
12,486 58.5%
12,235 46.4%
19,458 51.5%
16,354 50.5%
18,179 65.0%
20,766 61.6%
25,917 52.6%
15,212 57.8%
11,448 54.0%
31,270 62.3%
30,675 58.6%
57,515 48.3%
43,853 58.9%
25,090 62.2%
10,934 50.2%
54,846 53.1%
814,068 54.7%

Non-ATP
Execution
Share
Events
13,197 45.4%
29282 40.4%
18,336 37.8%
28,698 43.6%
17,181 46.6%
32,967 54.1%
17,267 48.2%
41,427 44.2%
37,945 47.5%
25,529 44.5%
21,997 54.4%
9,420 46.4%
26,316 57.1%
24,179 37.6%
8,846 41.5%
14,119 53.6%
18,359 48.5%
16,014 49.5%
9,807 35.0%
12,933 38.4%
23,362 47.4%
11,109 42.2%
9,755 46.0%
18889 37.7%
21,650 41.4%
61,624 51.7%
30,554 41.1%
15,228 37.8%
10849 49.8%
48,460 46.9%
675,299 45.3%
427,333,950
1,520,815,223
874,243,332
1,322,869,917
576,016,375
1,516,002,668
624,238,212
2,525,344,806
2,164,330,852
1,262,130,542
1,141,429,062
236,203,732
2,032,482,945
1,201,021,115
140,433,106
332,312,755
508,504,512
687,986,171
265,955,747
361,770,805
1,006,148,736
393,078,246
324,935,561
848,785,177
1,060,193,847
3,354,701,075
1,757,568,726
500,784,030
316,040,687
3,092,226,374
32,375,888,285

Value (€)

58.1%
51.1%
49.6%
57.2%
56.5%
67.1%
60.8%
54.0%
57.6%
56.8%
66.2%
58.4%
66.8%
44.0%
47.0%
64.2%
58.5%
59.5%
44.4%
48.5%
61.9%
56.1%
59.4%
49.1%
51.5%
62.1%
49.6%
53.1%
63.6%
63.5%
57.0%

Share

17th European Conference on Information Systems
Page 7 of 13


The value associated with ATP trading is 43.0% on average for all securities. As execution events are considered in Table 1, the traded value is double-counted; once for the buyer and once for the seller. 52.7% of the orders entered during continuous trading are ATP orders. ATP and Non-ATP submission exhibit a similar share of aggressive orders (ATP: 14.6%; Non-ATP: 11.2%) (order data not shown additionally in Table 1 to assure readability).

In the following a deeper analysis distinguishing aggressor and non-aggressor orders is performed: If aggressive executions are considered (Table 2), in total 56% of all executions are triggered by an ATP order as the aggressor. Linde has the highest share of ATP aggressor executions as for 68.4% an ATP trader was the aggressor. Further E.ON catches attention, as about two out of three executions were triggered by ATP orders, which sum up to 63.6% of the total value traded. For Deutsche Post only 30.5% of the executed value has been triggered by aggressive ATP orders. Please note that for the execution perspective of aggressors, the traded value is only single-counted.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>ATP Aggressor</th>
<th>Non-ATP Aggressor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#Exec.</td>
<td>Share</td>
</tr>
<tr>
<td>Adidas</td>
<td>5,871</td>
<td>51.8%</td>
</tr>
<tr>
<td>Allianz</td>
<td>17,939</td>
<td>58.3%</td>
</tr>
<tr>
<td>BASF</td>
<td>12,284</td>
<td>63.6%</td>
</tr>
<tr>
<td>Bayer</td>
<td>14,661</td>
<td>59.1%</td>
</tr>
<tr>
<td>BMW</td>
<td>7,759</td>
<td>54.4%</td>
</tr>
<tr>
<td>Commerzbank</td>
<td>10,674</td>
<td>47.3%</td>
</tr>
<tr>
<td>Continental</td>
<td>7,428</td>
<td>50.9%</td>
</tr>
<tr>
<td>Daimler</td>
<td>22,544</td>
<td>58.5%</td>
</tr>
<tr>
<td>Deutsche Bank</td>
<td>17,440</td>
<td>54.1%</td>
</tr>
<tr>
<td>Deutsche Börse</td>
<td>13,079</td>
<td>55.1%</td>
</tr>
<tr>
<td>Deutsche Post</td>
<td>6,329</td>
<td>45.1%</td>
</tr>
<tr>
<td>Deutsche Postbank</td>
<td>4,343</td>
<td>54.5%</td>
</tr>
<tr>
<td>Deutsche Telekom</td>
<td>8,217</td>
<td>47.7%</td>
</tr>
<tr>
<td>E.ON</td>
<td>17,075</td>
<td>66.0%</td>
</tr>
<tr>
<td>Fresenius Med. Care</td>
<td>4,930</td>
<td>57.1%</td>
</tr>
<tr>
<td>Henkel</td>
<td>4,603</td>
<td>44.7%</td>
</tr>
<tr>
<td>Hypo Real Estate</td>
<td>7,587</td>
<td>50.5%</td>
</tr>
<tr>
<td>Infineon</td>
<td>6,232</td>
<td>53.1%</td>
</tr>
<tr>
<td>Linde</td>
<td>7,916</td>
<td>68.4%</td>
</tr>
<tr>
<td>Luftansa</td>
<td>8,285</td>
<td>65.6%</td>
</tr>
<tr>
<td>MAN</td>
<td>10,385</td>
<td>50.6%</td>
</tr>
<tr>
<td>Merck</td>
<td>6,132</td>
<td>58.0%</td>
</tr>
<tr>
<td>Metro</td>
<td>4,890</td>
<td>57.5%</td>
</tr>
<tr>
<td>Münchner Rück</td>
<td>13,263</td>
<td>64.5%</td>
</tr>
<tr>
<td>RWE</td>
<td>13,280</td>
<td>63.4%</td>
</tr>
<tr>
<td>SAP</td>
<td>23,196</td>
<td>50.3%</td>
</tr>
<tr>
<td>Siemens</td>
<td>18,266</td>
<td>59.3%</td>
</tr>
<tr>
<td>ThyssenKrupp</td>
<td>9,446</td>
<td>59.7%</td>
</tr>
<tr>
<td>TUI</td>
<td>4,101</td>
<td>48.9%</td>
</tr>
<tr>
<td>Volkswagen</td>
<td>24,282</td>
<td>55.1%</td>
</tr>
<tr>
<td>Total</td>
<td>332,437</td>
<td>56.0%</td>
</tr>
</tbody>
</table>

Table 2. Number of executions and executed value for ATP- and Non-ATP aggressors in continuous trading

Table 2 further discloses that though across all securities 56.0% of all executions are triggered by aggressive ATP orders, they represent only 46.2% of the totally traded value. The average value per
executed order for ATP aggressor executions is 39,444 € and the average value per executed order for Non-ATP aggressor executions is 58,512 €. There are two possible explanations:

1) **ATP users have more partial executions**

A possible explanation for the fact that more executions result in less executed volume would be that aggressive Algorithmic Trading orders have more partial executions boosting their total number of executions. If the algorithms submit orders that hit orders at several price levels, this would result in several executions and the value per execution would be lowered. However, this argument does not hold true. This can be checked by measuring the number of different timestamps for the respective executions as multiple executions at the same timestamp represent partial executions. As there are 332,437 ATP aggressor executions at only 278,374 different timestamps, 19.4% of the aggressor orders seem to cause executions at several price levels. But for the Non-ATP aggressors there are 260,965 executions at 209,919 distinct timestamps indicating that 24.3% of Non-ATP aggressor orders cause more than one execution.

2) **Algorithms submit more but smaller orders**

Assuming that the algorithms can monitor changes in the order book and react in real-time and given that they still get what they saw when their order is arriving at the market, algorithms look for advantageous limits in the order book and snap at the chance and execute the best bid or offer. As the top of the book most often is thin – as the most volume is just behind the best bid and offer – this results in more but smaller executions. Furthermore the smaller executions lead to less market impact than larger executions that would potentially hit more than one price level.

For non-aggressor orders, Table 3 presents evidence that ATP non-aggressor orders (i.e. limit orders that are not immediately executable) are also smaller than their Non-ATP counterparts as it depicts the average order volumes and average order values (order volume times order limit). The surplus of Non-ATP orders’ average value over ATP orders’ average value is 143.5% for all securities. The values for the individual instruments range from 40.6% (Linde) to 455.7% (Volkswagen). Please note that iceberg orders and their peaks have not been considered for the calculation of the averages.

<table>
<thead>
<tr>
<th>ATP Non-Aggressor Orders</th>
<th>Non-ATP Non-Aggressor Orders</th>
<th>Avg. Value Surplus</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>Share</td>
<td>Avg. Volume</td>
</tr>
<tr>
<td>Adidas</td>
<td>37,007</td>
<td>53.9%</td>
</tr>
<tr>
<td>Allianz</td>
<td>119,618</td>
<td>51.2%</td>
</tr>
<tr>
<td>BASF</td>
<td>68,921</td>
<td>59.3%</td>
</tr>
<tr>
<td>Bayer</td>
<td>65,683</td>
<td>59.4%</td>
</tr>
<tr>
<td>BMW</td>
<td>41,761</td>
<td>51.5%</td>
</tr>
<tr>
<td>Commerzbank</td>
<td>47,947</td>
<td>48.1%</td>
</tr>
<tr>
<td>Continental</td>
<td>46,749</td>
<td>62.9%</td>
</tr>
<tr>
<td>Daimler</td>
<td>48,249</td>
<td>48.1%</td>
</tr>
<tr>
<td>Deutsche Bank</td>
<td>108,964</td>
<td>54.1%</td>
</tr>
<tr>
<td>Deutsche Börse</td>
<td>97,228</td>
<td>48.6%</td>
</tr>
<tr>
<td>Deutsche Post</td>
<td>23,812</td>
<td>60.2%</td>
</tr>
<tr>
<td>Deutsche Postbank</td>
<td>41,370</td>
<td>56.3%</td>
</tr>
<tr>
<td>Deutsche Telekom</td>
<td>43,692</td>
<td>54.8%</td>
</tr>
<tr>
<td>EON</td>
<td>108,008</td>
<td>54.7%</td>
</tr>
<tr>
<td>Fresenius Med. Care</td>
<td>45,735</td>
<td>71.1%</td>
</tr>
<tr>
<td>Henkel</td>
<td>31,483</td>
<td>55.5%</td>
</tr>
<tr>
<td>Hypo Real Estate</td>
<td>50,448</td>
<td>56.7%</td>
</tr>
<tr>
<td>Infineon</td>
<td>40,143</td>
<td>66.5%</td>
</tr>
<tr>
<td>Linde</td>
<td>55,312</td>
<td>68.2%</td>
</tr>
<tr>
<td>Lufthansa</td>
<td>56,682</td>
<td>62.7%</td>
</tr>
<tr>
<td>MAN</td>
<td>71,643</td>
<td>52.0%</td>
</tr>
</tbody>
</table>
Concerning research question 1, the data reveals that Algorithmic Trading is a relevant part of technical events, actual executions as well as order submissions. Algorithms tend to use smaller order volumes both for aggressive as well as for non-aggressive orders than (human) Non-ATP counterparts.

Research question 2: Does Algorithmic Trading activity, i.e. do actual orders submitted by Algorithmic Trading engines, reflect their technical ability to monitor and exploit real-time market movements and market information when algorithms execute orders aggressively?

Research question 2 relates to the (aggressive) execution behavior of algorithms and is addressed in two dimensions: 2a) concerning the usage of order types by algorithms versus human traders and 2b) concerning the submitted limits in case of aggressive limit orders by algorithms versus human traders.

2a) If ATP traders would be more aggressive one might assume that they will utilize market orders to a larger extent than Non-ATP traders. As Table 4 points out, this is not the case, as although there is a similar number of ATP and Non-ATP orders involved in continuous trading, only 6.2% of the market orders have been submitted by ATP users. A straightforward chi-square test shows that the null-hypothesis of equal likelihood for ATP and Non-ATP participants to either utilize limited orders (limit and iceberg orders) or market orders can be rejected at a p-value of 0.01, which reveals a highly significant difference. The vast majority of ATP submitted orders are limit orders, as an aggressive strategy can be implemented with limit orders and a smart setting of the limits as well. Such a strategy is eased by speed and low latency to monitor market movements in real-time and to react with minimum delay – a prerequisite that can be matched by machines.

2b) In the following the focus is laid on how algorithms set order limits in relation to the prevailing order book situation when implementing an aggressive strategy. Although ATP and Non-ATP exhibit a similar share of aggressive orders (ATP: 14.6%, Non-ATP: 11.2%; see research question 1), the applied limits relative to current best bids and offers differ clearly (Table 5). 67.8% of all order submissions that exactly match the best bid or offer are ATP orders. Nearly two thirds of the other aggressive submissions are Non-ATP orders. For ATP orders, even 85.1% of the aggressive orders are limited exactly to the best available limit in the order book. A chi-square test shows that the null-hypothesis of equal likelihood of ATP and Non-ATP orders to submit exact limit matches can be rejected at a p-value of 0.01, i.e. reveals a highly significant difference.
Further, 17.7% of aggressive ATP orders are also exactly matching the volume available at the best limit (Non-ATP: 7.9%). Out of all submissions that exactly match the opposite side’s limit and volume 76.7% are ATP orders. For high-liquid securities, such as Siemens or E.ON, this proportion is even higher (92.1% respectively 91.9%). Referring to research question 2, these results indicate that ATP orders’ limits and volumes are based on a real-time monitoring of the market and are set based on latest market movements.

Research question 3: Does Algorithmic Trading activity, i.e. do actual orders submitted by Algorithmic Trading engines, reflect their technical ability to monitor and exploit real-time market movements and market information when algorithms position non-aggressive orders in the order book?

Research question 3 relates to the (non-aggressive) submission behavior and can be addressed in two dimensions: 3a) concerning the positioning of non-aggressive orders by algorithms versus human traders relative to the current best bids or best offers and 3b) concerning the adaptation of limits by algorithms versus human traders in case of changing best bids or best offers.

3a) Table 6 points out the different positioning of non-aggressive ATP and Non-ATP orders. Of the non-aggressive orders that improve the spread, 75.9% are ATP orders, while of the orders that do not affect the spread 62.5% are Non-ATP. A chi-square test (null-hypothesis: equal likelihood of ATP and Non-ATP orders to improve the spread) can be rejected at a p-value of 0.01, i.e. again reveals a highly significant difference among ATP and Non-ATP orders. The different order positioning behavior can also be seen from the weighted-average absolute variation in cents by which orders narrow the spread. ATP orders improve the best limit on average by 1.38 cents while Non-ATP orders improve it by 1.95 cents, i.e. algorithms are able to position orders at the top of the book with a lower concession in terms of price improvement.

3b) As shown above, ATP market participants limit the majority of their non-aggressive order in a way to be at the spread by either adding volume to the existing spread limit or setting a better limit. As these orders are positioned at the top of the book, i.e. they have a high likelihood of execution, it is of interest to investigate their further lifetime. The following table 7 depicts what happens to orders that are part of the spread when their lifetime ends. The absolute figures reveal that there are by far more ATP orders that end their lifetime being at the spread. As for table 7 the distribution of the termination reasons is in the focus, the percentages are calculated in relation to all termination reasons at the spread for each ATP and Non-ATP. About two thirds of Non-ATP orders end their lifetime at the spread by getting executed, while about one third gets deleted. For ATP orders these ratios are nearly vice versa. 63.2% of the ATP orders that end their lifetime being part of the spread are deleted. Again,
a chi-square test (null-hypothesis: equal likelihood of ATP and Non-ATP orders to be terminated by execution; p-value of 0.01) reveals a highly significant difference between ATP and Non-ATP orders.

The third and sixth column (table 7) list the average time in milliseconds, that the orders were continuously part of the spread before termination. The average survival times for orders terminated by execution are similar. This meets the expectation, as for executions the survival time is determined by other (aggressive) orders and therefore can not be influenced by the order positioned at the spread. However, for deleted Non-ATP orders the survival time is nearly twice the one for deleted ATP orders. 63.2% of ATP orders are deleted on average 6.519 seconds after becoming part of the spread. At a first glance it seems as if ATP users initially submit orders at the spread and then get cold feet and delete their orders to avoid execution. But what seems to be a deletion is actually a modification, as within the Xetra market model only a reduction of order volume does not affect price-time priority while all other modifications are mapped to a deletion event and a subsequent submission event for the ‘new’ modified order. For 26.9% of the ATP orders deleted at the spread there is an ATP submission event of a new order in the same instrument at exactly the same timestamp, with the same direction (buy or sell) and exactly the same volume (Non-ATP: 23.2%). If the restrictions are relaxed (up to 1 second delay, +/-5% volume), there is a corresponding submission event for 40.2% of the ATP deletions (Non-ATP: 33.4%) out of which 115,331 again improve the spread (Non-ATP: 16,822). This indicates that ATP traders want their orders to be at the top of the book. Therefore, they emulate pegging orders at their front end (an order type where the limit tracks the best bid or offer and moves with the market) as this order type – contrary to other markets – is not provided by the Xetra back-end.

<table>
<thead>
<tr>
<th></th>
<th>ATP</th>
<th></th>
<th>Non-ATP</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Distribution</td>
<td>Avg. Survival time (ms)</td>
<td>Events</td>
</tr>
<tr>
<td>Execution</td>
<td>293,142</td>
<td>36.8%</td>
<td>17,927</td>
<td>184,009</td>
</tr>
<tr>
<td>Deletion</td>
<td>503,650</td>
<td>63.2%</td>
<td>6,519</td>
<td>99,674</td>
</tr>
<tr>
<td>Total</td>
<td>796,792</td>
<td>100.0%</td>
<td>10,716</td>
<td>283,410</td>
</tr>
</tbody>
</table>

Table 7. Survival time and termination reason for orders at the spread

To sum up the data analyzed concerning research question 3, the results indicate that ATP traders are more aware of the current spread, as they more often reflect the current best bid or offer when limiting their orders, and that their orders about three times more often narrow the current spread. Furthermore, they control their orders relative to the current market situation and delete and reinsert their orders based on changes in the current spread more extensively than Non-ATP traders do.

5 CONCLUSION

The detailed implementations of Algorithmic Trading systems are not published as these constitute important intellectual property rights of investment firms and are a key component of their business models both for proprietary trading and when providing algorithms to customers (regularly as a black box). Therefore, only little is known about how Algorithmic Trading engines schedule their trading and adapt to current market movements. Based on a unique dataset from a market operator that includes a tag enabling to distinguish Algorithmic Trading engines and human traders, the order submission behavior of both groups has been compared. An implication for theory is the result that the submission and deletion behavior of Algorithmic Trading systems (statistically) significantly differs from other market participants’. Evidence has been presented that they submit orders that are noticeably smaller. Additionally they show the ability to monitor their orders and modify them to be at the top of the book. Applying chi-square tests shows that Algorithmic Trading behavior is fundamentally different to human trading concerning the use of order types, the positioning of limits in case of executions and submissions as well as their modification/deletion behavior. These results let us conclude that Algorithmic Trading systems capitalize on their advantageous ability to process high-
speed data feeds and react instantaneously to market movements by submitting corresponding orders or modifying existing ones.

Future research based on the dataset will both investigate order submission strategies in auctions and the contribution of Algorithmic Trading engines to overall market liquidity. This contributes not only to the understanding of the algorithmic implementations but also has theoretical implications for market design and market surveillance issues. Furthermore, the understanding of algorithmic behavior has practical implications, as it enables to identify potential functional or technical bottlenecks.

References


Proceedings ECIS 2009
Coordinating Service Composition

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0072.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Auctions, Pricing, Business Models, Economics of IS</td>
</tr>
</tbody>
</table>
COORDINATING SERVICE COMPOSITION

The fundamental paradigm shift from traditional value chains to agile service value networks implies new economic and organizational challenges. As coordination mechanisms, auctions have proven to perform quite well in situations where intangible and heterogeneous goods are traded. Nevertheless, traditional approaches in the area of multiattribute combinatorial auctions are not quite suitable to enable the trade of composite services. A flawless service execution and therefore the requester's valuation highly depends on the accurate sequence of the functional parts of the composition, meaning that in contrary to service bundles, composite services only generate value through a valid order of their components. We present an abstract model as a formalization of a service value network. The model comprehends a graph-based mechanism design to allocate multiattribute service offers within the network, to impose penalties for non-performance and to determine prices for complex services. The mechanism and the bidding language support various types of QoS attributes and their (semantic) aggregation. We analytically show that this variant is incentive compatible with respect to all dimensions of the service offer (quality and price).

Keywords: Mechanism Design, Coordination, Service Value Network, Pricing Model, Semantics.

The research was funded by means of the German Federal Ministry of Economy and Technology under the promotional reference “01MQ07012”.

1 INTRODUCTION

The paradigm shift from a product- to a service-oriented economy fosters the movement of complete industries from vertical integration to horizontal specialization. Hierarchically organized firms start to cooperate in firmly-coupled strategic networks with stable inter-organizational ties, recently exploring the benefits of moving to more loosely-coupled configurations of legally independent firms. In theory, complex products or services can be produced by a single vertically integrated company. But in this case the company is not able to focus on its core competencies, having to cover the whole spectrum of the value chain. Also, it has to burden all risks in a complex, changing and uncertain environment by itself. This is why companies tend to engage in networked value creation which allows participants to focus on their strengths. At the same time rapid innovation in the ICT sector enables promising opportunities in B2B communication which also supports the current trend. However, especially in complex and highly dynamic industries, forming value networks – especially business webs with their open structure – is more than an attractive strategic alternative. Prominent advocates of this new paradigm are (Tapscott et al., 2000, Hagel III, 1996, Zerdick et al., 2000, Steiner, 2005). As (Tapscott et al., 2000, Steiner, 2005) express it, business webs bring together mutually networked, permanently changing legally independent actors in customer centric, mostly heterarchical organizational forms in order to create (joint) value for customers. Specialized firms co-operetively contribute modules to an overall value proposition under the presence of network externalities. A prime example for such highly dynamic fields of application is the internet of services. We briefly outline the advantages of business webs related to modularization and specialization (Zerdick et al., 2000): Concentration on core competencies strengthens specialization (C1); Sharing the risk involved (C2); High level of flexibility (C3); Modularization brings potential for innovation and allows for rapid market penetration (C4); Fruitful interplay of competition and partnership (C5).

Auctions have proven to perform well under these conditions to coordinate value generation while addressing mentioned network characteristics. Nevertheless traditional approaches in the area of multiattribute combinatorial auctions are not quite suitable to enable the trade of composite services. Auctions for composite services are much more complex than simple procuring auctions, where the suppliers themselves offer a full solution to the procurer. In composite services, this is not the case, as a flawless service execution and therefore the requester's valuation highly depends on the accurate sequence of its functional parts, meaning that in contrary to service bundles, composite services only generate value through a valid order of their components.

As a coordination mechanism in networked economies we propose a multidimensional procurement auction for trading composite services. We present a graph-based model that captures the main components and characteristics of service value networks. Based on this model we introduced a mechanism design that enables allocation and pricing of service components that together form a requested complex service. The mechanism is capable of handling a wide range of aggregation operations for service attributes also supporting rich semantic approaches for dealing with complex non-functional service specifications. Due to the combinatorial restrictions imposed by the underlying graph topology and the absence of capacity constraints, the winner determination problem can be solved in polynomial time which is a crucial issue when it comes to implementing online systems. We furthermore show that the proposed mechanism is individual rational, allocation efficient and incentive compatible with respect to QoS characteristics and prices of service offers. Hence, reporting the true type regarding configuration and price is a weakly dominant strategy for all service providers.

This paper is structured as follows: The next section gives a brief overview over the literature. Section 3 illustrates the idea of on-demand service procurement in networked economies based on an integration scenario from SAP BusinessByDesign. In Section 4 we propose a multidimensional procurement auction for trading composite services based on an abstract model of a service value network. The mechanism comprehends a multiattribute bidding language (Section 4.1) and the central allocation function (Section 4.2). Section 4.3 demonstrates the semantic aggregation of service
attributes and the auction conduction by providing a numerical example. An extension regarding service level guarantees and penalties for non-performance is introduced in Section 4.4. In Section 4.5 we analytically show the providers’ bidding strategies and valuable properties of proposed mechanism design. Section 5 concludes with a summary, the practical realization of our approach and future work.

2 RELATED WORK

Recently, an enormous body of work has been done that investigates problems of coordination from a game theoretic and computer science perspective (Papadimitriou, 2001). Especially the discipline of mechanism design that focuses on the problem to coordinate self-interested participants in pursuing an overall goal (Nisan and Ronen, 2001). The authors design suitable mechanisms to standard optimization problems in the area of task scheduling and routing. In incentive compatible mechanisms agents are incentivized to choose the strategy of revealing their true type. Incentive compatible mechanisms such as the celebrated Vickrey-Clarke-Groves (VCG) mechanism are firstly introduced and extensively investigated by (Clarke, 1971, Groves, 1973, Vickrey, 1961). It is important to notice that incentive compatibility in VCG-based mechanisms may fail in repeated games (Binmore and Swierzbinski, 2000) due to the possibility to learn from past situations and adjust ones strategy in a trial-and-error process.

Most of the research has been done with respect to truth-telling of one-dimensional types. The field of designing incentive compatible mechanisms, that induce truth-telling of multidimensional properties of goods or services, still lacks deeper research. A thorough analysis and investigation in the area of multidimensional auctions and the design of optimal scoring rules has been done in (Branco, 1997, Che et al., 1993). In (Bichler and Kalagnanam, 2005) the winner determination problem in configurable multiattribute auctions is investigated from an operational research perspective without accounting for mechanism design aspects such as incentive compatibility. In (Parkes and Kalagnanam, 2002, Parkes and Kalagnanam, 2005) the authors introduce iterative multiattribute procurement auctions focusing on mechanism design issues and solving the multiattribute allocation problem. Preferences for multidimensional goods and multidimensional types in scoring auctions are extensively investigated in (Asker and Cantillon, 2008) and extended to combinatorial auctions in (Müller et al., 2008). Nevertheless their work does not consider value chains and sequences of services as well as their technically feasible interrelations in order to coordinate value generation in service networks. All of these approaches assume bundles of goods in scenarios where the sequence and order does not matter and therefore cannot be applied to composite services that only fulfil their objectives in the right sequence of execution.

Nevertheless, combinatorial auctions yield major drawbacks regarding computational feasibility that result from an NP-hard complexity. Computational feasibility implies a trade-off between optimality and valuable mechanism properties such as incentive compatibility. Several authors propose approximate solutions for incentive compatible mechanisms to overcome issues of computational complexity (Nisan and Ronen, 2007, Ronen, 2001, Ronen and Lehmann, 2005). Path auctions as a subset of combinatorial auctions reduce complexity through predefining all feasible service combinations in an underlying graph topology and are investigated in (Archer and Tardos, 2007, Feigenbaum et al., 2006, Hershberger and Suri, 2001). In their work, path auctions are utilized for pricing and routing in networks of resources such as computation or electricity. Application-related issues of auctions to optimal routing are examined by (Feldman et al., 2005, Maille and Tuffin, 2007). All of these approaches deal with the utility services layer according to the service classification in (Blau et al., 2008) and hence do not cover the problems related to complex services.
3 BUSINESS SCENARIO

To illustrate the idea of a service value network we introduce a business scenario which is actually delivered to customers as part of SAP’s BusinessByDesign\(^1\). The scenario consists of modular service components that can be provided by decentralized service providers. The integration scenario “Service Request and Order Management” (cp Figure 1) describes operational processes in a customer service based on service requests, service orders and service confirmations. From an end-to-end perspective the scenario includes the integration into related applications such as logistics planning and execution, invoicing and payment, as well as financial accounting.

![Figure 1 Business Process "Service Request and Order Management"](image)

A service value network is formed by many decentralized service providers that contribute to the achievement of an overall goal. In our scenario this goal is the flawless execution of a business scenario in order to provide defined functionality to the customer. From now on we call this overall goal a complex service. Recalling the main characteristics of service value networks there are many service providers that offer differentiated and specialized services covering various types of functionality within the network. In our scenario the functionality of each component can be modularized and therefore performed by different software-as-a-service providers as depicted in Table 1. The rapid upcoming of on-demand service providers shows the high degree of innovation and market penetration as a result of modularization (C4). Service providers offer specialized services and concentrate on their core competencies (C1). Each service provider is responsible for a certain part of the overall functionality which consequently spreads the risk of an erroneous business process over all contributing service providers (C2). Furthermore they partly grant access to their own resource supporting the realization of the overall business scenario (C5). The potential of substituting service providers on demand enables flexibility and rapid reaction on changing market requirements (C3).

<table>
<thead>
<tr>
<th>CRM</th>
<th>SCM</th>
<th>FIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rightnow (<a href="http://www.rightnow.com/">http://www.rightnow.com/</a>)</td>
<td>7Hills (<a href="http://www.7hillsbiz.com/">http://www.7hillsbiz.com/</a>)</td>
<td>Opsource (<a href="http://www.opsource.net/">http://www.opsource.net/</a>)</td>
</tr>
<tr>
<td>Oracle (<a href="http://www.oracle.com/crmondemand/">http://www.oracle.com/crmondemand/</a>)</td>
<td>Intacct (<a href="http://www.intacct.com/">http://www.intacct.com/</a>)</td>
<td></td>
</tr>
<tr>
<td>SAP (<a href="http://www.sap.com/solutions/sme/businessbydesign/">http://www.sap.com/solutions/sme/businessbydesign/</a>)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. SaaS Providers for CRM, SCM and FIN Components and their Functional Coverage

\(^1\) http://www.sap.com/solutions/sme/businessbydesign/
4 ABSTRACT MODEL & MECHANISM DESIGN

A service value network is represented by an $k$-partite, directed and acyclic graph $G = (V, E)$. Each partition $v_1, \ldots, v_k$ of the graph represents a functionality cluster that entails services that provide the same functionality (substitutes). The set of $N$ nodes $V = \{v_1, \ldots, v_N\}$ represents the set of service offers with $v$ is an arbitrary service offer. Services are offered by a set of $Q$ Service Providers $S = \{s_1, \ldots, s_Q\}$ with $s$ is an arbitrary service provider. The ownership information $\sigma: S \rightarrow V$ that reveals which service provider owns which service within the network is public knowledge. There are two designated nodes $v_s$ and $v_f$ standing for source and sink in the network. The set of $M$ edges $E = \{e_1, \ldots, e_M\}$ denotes service compatibilities and interoperabilities such that $e_j$ represents interoperability of service $j$ with service $i$ and their sequence of execution. A service configuration $A_j$ of service $j$ is fully characterized by a set of attributes $A_j = \{a_j^1, \ldots, a_j^l\}$ where $a_j^l$ is an attribute value of attribute type $l$ of service $j$’s configuration. Let furthermore $c_j(e_j, A_j)$ denote a cost function that maps service $j$’s configuration to corresponding costs such that $c: E \times A \rightarrow R$. $c_j$ denotes costs that the service provider who owns service $j$ has to bear for developing a service that is interoperable with service $i$ (development and production costs) and for performing it during execution (execution costs). Configuration and costs are private knowledge to the service provider who owns a particular service (type). If two services are not interoperable at all, they are not connected within the network. Value is created through the network by performing a sequence of services that form a connected path from source to sink. We call such created value a complex service. Let $F$ denote the set of all feasible paths from source to sink. Every $f \in F$ represents a possible instantiation of the complex service. $F_i$ represents the set of all feasible paths from source to sink without node $i$ and its incoming and outgoing edges. Let $F_i$ be the set of all feasible paths from source to sink that entail node $i$. In our model we focus on the core process of realizing an overall goal without going into process-related details such as parallel or cyclic components. We apply a business and management-oriented view addressing the question of how an overall goal can be achieved maximizing the systems welfare and to dynamically determine prices.

4.1 Bidding Language

As a formalization of information objects which are exchanged during auction conduction we introduce a bidding language based on bidding languages for products with multiple attributes as discussed in (Engel et al., 2006). Our formalization is aligned to multiattribute auction theory as presented in (Parkes and Kalagnanam, 2002, Ronen and Lehmann, 2005) and assures compliance with the WS-Agreement specification in order to enable realization in decentralized environments such as the Web.

A service requester wants to purchase a complex service $f$ which is characterized by a configuration $A_f$. The importance of certain attributes and prices of a requested complex service is idiosyncratic and depends on the preferences of the requester. The requester's preferences are represented by a utility function $U$ of the form:

$$ U_f(a, \Lambda, A, \mathcal{P}) = aS(A_f) - T_f $$

$T_f$ denotes the sum of all transfer payments the requester has to transact to service providers that contribute to the complex service such that $T_f = \sum_{e\in f} t_e$. The configuration $A_f$ of the complex service is the aggregation of all attribute values of contributing services on the path $f$ such that
\( A_j = (A_j', ..., A_j') \) with \( A_j' = \oplus_{a \in A_j} a_j' \). The aggregation of attributes values depends on their type (i.e. encryption can be aggregated by an AND operator whereas response time is aggregated by a sum operator). Different methods for aggregating service attributes are presented in Section 4.3.

The scoring rule \( S(A_j) = \left( \sum_{l=1}^{L} \lambda_l A_j^l \right) \) represents the requester’s valuation for a configuration \( A_j \) of the complex service represented by path \( f \). The scoring rule is specified by a set of weights \( \Lambda = \{\lambda_1, ..., \lambda_L\} \) with \( \sum_{l=1}^{L} \lambda_l = 1 \) that defines the requester’s preferences of each attribute type analog to the definition of scoring rules in (Asker and Cantillon, 2008). To assure comparability of attribute values from different attribute types the aggregated attribute values \( A_j^l \) are mapped on an interval \([0;1]\). \( T_f \) represents the overall price of the complex service. \( \alpha \) can be interpreted as the willingness to pay for an optimal configuration \( S(A_j) = 1 \) based on the requester’s score. In other words \( \alpha \) defines the substitution rate between configuration and price based on the requester’s preferences.

Definition 1. Multiattribute Service Request

A request for a complex service is a vector of the form

\[ R := (G,F,\alpha,\Lambda,\Gamma) \]

with \( G \) represents a complex service network, \( F \) represents all feasible paths from source to sink that form a possible instantiation of a complex service, \( \Lambda \) the requester’s preferences and \( \alpha \) the willingness to pay. \( \Gamma \) denotes the set of lower and upper boundaries for each attribute type.

A service offer consists of an announced service configuration \( A_j \) and a corresponding price bid \( p_y \) that a service provider wants to charge for service \( j \) being invoked depending on the predecessor service \( i \) such that \( b_y(e_y) = (A_j, p_y) \) is a service offer bid for invocation of service \( j \) which interoperable with a predecessor service \( i \) with \( b : E \rightarrow A \times R \). A service provider \( s \) bids for all incoming edges to every service it owns.

Definition 2. Multiattribute Service Offer

A multiattribute service offer is a bid matrix of the form

\[ B' := \begin{cases} b_y(e_y) = (A_j, p_y), & i \in \tau(f), j \in \sigma(s) \\ 0, & \text{otherwise} \end{cases} \]

with \( \tau(v) \) denotes the set of all predecessor services to service \( v \) with \( \tau : V \rightarrow V \) and \( \sigma(s) \) the set of all services owned by service provider \( s \).

4.2 Mechanism Design

The mechanism maximizes welfare by allocating a path \( f^* \) within the service value network that yields the highest overall utility. Let \( \mathcal{U}_f \) denote welfare induced by path \( f \) with \( \mathcal{U}_f = \alpha S(A_j) - P_f \).

\[ o := \arg\max_{f \in F} \mathcal{U}_f \]

Let \( \mathcal{U}^* \) denote the utility of the winning path meaning the utility of a path \( f^* \) that maximizes welfare. Let \( \mathcal{U}_f^* \) denote the utility of a path \( f^*_s \) that yields a maximum overall utility in the reduced graph without every service owned by service provider \( s \) and its incoming and outgoing edges.
Every service provider $s$ receives a payment or transfer $t^i = \sum_{i \in \sigma(s), j \in \sigma^c(s)} t^i_j$ for all services it owns which are on the winning path. A payment $t^i_j$ for service $j$ corresponds to the monetary equivalent of the utility gap between the “winning path” and “second best path”. In other words a monetary equivalent to the utility service $j$ contributes to the systems welfare. This monetary equivalent represents the price that service provider $s$ could have charged without losing her participation in the winning allocation.

(5) $t^i_j := p^i_j + (U^i - U^i_s)$

Consequently the payment function $t^i$ for service provider $s$ is defined as

(6) $t^i := \sum_{i \in \sigma(s)} \sum_{j \in \sigma^c(s)} p^i_j + (U^i - U^i_s), \text{ if } e^i_j \in o
\left\{ \begin{array}{ll}
0, & \text{otherwise}
\end{array} \right.$

Costs $c^i$ that service provider $s$ has to bear for performing offered and allocated services result accordingly:

(7) $c^i := \sum_{i \in \sigma(s)} \sum_{j \in \sigma^c(s)} c^i_j(e^i_j, A^i_j), \text{ if } e^i_j \in o
\left\{ \begin{array}{ll}
0, & \text{otherwise}
\end{array} \right.$

The solution to the allocation problem in (4) can be computed in polynomial time using well-known graph algorithms to determine the "shortest" within a network such as the Dijkstra algorithm. Using a Fibonacci heap data structure the time complexity can be reduced to $O(n \log(n) + m)$ with $m$ is the number of edges and $n$ the number of nodes within the graph. According to the payment scheme in (6) the allocation must be computed twice: Based on the graph with the service offerings of the service provider receiving the payment and without its participation. In the second case the graph can be pre-processed and reduced by all service offerings owned by the service provider that receives the payment. After the reduction the shortest path can be computed accordingly which yields the same time complexity. In contrary to the NP-hard complexity in general combinatorial auctions this is a valuable achievement that enables the conduction of our auction in online systems.

4.3 Aggregation and Preference Mapping of Service Attributes

In order to determine the overall score for a provider based on its scoring function, the attribute values of the complex service have to be computed. Recall, the type of function for aggregating attribute value highly depends on the attribute type. Traditional quality of service attributes such as response time for example can be aggregated with basic mathematic operations such a sum operator. Table 2 shows different types of aggregation functions for multiple attribute types exemplarily. For example, the overall throughput of a complex service that consists of multiple service components is determined by the lowest throughput rate within the allocation and can therefore be computed using a minimum operator.

<table>
<thead>
<tr>
<th>Attribute Type</th>
<th>Aggregation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l$</td>
<td>$\oplus_{a \in f} d^l_j$</td>
</tr>
<tr>
<td>Response Time (rt)</td>
<td>$\sum_{a \in f} a^r_{j}$</td>
</tr>
<tr>
<td>Encryption Type (et)</td>
<td>$\land_{a \in f} a^e_{j}$</td>
</tr>
<tr>
<td>Error Rate (er)</td>
<td>$\max_{a \in f} a^e_{j}$</td>
</tr>
</tbody>
</table>
Table 2. Aggregation Functions for Different Types of Attributes

Nevertheless, only considering basic quality of service attributes is not sufficient for dealing with complex non-functional service characteristics that express rich semantic information. The auction mechanism must be capable of aggregating a broad range of descriptive service attributes that express multiple quality aspects. The following example depicted in Figure 2 shows a service value network with four service offers and three possible paths from source to sink (top, middle, bottom).

Figure 2 Numerical Example

For simplicity and without loss of generality we assume that each service provider owns only a single service. Price values on the edges represent price bids announced by service providers. Each service configuration consists of attribute values for the types encryption \(a^e\) and probability of success \(a^s\). Attribute values are aggregated according to the aggregation operations in Table 2. Encryption types are derived from the concepts in the security algorithm ontology as illustrated in Figure 3.
The service requester announces its willingness to pay and weights for each attribute type representing its scoring function such that $\lambda_p = 0.2$, $\lambda_m = 0.8$ and $\alpha = 100$. Furthermore it specifies the *individual encryption* attribute type in first order logic:

$$\text{IndividualEncryption}(x) \leftarrow (\text{BlockCipher}(x) \wedge \text{hasKeyLength}(x, k) \wedge \text{isGreaterOrEqual}(k, '128')) \lor (\text{AsymmetricAlgorithm}(x) \wedge \text{hasKeyLength}(x, k) \wedge \text{isGreaterOrEqual}(k, '256'))$$

The mechanism allocates service offers on a path from source to sink based on the service request and announced multiattribute offers according to (4). The welfare level, generated by each allocation evolves as follows:

- Middle: $\mathcal{U}_{middle}^f = 100(0.2(1 \wedge 1) + 0.8(0.9 \times 0.8)) - (13 + 17) = 47.6$
- Top: $\mathcal{U}_{top}^f = 100(0.2(1 \wedge 0) + 0.8(0.9 \times 0.7)) - (13 + 16) = 21.4$
- Bottom: $\mathcal{U}_{bottom}^f = 100(0.2(0 \wedge 1) + 0.8(0.9 \times 0.8)) - (10 + 20) = 27.6$

Therefore $f_{middle}$ is allocated as it yields the highest welfare and each service provider that owns a service on it receives a payment according to (6) such that $t^1 = 13 + (47.6 - 27.6) = 33$ and $t^4 = 17 + (47.6 - 21.4) = 43.2$. The transfer is designed to compensate service providers for their contribution to the system’s welfare which implies that i.e. provider 1 could have bid a price of 33 without having lost its participation in the allocation.

### 4.4 Verification of Service-Level-Agreements

As introduced in Section 4.1 service providers’ bids contain a configuration and a price component. The allocation function maximizes welfare based on the achieved quality for the service requester and the costs that occur on the producer’s side. This shows that the announced quality also determines the likelihood of being allocated which might induce and incentive for service providers to lie about their configuration. Therefore proposed mechanism is extended with a so called ex-post verification term which is explained in detail in this section.

Let $a^j_l$ be the announced attribute value of attribute type $l$ of service $j$’s configuration. Furthermore let $\tilde{a}^j_l$ be the actual attribute value of attribute type $l$ realized ex-post by service $j$ during execution. $\tilde{U}^f$ is the overall winning path utility with the actual realized attribute values $\tilde{a}^j_1, ..., \tilde{a}^j_l$ of service $j$.

Auctioning services based on a platform approach opens up the possibility of ex-post verification. This means that the actual delivered quality of participating services can be measured and monitored after execution. Therefore we can ex-ante enforce a true announcement of quality to be delivered by verifying it ex-post. According to the *Compensation-and-Bonus Mechanism* introduced in (Nisan and Ronen, 2001) a compensation function is constructed as follows

$$\Delta f_j = \begin{cases} (\tilde{U}^f - \tilde{U}^j), & \text{if } e_j \in o \\ 0, & \text{otherwise} \end{cases}$$

The compensation function represents the utility gap that results from the utility difference of the announced attribute values and the actual performed ones from the service requester's perspective. In other words the gap that results from the utility loss the systems incurs because of the service provider's untruthful announcement. The monetary equivalent to this utility gap according to the requester’s preferences represents the penalty payment the service provider has to bear for deviating from the announced attribute values. This negative consequence can be interpreted as a contractual penalty for not realizing specified service-level-agreements as defined in (Salle and Bartolini, 2004).

Taking the compensation function into account the payment function is extended as follows:
4.5 An Analytical Analysis of Bidding Strategies of Service Providers

The bidding strategy of each service provider comprehends a price announcement and a corresponding service configuration consisting of a set of attribute values as introduced in the Section 4.1. In this section we analytically analyze providers’ bidding strategies in proposed mechanism design:

**Lemma 1.** In a Multiattribute Verification Mechanism for each service provider $s \in S$ the reward is independent from its bids consisting of the announced attribute values $a^s_1, \ldots, a^s_L$ and the announced prices $p_{ij} \forall i \in \tau(j), \forall j \in \sigma(s)$.

**Proof of Lemma 1.** Assuming without loss of generality that service provider $s$ only owns one service $z$ with a configuration $A_z$. $F_z$ denotes the set of all feasible paths from source to sink without service $z$ and its incoming and outgoing edges. We denote $f^*$ as the path which is allocated by $o$. Let $U'_z$ be the utility of path $f^*$ in the graph with service $z$. Let $U_z$ be the utility of path $f'_z$ in the reduced graph without service $z$ and its incoming and outgoing edges. Let $U'_{z}(e_{s})$ denote the utility of path $f'_{s}$ with $e_{s} \in f'_{s}$. $U'_z$ is the overall winning path utility with the actual realized attribute values $\hat{a}_1^s, \ldots, \hat{a}_L^s$ of service $z$. An invocation $e_{s}$ of service $z$ is allocated by $o$ iff $U'_z(e_{s}) \geq U_z$. In this case the profit of service provider $s$ evolves as follows:

$$\pi^s = p_{s}^a + t^s - c^s$$

$$\pi^s = p_{s}^a + (U'_z - U_z) - c_{s}(e_{s}, \hat{A}_z)$$

$$\pi^s = p_{s}^a + ((U'_z - U_z) - (U'_z - U'_z)) - c_{s}(e_{s}, \hat{A}_z)$$

$$\pi^s = p_{s}^a + (U'_z - U_z) - c_{s}(e_{s}, \hat{A}_z)$$

$$\pi^s = \left(\alpha \left( \sum_{l=1}^{L} \check{A}_l \right) - \sum_{e_{s} \in \tau(j)} p_{ij} - U_z\right) - c_{s}(e_{s}, \hat{A}_z)$$

Equation (10) shows that once a service $z$ is allocated, its reward is independent from its announced price $p_{s}$ and all announced attribute values $a^s_1, \ldots, a^s_L$. In other words $s$’s bid does not have an impact on the transfer function $t^s$. □

**Theorem 1.** In a Multiattribute Verification Mechanism, for each service provider $s \in S$ the bidding strategy $b_{ij}(e_{s}) = (A_{ij}, p_{ij})$ with $\hat{a}^s_l = a^s_l \forall l \in L, p_{ij} = c_{ij}(e_{s}, \hat{A}_z)$ (truth telling with respect to configuration and price) $\forall i \in \tau(j), \forall j \in \sigma(s)$ is a weakly dominant strategy.

**Proof of Theorem 1.** Lemma 1 shows that once service providers are allocated they are not able to influence their reward as $\pi$ is independent from the announced attribute values and prices. Nevertheless, bids have an impact on the chance of being allocated. Assuming without loss of generality that service provider $s$ only owns one service $z$ with a configuration $A_z$. A service provider $s$ wants to be allocated iff $\pi^s > 0$.
A possible solution that satisfies (11) is truth-telling with respect to configuration and price such that $p_a = c_a(e_a, \tilde{A})$ and $\mathcal{U}'(e_a) = \mathcal{U}'$. As shown in Lemma 1 service providers have no control about their reward once they are allocated which implies that any other possible solution besides truth-telling that satisfies (11) is not better than truth-telling. Hence, reporting attribute values $a'_1, \ldots, a'_l$ truthfully meaning that announced values are actually realized through execution such that $a'_l = a'_l \forall l \in L$ and consequently $\mathcal{U}'(e_a) = \mathcal{U}'^*$ as well as $p_a = c_a(e_a, \tilde{A})$ is a weakly dominant strategy.

Theorem 1 shows that the provider’s bidding strategy is determined through the mechanism design. Service providers act best (or at least as good as any other alternative) by reporting their services’ configurations and internal costs truthfully which is a valuable mechanism property from a requester’s perspective. This property assures that although all service providers act self-interested and therefore try to maximize their profit, their dominant strategy maximizes the system’s welfare and the requester receives a technical feasible instantiation of the desired complex service at a guaranteed service level. This is a valuable property as it tremendously lowers strategic complexity for service providers and fosters a trustful requester-provider-relationship. It is well-known in literature that incentive compatibility in VCG-based mechanisms may fail in repeated games (Binmore and Swierzbinski, 2000). Nevertheless, in service value networks we observe a high degree of dynamicity with respect to changing service providers, variable costs and network topologies. Thus, each auction setting is different from the preceding one which makes learning from past situations impossible and each game can therefore be treated as a one-shot game.

5 CONCLUSION

We proposed a multidimensional procurement auction for trading composite services in networked economies. We presented a graph-based model that captures the main components and characteristics of service value networks. Based on this model we introduced a mechanism design that enables allocation and pricing of service components that together form a requested complex service. However, auctions for composite services are much more complex than simple procuring auctions, where the suppliers themselves offer a full solution to the procurer. In composite services, this is not the case, as a flawless service execution and therefore the requestor’s valuation highly depends on the accurate sequence of its functional parts, meaning that in contrary to service bundles, composite services only generate value through a valid order of their components. The allocation is computed based on the requestor’s score for QoS characteristics of the complex service. At the same time, the mechanism is capable of handling a wide range of aggregation operations for service attributes also supporting rich semantic approaches for dealing with complex non-functional service specifications. Due to the combinatorial restrictions imposed by the underlying graph topology and the absence of capacity constraints, the winner determination problem can be solved in polynomial time which is a crucial issue when it comes to implementing online systems. We furthermore showed that proposed mechanism is individual rational, allocation efficient and incentive compatible with respect to QoS characteristics and prices of service offers. Hence, reporting the true type regarding configuration and price is a weakly dominant strategy for all service providers. This is a valuable property as it tremendously lowers strategic complexity for service providers and fosters a trustful requester-provider-relationship.
Proposed graph-based scoring auction is evaluated in the TEXO use case of the THESEUS² project. TEXO is a research project, within the Theseus research program initiated by the Federal Ministry of Economy and Technology (BMWi). Within the Theseus program, TEXO contributes to service economy by creating infrastructure components for business webs in the Internet of Services. Via intuitive interfaces and technical systems TEXO addresses the full lifecycle of these services from innovation to productive usage. Addressing these demands requires an interdisciplinary approach to create an integrated platform for the internet of services which supports all phases of the lifecycle. For all stakeholders and participants in such a service value network, innovative business models being as flexible as the network itself are required. Especially the novel requirements for pricing models are addressed by proposed graph-based multidimensional procurement auction. The auction mechanism is capable of allocating and pricing of composite services in an efficient and truthful manner. It enables flexible participation and switching for service providers and at the same time it does not require complete information about configurations, prices and interrelations from the service requester’s perspective which makes the mechanism favourable for ad-hoc and situational environments such as service value networks.

References


² http://theseus-programm.de/front


UNDERSTANDING THE MANIFOLD FORMS OF B2B INTEGRATION - A TRANSACTION COST PERSPECTIVE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0685.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
</tbody>
</table>
UNDERSTANDING THE MANIFOLD FORMS OF B2B INTEGRATION - A TRANSACTION COST PERSPECTIVE

Christine Legner, European Business School (EBS), Rheingaustr. 1, 65375 Oestrich-Winkel, Germany, christine.legner@ebs.edu

Abstract

The need for collaboration within value chains is rapidly increasing and drives enterprise to align and electronically integrate their business processes with their business partners. As technologies evolve, manifold forms of B2B integration have emerged – from e-mail communication to customer or supplier portals, the exchange of EDIFACT- to XML documents, and Web Services. Although service-oriented architectures (SOA) are considered the future of inter-organizational linkages, no empirical studies have been found which surveyed the impact of SOA on B2B integration costs and benefits. From a research perspective, we still lack a systematic analysis that explains how a specific B2B integration technology impacts the effectiveness of B2B integration.

Building on transaction cost theory, this research analyzes the different forms of B2B integration with regard to their impact on connectivity and coordination costs. Based on a field study from the automotive industry, it demonstrates that there is economic rationale for preferring supplier portals to machine-to-machine integration based on EDIFACT or XML messages. Compared to prior technologies, SOA reduces the costs of external integration by eliminating separate B2B integration infrastructures and improving connectivity of internal applications. However, we find that prior literature tends to overestimate the impact of open Internet and Web service technologies on connectivity costs.

Keywords: B2B integration, e-business, electronic data interchange, inter-organizational systems, transaction costs, Service-oriented architecture (SOA)
1 INTRODUCTION

The need for collaboration within value chains is rapidly increasing in many industries and drives enterprise to align and electronically integrate their business processes with their business partners. A huge growth of inter-organizational process integration is projected in the coming years due to the following three factors. First, enterprises increasingly realize that they gain competitive advantage by intensifying their customer interactions, concentrating on core competencies and augmenting the level of external sourcing (Dyer and Singh 1998; Venkatraman and Henderson 1998). This is underpinned by prominent examples from the automotive industry (Dannenberg and Kleinhans 2004) and financial services (Sydow et al. 1998). Second, companies have significantly increased internal process and systems integration over the past decade which is considered a major enabler of external integration (Zhu et al. 2006; Zhu et al. 2004). Third, inter-organizational systems technology has been completely overhauled since the mid 1990s by the emergence of the Internet. Today, the Internet provides a widely accepted infrastructure for e-business. Web services and service-oriented architectures (SOA) have emerged as an enhanced concept for integration in heterogeneous environments and are expected to stimulate inter-organizational process integration (Daniel and White 2005; Hagel and Brown 2001).

Inter-organizational systems (IOS) have been intensively studied since the 1960s, when the online airline ticketing system SABRE and other early forms of IOS established electronic linkages between business partners. Despite the vast body of IOS research that analyzes characteristics and benefits of B2B integration (e.g. Choudhury 1997; Massetti and Zmud 1996; Mukhopadhyay and Kekre 2002; Saeed et al. 2005), little research has been conducted so far with regard to analyzing and comparing the different forms of B2B integration and their effectiveness. While there is consent that the use of electronic channels reduces transaction costs (Grover et al. 2002; Malone et al. 1987), it is not well understood how technology choices affect the savings. In collaborative B2B relationships, business partners negotiate mid- to long-term contractual agreements that govern a larger number of transactions (Christiaanse et al. 2004; Clemons and Row 1992; Grover et al. 2002). As they seek investments in electronic linkages in order to reduce high-perceived transaction costs, they have the choice between manifold B2B integration options, from e-mail communication to customer or supplier portals, the exchange of EDIFACT- to XML documents, and Web Services. This paper takes on these challenges and aims at answering the following research questions:

1. How do different forms of B2B integration impact transaction costs in B2B relationships?
2. Is there empirical evidence that Web services and SOA overcome the shortcomings of prior forms of B2B integration and thereby are more likely to experience broad adoption?

Building on transaction cost theory and prior IOS literature, the author suggests distinguishing connectivity costs and coordination costs when analyzing the effectiveness of different forms of B2B integration. The resulting model has been applied in the automotive industry in order to explore the costs and benefits of B2B integration. While the field study presents a first attempt towards measuring transaction costs of electronic B2B relationships, the main contribution of this paper is the conceptual model and the operationalization of transaction cost elements. Hence, this research is intended to further stimulate the academic discourse on the effectiveness of B2B integration and the measurement of transaction costs in electronic B2B relationships.

The remainder of this article is structured as follows: The next section outlines the research methodology. From the review of prior research on IOS and transaction cost theory, we subsequently derive a conceptual model for analyzing different forms of B2B integration and their effectiveness. By applying the model in a field study in the automotive industry, we were able to explore the impact of five different forms of B2B integration on transaction costs. The article concludes with a summary of the insights related to current forms of B2B integration and the impact of service-oriented concepts.
2 RESEARCH METHOD

In view of the research objectives, we adopted an exploratory research design. Using transaction cost economics as a framework of analysis, the aim was to explore how the different forms of B2B integration impact the transaction costs. In a first step, a literature review was performed in order to systemize the different forms of IOS and to identify different transaction cost elements that characterize B2B integration. Given the lacking measurement approaches employed in the transaction cost literature, we derived a refined model that comprises the main transaction cost elements in B2B integration. The second step in the research process involved the collection and analysis of data that are related to the different forms of B2B integration and their impact on transaction costs. Since many industries either have limited experience with electronic integration or strongly rely on a dominant form of B2B integration (European Commission 2007), we had to carefully select a scenario that would provide the possibilities of analyzing and comparing different forms of B2B integration. In addition, accessibility of data was an important factor. For the purpose of gathering data from different B2B integration variants, we decided to focus on the specific scenario of engineering change management in the automotive industry. The following criteria were decisive: (1) The automotive industry has broad experience in B2B integration due to its long history in EDI-based supplier relationships. In this regard, automotive manufacturers and suppliers are aware of the various issues involved in B2B integration and can also be considered “IT-savvy”. (2) Engineering change management has been subject to a recent industry standardization initiative by the Association of German Automotive Manufacturers (VDA). This initiative resulted in VDA Recommendation 4965 which represents a well-documented and comprehensive industry standard. In view of the wide range of implementation variants, which range from manufacturer-neutral clients to EDI, automotive companies are pressing for solutions that have greater interoperability through SOA and Web services.

Over a period of 15 months, from October 2005 to February 2007, the author was heading a research team that conducted a field study in the automotive industry. As an active member of the VDA initiative on Engineering Change Management (VDA ECM), the research team participated in regular working meetings and contributed to the specification of XML messages. In addition, it supported a sub-group of the VDA initiative in designing a service-oriented architecture for the electronic integration of their engineering change management processes. This sub-group consisted of one automotive OEM and four suppliers, and was supported by several technology providers. This activity implied intensive collaboration between all participants. The results were a pilot implementation of the ECM scenarios based on SOA and an in-depth evaluation of the suggested approach compared to the more traditional forms of B2B integration. The evaluation was performed based on the conceptual model presented in this paper. It builds on the experiences from the pilot implementation and estimations that the automotive experts gained when realizing other B2B integration projects.

3 PRIOR RESEARCH

3.1 Inter-organizational Systems (IOS)

According to Johnston and Vitale (1988) and Hong (2002), inter-organizational systems (IOS) are network-based information systems that transcend organizational boundaries. The many different forms of IOS that exist currently reflect the evolution of B2B integration technologies. They range from the first EDI-based systems that were established in the 1960s to the Internet that gave raise to the concepts of e-business and e-commerce, and, more recently, Web services and Service-oriented Architectures. Many scholars argue that service-oriented concepts will provide a more cost-effective e-business platform than traditional EDI systems (Dorn et al. 2009; Legner and Vogel 2008). They are expected to cope better with the differences in semantics and pragmatics among the different actors,
and to replace the traditional document-centric approaches to B2B integration by a process-centric approach. IS researchers have proposed various categorization schemes and modes in order to systemize the different levels of external integration and their support for different types of supplier-buyer relationships (e.g. Choudhury 1997; Massetti and Zmud 1996; Mukhopadhyay and Kekre 2002; Saeed et al. 2005). IOS can be classified according to the type of electronic interaction that they support and the topology of the IT-supported inter-firm relationship (c.f. Table 1):

- The interaction type depends on the communication channel (Löwer 2005; McAfee 2005; Reimers 2001): Electronic human-human-interaction describes traditional forms of interaction between humans that are supported by electronic means, e.g. e-mail or video-conferencing. In the case of human-machine-interaction, external users are gaining direct access to shared data and applications. This is typically realized by Web front-ends or portals that bundle data and applications on the basis of users and roles. Machine-machine interaction finally describes the direct communication between two information systems which eliminates human intervention. It can be achieved by file transfer or by message exchange, which are both associated with asynchronous communication, as well as by service calls.

- The topology of the IT-enabled inter-firm relationship refers to the relationships between the business partners (Alt 2008). 1:1 connections represent dyadic relationships, e.g. between a customer and a supplier. Over time, they often evolve into 1:n or n:1 connections when a focal firm starts linking up with a larger number of (smaller) business partners. The shift to m:n relationships requires either the adherence to widely accepted standards (Damsgaard and Truex 2000) or the existence of an intermediary that facilitates multilateral electronic relationships (Giaglis et al. 2002).

<table>
<thead>
<tr>
<th>Interaction Type (Löwer 2005; McAfee 2005; Reimers 2001)</th>
<th>Topology of IT-enabled inter-firm relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>human-human-interaction: interaction between humans supported by electronic means</td>
<td>1:1 (dyadic)</td>
</tr>
<tr>
<td>human-machine-interaction: external users are getting direct access to data and applications</td>
<td>1:n or n:1 (focal)</td>
</tr>
<tr>
<td>machine-machine-interaction: direct communication between two information systems without human intervention</td>
<td>m:n (multilateral / exchange)</td>
</tr>
<tr>
<td>Direct communication between information systems based on bilaterally agreed specifications:</td>
<td>Direct communication between information systems based on specifications that are imposed by one partner:</td>
</tr>
<tr>
<td>• File transfer</td>
<td>• File transfer</td>
</tr>
<tr>
<td>• Message exchange</td>
<td>• Message exchange</td>
</tr>
<tr>
<td>• Web services</td>
<td>• Web services</td>
</tr>
<tr>
<td>Multilateral communication between information systems mediated by an intermediary:</td>
<td>Multilateral communication between information systems based on standards:</td>
</tr>
<tr>
<td>• Electronic marketplace</td>
<td>• File transfer</td>
</tr>
<tr>
<td>• Exchange</td>
<td>• Message exchange</td>
</tr>
<tr>
<td>• Shared service</td>
<td>• Web services</td>
</tr>
</tbody>
</table>

Table 1. Categorization of Different Forms of IOS
3.2 Transaction Cost Theory

Transaction costs, originally developed by Coase (1937), are key in explaining the impact of electronic integration and assessing its benefits. According to transaction cost theory, the most efficient form of the inter- and intra-organizational structure minimizes transaction and production costs (Rindfleisch and Heide 1997). Although the use of IT is generally considered to reduce transaction costs (Malone et al. 1987), existing studies apply transaction cost theory rather at a conceptual rather than at a measurement level. Williamson (1985) initially proposed a division between ex-ante costs (costs prior to the execution of a transaction, i.e. partner search and contract negotiation) and ex-post costs (occurring during and after a transaction, i.e. policing and enforcement costs). With respect to the shortcomings of Williamson’s approach, Milgrom and Roberts (1992) propose the categorization of transaction costs into motivation and coordination costs which encompass the cost of obtaining information, the cost of coordinating the production process as well as the cost of measurement. Clemons and Row (1992) and Clemons et al. (1993) distinguish between two components of transaction costs, namely costs of coordination and costs of transaction risks. They argue that IT investments are asset-specific since they are idiosyncratic to the relationship with the other firm.

While the transaction cost-standpoint has been widely adopted in studies related to supply chain integration, outsourcing, and electronic markets, most studies are focussing on the negotiation and transaction phase. The relational perspective in transaction cost research (Christiaanse et al. 2004; Clemons and Row 1992; Grover et al. 2002) argues that it is too narrow to restrict the unit of analysis to a pure sales transaction (instead of the exchange relationship). In B2B networks, firms negotiate mid- to long-term contractual agreements that govern a larger number of transactions. In such relational structures, they may seek bilateral investments for setting up electronic linkages in order to reduce high-perceived transaction costs (Grover et al. 2002). In terms of transaction cost theory, this translates into asset-specific (or relationship-specific) investments and high transaction frequency.

4 CONCEPTUAL MODEL

Based on the relational perspective in transaction cost theory and propositions from IOS research, we suggest that two main cost elements characterize electronic B2B integration, namely connectivity and coordination costs (c.f. Table 2).

Connectivity costs denote non-recurring costs to establish an electronic business relationship. Prior literature on transaction costs emphasizes partner finding and contractual negotiation, i.e. the time spent by both transacting sides on agreeing on contractual terms. In the context of B2B integration, significant costs are incurred for the design and implementation costs of electronic process integration. Prior literature on IOS adoption emphasizes the effort to establish inter-organizational agreements (Kubicke 1992; McAfee 2005; Reimers 2001) and the internal costs incurred for implementing external process integration (Zhu et al. 2006). Since early IOS have mostly been built as proprietary systems, relationship-specific or asset-specific investments in joint infrastructure are considered to be more important in the case of older communication technology applications such as EDI and other proprietary networks. The use of open technologies, such as Web service- or XML-based applications, is expected to reduce asset-specific investments for firms (Christiaanse et al. 2004). Hence, the following cost elements that make up the connectivity costs of enterprise $i$ connecting with $n$ business partners can be derived:

$$C_{connect_i} = \sum_{j=1}^{n}(C_{P.agr.ij} + C_{P.impl.ij} + C_{IS.agr.ij} + C_{IS.impl.ij})$$

Costs occur for establishing inter-organizational agreements between two business partners $i$ and $j$ and for their subsequent (internal) implementation. Since we have to consider the organizational and the
technical aspects in setting up electronic B2B relationships, we distinguish costs at the process and IS layer. With regard to the process layer, the efforts in establishing an inter-organizational agreement on the process interaction between enterprises \(i\) and \(j\) result in \(c_{P,agr,ij}\). The costs for the internal implementation of the partner-specific process and organizational changes are denoted by \(c_{P,impl,ij}\). On the IS layer, \(c_{IS,agr,ij}\) are the costs for the inter-organizational agreement on the IS interface (or services) and communication infrastructure, whereas \(c_{IS,impl,ij}\) are the costs incurred for their implementation.

Coordination costs are the recurring costs to enable and execute a transaction. Electronic integration reduces coordination costs (Grover et al. 2002; Malone et al. 1987). However, if an IOS is used to automate an existing process, its effects are limited to reducing manual data processing and improving the reliability as well as the timeliness of information (Hoogewegen and Wagenaar 1996). As the firms progress to using IOS for closely coupling business processes between firms, they are able to realize additional benefits of vertical integration (Mukhopadhyay and Kekre 2002; Saeed et al. 2005; Zhu et al. 2004). The use of IOS is considered most beneficial if applied in cooperative relationships (Chatfield and Yetton 2000; Johnston and Vitale 1988) and accompanied by process innovation such as vendor-managed inventory or continuous replenishment in the retail and consumer goods industries (Clark and Stoddard 1996; Riggins and Mukhopadhyay 1994). The effectiveness of B2B integration can be measured by reduced coordination costs, i.e. the savings realized in executing and monitoring transactions. These savings depend on the number of transactions \(t_n\) conducted with all \(n\) business partners and the savings in coordination costs per transaction.

\[
\Delta C_{coord} = \sum_{j=1}^{n} \sum_{k=1}^{t_n} (\Delta C_{coord,ijk})
\]

<table>
<thead>
<tr>
<th>Transaction phase / costs</th>
<th>Description</th>
<th>Propositions from IOS literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex-post: Coordination costs</td>
<td>Recurring costs to enable and execute a transaction</td>
<td>Electronic integration reduces coordination costs (Grover et al. 2002; Malone et al. 1987). IOS are most beneficial if applied in cooperative relationships (Chatfield and Yetton 2000; Johnston and Vitale 1988) and accompanied by process innovation (Clark and Stoddard 1996; Riggins and Mukhopadhyay 1994).</td>
</tr>
</tbody>
</table>

Table 2. Transaction Cost Breakdown for Assessing B2B Integration

5 FIELD STUDY IN THE AUTOMOTIVE INDUSTRY

For the purpose of exploring the impact of different forms of B2B integration on transaction costs, the conceptual model has been applied in a field study in the automotive industry. The following section presents the findings of the study.
5.1 Background: Engineering Change Management in the Automotive Industry

Although it has been intensively studied for a long time, the automotive value chain is currently undergoing significant changes (Dannenberg and Kleinhans 2004; Doran et al. 2007; Seidel et al. 2005): Manufacturers are concentrating on branding and downstream activities such as marketing and after-sales, while tier-1 suppliers are increasingly taking over engineering, production and assembly of major components and even entire vehicles (Coronado Mondragon et al. 2006). Today, more than 65% of the value is created within the supplier network with a projected increase to 77% over the next decade (Dannenberg and Kleinhans 2004). With the changing roles in the automotive industry, the need for inter-organizational coordination increases and encompasses not only supply chain management, but also the innovation and product development processes. As part of product life cycle management, engineering change management refers to evaluating and deciding ideas for change as well as the implementation of the changes in development and production. Possible triggers include changes in product design or the elimination of quality and/or safety defects. Currently, suppliers are directly affected by approximately 30% of the more than 10 000 engineering changes an OEM processes per year. Engineering change management is the subject of current standardization efforts which have resulted in VDA Recommendation 4965 of the German Association of the Automotive Industry (VDA 2005) and which are currently brought to the international level.

5.2 Forms of B2B Integration

As part of the field study, five different forms of B2B integration have been analyzed. Interestingly, the group of automotive companies has never used intermediaries to electronically support interactions between OEMs and tier-1 suppliers, despite the intensive discussions in practitioner and academic publications about the role of Covisint and its recent shutdown. The companies had gained experiences with the following five different forms of B2B integration:

- **E-mail communication**: While e-mail communication represents the most widespread form of electronic interaction in the engineering change management scenario, it is mostly used to notify external partners about engineering changes and to share documents with them. Since no structured information is exchanged, e-mail communication corresponds to an electronic human-human interaction in a dyadic OEM-supplier relationship (1:1).

- **Portals**: In order to further automate and streamline their B2B interactions, OEMs have established supplier portals. These portals support interactions with suppliers and provide them with front-end access to the OEM’s internal applications, including engineering change management systems. Supplier portals are to be classified as human-machine interaction and support 1:n relationships between one OEM and its suppliers.

- **Message exchange based on individual data format**: This form of B2B integration is used if two firms bilaterally agree on a data format for exchanging engineering change messages electronically. It represents a dyadic relationship with bilateral agreements between the OEM and supplier (1:1).

- **Message exchange based on industry standard**: This option corresponds to the situation when the message format specified by the VDA ECM standard is used to electronically exchange engineering changes between business partners. In this case, multi-lateral agreements on ECM messages are a prerequisite (m:n).

- **SOA-based process integration based on industry standard**: The SOA-based approach allows for external process integration based on Web service calls. For the engineering change management scenario, the reference architecture outlined by Legner and Vogel (2008) translates the VDA recommendation into a public process model and public service interface (m:n). It has been subject of a pilot implementation.
As part of the field study, data was gathered based on prior experiences with these five forms of B2B integration and complemented with data gathered by VDA (2005). In the case of SOA-based connections, results from the pilot implementation were used. Table 3 presents an overview of the findings which will be discussed in more details in the following two sections.

5.3 Findings Related to Coordination Costs

As discussed earlier, the effectiveness of B2B integration depends on its effect on coordination costs, i.e. the savings in executing and monitoring transactions. In the specific case of engineering change management, the benefits of electronic integration only materialize if engineering changes are exchanged as structured messages that can be processed and imported into the company-specific engineering change management systems. The exchange of electronic messages significantly reduces the manual efforts that are caused by business partners requesting to comment on engineering changes by using their company-specific templates and terminology. With regard to coordination costs, no difference has been identified between the three forms of electronic machine-to-machine integration. Since all of them realize electronic system integration, the reduction in coordination costs amounts to 0.75 person-days per engineering change request (for both parties). On average, three external interactions are required for processing an engineering change request, with an estimated effort of 0.25 person-days for collecting and organizing information. Compared to machine-to-machine integration, the drawbacks from portal-based human-to-machine integration are obvious. Given that suppliers are required to manually re-enter data and to adopt the OEM-specific business and process logic, the estimated savings of 0.75 person-days only apply for the OEM.

5.4 Findings Related to Connectivity Costs

Connectivity costs are incurred for establishing inter-organizational agreements between the different parties and their subsequent (internal) implementation. While connectivity costs are insignificant in the case of e-mail communication, they are also a minor factor in the case of supplier portals where the OEM incurred most of the costs. However, connectivity costs represent an “entry barrier” when it comes to the establishment of machine-to-machine linkages between manufacturers and their suppliers. Table 3 presents the connectivity costs which we identified for the different forms of B2B integration in the engineering change management scenario. Estimations by the VDA ECM working group were used for quantifying the efforts related to inter-organizational agreements (here: $c_{P,agr,i}$ and $c_{IS,agr,i}$). It has proven impossible to further quantify the internal efforts for linking up $c_{P,impl,i}$ and $c_{IS,impl,i}$ since they largely depend on the internal processes and applications. However, Table 3 is quite informative in several ways: First, the inter-organizational agreements at the process level generate significantly more effort (15 days) than those at the IS level (5 days). Second, automotive companies experience that none of the approaches to B2B integration completely eliminates bilateral efforts. Even if public constructs are defined, some bilateral negotiations will be necessary to analyze and set up the collaboration (5 days). Third, Table 3 confirms that SOA-based process integration addresses the shortcomings of existing B2B integration approaches and significantly reduces the relationship-specific investments at process and IS level. Whereas the reduction of relation-specific investments compared to point-to-point connections is obvious, the comparison with the document-centric B2B standards is more interesting. The latter indicates significant relationship-specific costs for inter-organizational process alignment. In the case of SOA, these costs can be reduced if the vertical standard specifies public SOA constructs in the form of process and IS interface specifications. Finally, automotive experts stated that SOA will further reduce the IS-level implementation efforts although they have had difficulties to quantify these effects in the concrete case. They argue the following. (1) An SOA eliminates the need for maintaining a separate B2B integration infrastructure and is more scalable and flexible than existing adapters; (2) widely used Internet technologies will require less proprietary integration knowledge in the future; (3) strongly typed interfaces and process-centric integration allow for better testing and earlier error detection.
<table>
<thead>
<tr>
<th>Form of B2B Integration</th>
<th>E-mail communication</th>
<th>Portal</th>
<th>Message exchange based on individual data format</th>
<th>Message exchange based on industry standard</th>
<th>SOA-based process integration based on industry standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interaction Type</td>
<td>Human-human</td>
<td>Human-machine</td>
<td>Machine-machine</td>
<td>Machine-machine</td>
<td>Machine-machine</td>
</tr>
<tr>
<td>Relationship</td>
<td>1:1 (Dyadic relationship with unstructured information exchange)</td>
<td>1:n (Focal relationship where OEMs provide suppliers with Web-based access to their information systems)</td>
<td>1:1 (Dyadic relationship with bilateral agreement on message format)</td>
<td>n:m (multi-lateral agreement on message formats, here: XML or EDI documents)</td>
<td>m:n (Multi-lateral agreement on public constructs covering process and IS layer)</td>
</tr>
<tr>
<td>Adoption in practice</td>
<td>High – used in all B2B interactions in engineering change management</td>
<td>Middle – OEMs have established around 30-50 different supplier portals, ECM scenario is partly implemented</td>
<td>Low – few implementations</td>
<td>Low – few implementations, e.g. BMW and Magna Steyr / EDI</td>
<td>Low – pilot implementation by BMW and suppliers</td>
</tr>
</tbody>
</table>

### Coordination costs – savings in executing and monitoring transactions: $\Delta c_{\text{coord},i}$

<table>
<thead>
<tr>
<th>Savings through electronic processing</th>
<th>No difference</th>
<th>0.75 PD per engineering change (OEM only)</th>
<th>0.75 PD per engineering change (both parties)</th>
<th>0.75 PD per engineering change (both parties)</th>
<th>0.75 PD per engineering change (both parties)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>$\Delta c_{\text{coord},i} = 0$</td>
<td>$\Delta c_{\text{coord},i} = 0.75 * n * t_n$</td>
<td>$\Delta c_{\text{coord},i} = 0.75 * n * t_n$</td>
<td>$\Delta c_{\text{coord},i} = 0.75 * n * t_n$</td>
<td>$\Delta c_{\text{coord},i} = 0.75 * n * t_n$</td>
</tr>
</tbody>
</table>

### Connectivity costs - inter-organizational agreement on the process interaction: $c_{\text{agr},i}$

<table>
<thead>
<tr>
<th>Costs for analyzing engineering partnership</th>
<th>N/A</th>
<th>N/A</th>
<th>5 PD per partner</th>
<th>5 PD per partner</th>
<th>5 PD per partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Costs for defining inter-organizational process interaction</td>
<td>N/A</td>
<td>N/A</td>
<td>10 PD per partner</td>
<td>10 PD per partner</td>
<td>0 PD (due to pre-defined public constructs)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>$c_{\text{agr},i} = 0$</td>
<td>$c_{\text{agr},i} = 0$</td>
<td>$c_{\text{agr},i} = \sum_{j=1}^{n} 15 = n * 15$</td>
<td>$c_{\text{agr},i} = \sum_{j=1}^{n} 15 = n * 15$</td>
<td>$c_{\text{agr},i} = \sum_{j=1}^{n} 5 = n * 5$</td>
</tr>
</tbody>
</table>

### Connectivity costs - internal implementation of the partner-specific process / organizational changes: $c_{\text{impl},i}$

<p>| Costs for specifying internal changes | N/A | 3 PD per OEM (suppliers only) | 3 PD per partner | 3 PD per partner | 3 PD (partner-independent) |</p>
<table>
<thead>
<tr>
<th>Partner-specific organizational changes</th>
<th>N/A</th>
<th>$c_{P,\text{change},ij}$ depending on internal processes (suppliers only)</th>
<th>$c_{P,\text{change},ij}$ depending on internal processes</th>
<th>$c_{P,\text{change},ij}$ depending on internal processes</th>
<th>$c_{P,\text{change},ij}$ (partner-independent) depending on internal processes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>$c_{P,\text{impl}ij} = 0$</td>
<td>$OEM: c_{P,\text{impl}ij} = 0$</td>
<td>$Supplier: c_{P,\text{impl}ij} = \sum_{j=1}^{n} (3 + c_{P,\text{change},ij})$</td>
<td>$c_{P,\text{impl}ij} = \sum_{j=1}^{n} (3 + c_{P,\text{change},ij})$</td>
</tr>
</tbody>
</table>

**Connectivity costs - inter-organizational agreement on the IS level: $c_{\text{IS,agr},i}$**

<table>
<thead>
<tr>
<th>Definition of the information that has to be exchanged</th>
<th>N/A</th>
<th>N/A</th>
<th>3 PD per partner</th>
<th>0 PD (due to pre-defined EDI/XML messages)</th>
<th>0 PD (due to pre-defined public constructs such as WSDL, …)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision on the data exchange format</td>
<td>N/A</td>
<td>N/A</td>
<td>1 PD per partner</td>
<td>0 PD (due to pre-defined EDI/XML messages)</td>
<td>0 PD (due to pre-defined public constructs such as WSDL, …)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>N/A</td>
<td>$c_{\text{IS,agr}i} = 0$</td>
<td>$c_{\text{IS,agr}i} = \sum_{j=1}^{n} 4 = n * 4$</td>
<td>$c_{\text{IS,agr}i} = 0$</td>
</tr>
</tbody>
</table>

**Connectivity costs - internal implementation of the IS interface and communication infrastructure: $c_{\text{IS,impl},i}$**

<table>
<thead>
<tr>
<th>Field mapping</th>
<th>N/A</th>
<th>N/A</th>
<th>1 PD per partner</th>
<th>1 PD in total</th>
<th>1 PD in total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface realization + communication infrastructure and backend integration</td>
<td>N/A</td>
<td>Only OEM: $c_{\text{IS,impl,change}}$, (setup of supplier portal, partner-independent, depending on internal application landscape)</td>
<td>10 PD per partner + $c_{P,\text{IS,change},ij}$ (depending on internal application landscape)</td>
<td>10 PD in total (adapter) + $c_{P,\text{IS,change},ij}$ (partner-independent, but depending on internal application landscape)</td>
<td>10 PD in total (service) + $c_{P,\text{IS,change},ij}$ (partner-independent, but depending on internal application landscape)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>$c_{\text{IS,impl}ij} = 0$</td>
<td>$OEM: c_{\text{IS,impl}ij} = c_{\text{IS,change}ij}$</td>
<td>$c_{\text{IS,impl}ij} = \sum_{j=1}^{n} (11 + c_{\text{IS,change}ij})$</td>
<td>$c_{\text{IS,impl}ij} = 11 + c_{\text{IS,change}ij}$</td>
<td>$c_{\text{IS,impl}ij} = 11 + c_{\text{IS,change}ij}$</td>
</tr>
</tbody>
</table>

**Legend:** PD – person days

*Table 3. Assessment of Five Different Forms of B2B Integration From a Transaction Cost Perspective*
6 CONCLUSION

The contribution of this research is two-fold: First, it derives a model for assessing different forms of B2B integration and their impact on transaction costs of collaborative B2B relationships. By distinguishing connectivity and coordination costs, it provides a systematic view on the cost and benefit structures of different forms of B2B integration. It thereby captures the inherent differences that exist between machine-to-machine and human-to-machine interaction. While existing IOS research focuses mostly on benefits, this paper draws the attention to the fact that different forms of B2B integration require different ex-ante investments in order to generate benefits. The second contribution of this research is the comparison of different forms of B2B integration. This study demonstrates that there is economic rationale for preferring e-mail communication and supplier portals to more complex forms of machine-to-machine integration. While machine-to-machine integration is associated with the highest savings in coordination costs, it comes with significant investments in electronic connectivity. This explains some of the phenomena revealed by current statistics on e-business adoption (European Commission 2007). Asymmetric investments and benefits may exist for the different parties in electronic B2B relationships, notably in the case of B2B portals. When comparing SOA and Web Services to prior forms of B2B integration, we find that service-oriented concepts potentially increase the interoperability and scalability of electronic B2B relationships. This is due to their (1) leveraging of open Internet standards, (2) elimination of the need for maintaining a separate B2B integration infrastructure and (3) ease of integrating with internal applications. However, IS implementation efforts only account for a smaller part of the overall investments in external process integration, and SOA does not alter the problem of implementing organizational changes. Another interesting finding relates to the role of standardization which has been extensively discussed by prior IOS literature (Damsgaard and Truex 2000; Reimers and Li 2005; Zhu et al. 2006). According to the automotive industry’s experiences, vertical standards reduce connectivity costs and relationship-specific investments, but do not completely eliminate them. Besides the need for analyzing bilateral relationships prior to setting up electronic linkages, this is due to the costs incurred for implementing internal process and system changes in order to connect with external partners. On the other hand, the scope and quality of vertical standards significantly affect connectivity costs. If standardization goes beyond message definition and specifies process-level agreements, it further reduces connectivity costs.

There are several limitations to our research. The most important limitation lies in the explorative approach and the limited empirical base for assessing the alternative forms of B2B integration. Although we chose an industry and a scenario which provided rich data related to different forms of B2B integration, there is a risk that our findings are only applicable to this particular scenario. Another limitation of our research is due to the lacking availability of data on B2B integration costs. Our analysis relies on estimations from pilot and productive implementations. Future work is necessary to gather more empirical data and to validate our findings in other B2B scenarios and industry settings. While this study demonstrates that transactions cost theory provides a very valuable framework for analyzing the different forms of B2B integration, the suggested model only presents a first step towards measuring transaction costs. We hope that our findings stimulate further research and encourage further studies that elaborate on the distinction between connectivity and coordination costs. Since the different forms of B2B integration are still understudied, more conceptual and empirical research is needed to improve our understanding of B2B integration variants and their impact.

References


E-DISPUTES AT THE CROSSROADS: A STAKEHOLDER ANALYSIS OF ON-LINE DISPUTE RESOLUTION MECHANISMS (ODR)

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0747.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Legacy applications, Legal aspects of IT, Online shopping, Regulation / Deregulation</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
E-DISPUTES AT THE CROSSROADS: A STAKEHOLDER ANALYSIS OF ON-LINE DISPUTE RESOLUTION MECHANISMS (ODR)

Sidiropoulou, Panagiota-Aikaterina, Middlesex University, Law Department, The Burroughs, Hendon, NW4 4BT London, UK, a.sidiropoulou@mdx.ac.uk

Moustakas, Evangelos, Middlesex University, Marketing and Enterprise Department, The Burroughs, Hendon, NW4 4BT London, UK, e.moustakas@mdx.ac.uk

Abstract

The growth in e-commerce has been accompanied by a large number of disputes related to the online commercial interactions causing high litigation expenses and outstanding delays during the decision process. The distinct of such disputes in e-commerce have also led to the growth of online dispute resolution mechanisms, popularly known as ODR. In this paper, we provide an exploratory understanding and conceptualization of ODR. Based on the critical characteristics of ODR, we identify the various types of on-line transactions where ODR could be used and highlight its positive contribution in tackling these problems. Furthermore, we identify the key stakeholders as well as the secondary players in the ODR process and enunciate the roles played by them in order to move towards the development of an effective international on-line dispute resolution framework.

Keywords: On-line Dispute Resolution, ODR, e-negotiations, stakeholder analysis

1 INTRODUCTION

Since the advent of internet as a medium for online commerce, it has emerged as a critical medium for firms to interact and indulge in commercial transactions with their customers and business partners. Yet, the enormous growth in e-commerce and online transactions across the globe has been accompanied by a fallout arising from the number of disputes related to the online commercial interactions. The disputes have arisen for various reasons, but primarily due to the nature of e-commerce as an asynchronous mode for online transactions. The distinct of such disputes in e-commerce have also led to the growth of online dispute resolution mechanisms. A modern e-alternative to traditional lawsuits is the On-line Dispute Resolution (ODR) – a collection of online, out-of-court techniques, that in recent years has become one of the key areas in the legal world (Katsh, 1995, 2007).

The origins of ODR can be traced back in 1996, with the making of ‘Virtual Magistrate’ project (ICC 2004), due to issues that arose regarding Internet Service Providers’ rights and responsibilities. This was the first on-line arbitration system which was supported by the Chicago Kent Law School (Reed, 2007) and offered its free services to on-line users regarding defamatory postings and corrupted data files. Over a period of time, ODR has become popular among public and got termed as a ‘legal or business fad’ (Kaufmann et al., 2004). ODR’s basic purpose is to make justice more accessible, by improving the flawed parts of the justice process. In assessing the effectiveness of ODR, the Office for Developed and Transition Economies (Consumers International, 2000) identified that consumers were not using various ODR systems and there were recommendations for its avoidance.

Trust is recognized as an important, yet obscure concept in on-line transactions; however it has become more needed for effective online dispute resolution. Most of the disclosure policies and
agreements in online sites are unclear while several privacy infringements have also become commonplace (Schultz, 2004). Disputing parties when involved in dispute resolution complain about the level of protection of personal data, business secrets, contracts and their communication messages (Hornle, 2002). Moreover, there is an insufficient legal framework. On-line dispute resolution legislation or even an accepted cross-border dispute resolution regulation are absent adding to the complexities (Katsh and Rifkin, 2001). Government’s oversight and the need for ODR regulation are still emergent. In addition, the rapid evolution of on-line dispute resolution scheme will always demand immediate technological improvements and the development of software, programs and techniques, which remain unknown to legal practitioners.

This study is an exploratory effort to examine ODR mechanisms that exist in the current environment. More specifically, we perform a stakeholders’ analysis to examine multiple players, their interactions according to their actions and roles, in order to move towards the development of an effective international on-line dispute resolution framework. The key players, also known as the ODR’s principal stakeholders, which are examined are claimants, consumer associations, technical experts, ODR service providers, accreditation entities and the Government.

The contribution of the current research paper is important, as it is the basis for improving the means of dispute resolution, in order to be competent to settle on-line disputes whenever arise from cross-border commercial transactions. The development of a harmonized framework will play a part in achieving a constructive dispute resolution mechanism which will support e-commerce transactions to get improved and become more reliable. Also, it should ascertain that the quality of ‘on-line justice’ in meeting the claimants’ expectations.

2 ON-LINE DISPUTE RESOLUTION

According to the Internet World Stats (2008) report, the number of Internet users internationally on June 30th, 2008 was 1,463,632,361 compared to the world population, which is estimated at 6,676,120,288. In addition, most of the legal information gathered, known as ‘discovery’, exist in electronic form which is an evidence that proves attorneys’ need to use technology, despite their lack of interest in using electronic means of communication. Statistics point out that over 98% of lawsuits typically get settled through negotiation and out-of-court settlements (Boyer, 2007). The restrictions in the conventional legal environment and the evolution of World Wide Web (WWW) encouraged the emergence of on-line dispute resolution mechanisms (Kaufmann-Kobler and Schultz, 2004). In 2003, the United Nations Economic Commission for Europe (UNECE) Forum on ODR explained that the “rationale behind the early ODR programs was to leverage the cost-efficiency, accessibility, detemporalization and depersonalization of on-line interaction while preserving all of the procedural advantages presumed to be perfected in existing forms of face-to-face dispute resolution” (Choi 2003 p.1).

The ODR is not incorporated in any statutes, laws or regulations, since no legal framework supports this concept yet. The ODR is known as the ‘offspring’ of the traditional Alternative Dispute Resolution (ADR), which is the universal and long-existing ‘alternative’ movement to judicial process (Hornle, 2002). In broad terms, it is the branch of dispute resolution that is primarily enacted on an online medium (Hornle, 2002). It deals with the management of on-line and/or off-line disputes, depending whether they arise from interactions in cyberspace or in the off-line world respectively. This means that except for the participation of disputants and the intermediary’s intervention (third party), the ODR system is supported by digital technology also called the ‘Fourth Party’ (Katsh at al., 2001). The digital technology works as an ‘electronic ally, partner, collaborator’ (Katsh, 2002), by improving human capabilities with advanced electronic infrastructure and providing ways so that “dispute resolution expertise can be of value on-line” (Brannigan, 2003). The EU project of the Joint Research Commissions on the specifications for eXtensible Mark-up Language (XML) confirmed that ODR tools are a proof of the significance of technology in the evolution of Alternative Dispute Resolution, since it is a matter of interactivity (Keane et al., 2004).
In the Executive Summary of Final Recommendations of the American Bar Association Task Force on e-Commerce and ADR, ODR was described as the set of ADR techniques (mediation, arbitration, negotiation and conciliation) when integrated with “the use of the internet, websites, e-mail communications” (American Bar Association, 2002). In addition, on-line arbitration and on-line mediation are considered as ODR tools (Katsh, 1995). Moreover, Kaufmann and Schultz (2004) refer to other ODR techniques, such as automated negotiation’ (or ‘blind-binding’ or ‘assisted negotiation’ or ‘enhanced negotiation’ or ‘technologically facilitated negotiation’), where no service is offered by a neutral party since a computer software automatically settles the disputants bids-demands, as well as the ‘cyber tribunals’ (or virtual/wired courts). Ethan Katsh (1995) adds the terms of ‘on-line settlement’ that deals with insurance and financial claims and that of ‘on-line resolution of consumer complaints’. Additionally, Katsh (2007, p.105) suggests that ODR is “a network application that illustrates how, as data is generated by interactions between humans and machines, boundaries between dispute prevention and dispute resolution, or between dispute resolution and standard setting, may become less clear”. The ODR concept can also be found with other alternative terms, such as Virtual ADR, internet Dispute Resolution (iDR), electronic Dispute Resolution (eDR), electronic ADR (e-ADR) or on-line ADR (oADR) (Haloush, 2007, Raines et al., 2007).

3 A TYPOLOGY OF ON-LINE DISPUTE RESOLUTION

Based on the definitions and characteristics of ODR identified from prior literature, ODR processes can be offered for the settlement of disputes between Businesses and Consumers (B2C), Consumers and Consumers (C2C), Government and Citizens (G2C) and between businesses (B2B), as presented in Table 1.

3.1 B2C

There are several occasions where problems appear in B2C transactions. These problems could be handled in a cost effective way for both parties. Consumers could receive the same support in the on-line environment with the ones in the traditional markets in order to be confident before, during and after the transaction. This can be achieved through the development of an on-dispute resolution mechanism. For most professionals, who are involved in developing e-commerce platforms, dispute resolution is one discrete part of the overall e-trust environment.

3.2 C2C

Consumers are using intensively C2C portals such as virtual auctions to buy, sell or exchange products and services such as books, CDs, tickets etc. Unfortunately, the number of complaints in C2C commercial activities has increased dramatically over the last years. Initially, ODR has been used at web-sites for auctions, a type of business which is particularly vulnerable to conflicts, due to the fact that this transaction environment leaves ‘alone’ buyers and sellers to agree on after sales details. It is almost impossible to identify who is on the other side of the transaction especially because repeated relationships between buyers and sellers are rare. As a result, it is hard to trust because of lack of personal contact.

3.3 G2C

Due to the extensive involvement of the Government to various activities the range of disputes and the number of complaints is overwhelming (Ramasastry, 2004). ODR can be used by Governmental agencies and support the general move towards a model where citizens can handle several tasks by themselves such as paying parking tickets, requesting tax returns or informing change of their address. The above infrastructure could be extended for ODR purposes and this could resolve a large number of cases electronically without requiring citizens to reach physically administration authorities. More
specifically, if there is a repeated problem with pensions of 10,000 elderly citizens, instead of every
citizen to wait in long queues to fill in a form or speak to a public administrator he can file his
complaint electronically. If the complaint matches the information in a database, the problem can be
handled automatically. ODR can enhance citizen satisfaction towards Government since it resolves
several cases on an effective and user friendly way.

3.4 B2B

Because of their higher value, Business to Business transactions could use ODR mechanisms even
more effectively than in the B2C environment. B2B transactions have also to deal with trust issues
because of problems related to language, shipping and currency. Companies could use the ODR
mechanisms to avoid litigation expenses. Buyers and sellers in B2B usually repeat their transactions
and this provides an incentive to generate outcomes more effectively by using ODR in order to
continue collaborating in the future. The fact that organisations have legal identity and reputation also
works in favour of ODR. Another important element in B2B transactions is the fact that due to their
nature, during online exchanges both parties are buyers and sellers. As a result there is no problem
about who is going to pay the mediator’s fee which means there is no risk of bias.

<table>
<thead>
<tr>
<th>Relation</th>
<th>Problems</th>
<th>The use of ODR</th>
</tr>
</thead>
</table>
| B2C      | - Undermine the credibility of on-line market space  
- Damages the corporate brand and overall reputation  
- Consumers as well as businesses can not afford legal expenses and lawyers’ fees | - Increases the consumer confidence on-line  
- Keeps customers satisfied  
- Reduces corporate liabilities  
- Removes the risk of failed transactions  
- Increases customer loyalty |
| C2C      | - This environment leaves ‘alone’ buyers and sellers to agree on after sales details.  
- Hard to identify who is on the other side of the transaction  
- Lack of trust | - Minimizes the anonymity between individuals who use on-line C2C platforms to buy and sell their products and services as well as tackles disagreements effectively  
- Removes the risk of fake transactions |
| G2C      | - Conflicts arise during the delivery of Governmental services to citizens. Disputes over tax collection cases, benefits payments etc.  
- Time consuming and bureaucratic procedures discourage citizens from addressing for solution of their problems  
- Due to the extensive involvement of the Government to various activities the range of disputes and the number of complaints is overwhelming | - Could resolves electronically and automatically a large number of cases without requiring citizens to reach physically administration authorities  
- Avoids bureaucratic procedures on a time effective way  
- Can enhance citizen satisfaction towards Government and keeps society functioning smoothly  
- Can assist older individuals or people with special needs |
| B2B      | - Have to deal with trust issues because of problems related to language, shipping and currency  
- Damages their legal identity and reputation | - ODR could be used very effectively in this environment due to the high value of purchases as well as because buyers and sellers repeat their transactions  
- Could avoid litigation expenses  
- Buyers and sellers in B2B usually repeat their transactions |

Table 1. Business relationships, their problems and the use of ODR
4 STAKEHOLDERS’ ANALYSIS

4.1 Major ODR Stakeholders

Stakeholder Analysis (SA) has been widely applied in strategic management and corporate governance (Burgoyne, 1994; Donaldson and Preston, 1995), as well in information systems studies. Following Donaldson and Preston (1995) and Mitchell et al. (1997), we extend stakeholder analysis to the context of ODR. Stakeholder analysis will contribute in understanding the potential relations that may be created between the different players in ODR. In addition, it is a way of identifying possible conflicts of interests and presenting clearly the expectations of stakeholders in relation to the problems and the objectives of the research as a whole. Particularly, stakeholder analysis is “the process of systematically gathering and analyzing qualitative information to determine whose interests should be taken into account when developing and implementing a project” (Schmeer, 1999).

Technically, stakeholders’ analysis is used, because it supports in structuring at the best way the project design, understanding the whole system, by identifying and assessing the different stakeholders’ perspectives, actions and goals of the key actors to the research problem. It also matches all the gathered information which can then be analyzed properly and all the necessary ones may generate a single and sound framework of presented findings. In advance, it promises the chance of accomplishing a successful project. Specifically, such method targets at developing a model, an integrated mechanism of alternative dispute resolution applied in on-line transactions and thus, enhancing the on-line development of such dispute resolution mechanism.

Prior to the analysis, the major stakeholders will be identified and classified. In 1708, the word ‘stakeholder’ denoted the bet or the deposit (Clarke, 2004). Nowadays, stakeholder’s terminology refers to “anyone significantly affecting or being affected by someone else’s decision-making activity” (Chevalier, 2001). A stakeholder can either be a group of people or individuals or even organisations and institutions, which depending on the nature of their position and status can “significantly influence or be important to the success of the project” (Social Development Department, 1995) or “share a common interest” (ASA 2008) or this generic term that covers “all those people who have interest in the wellbeing of an area” (IDEA, 2008).

When the stakeholder influences significantly the project, then he considers being a powerful stakeholder (Social Development Department, 1995) either because of the position he holds in relation to the others stakeholders and his power “to control what decisions are made, facilitate its implementation or exert influence which affects the project negatively” (ASA 2008). In the current research, in order a stakeholder to be regarded as important, its interests and needs should be a priority. As a result, the decision regarding which players should be included in the stakeholders’ analysis depends on their major contribution and interest in the outcome to the current research.

The following stakeholder analysis’ table contains the list of the key ODR stakeholders that participate in the current research. These are the claimants, the governmental bodies, the ODR Service Providers, the ODR clearinghouses, the accreditation bodies, the technical experts and the consumer associations. The row data allows showing stakeholders’ relations and interactions between each other, regarding their roles and impact of intervention on the outcome of research, depending on their respective set of actions and goals in relation to the research problems addressed.

The following table (Table 2) provides an overview of the On-line Dispute Resolution stakeholders. In the following sections the key stakeholders are identified and their potential responses and objectives are analysed in detail.
### Table 2. Stakeholders’ Analysis for ODR

<table>
<thead>
<tr>
<th>Stakeholder</th>
<th>Role</th>
<th>Actions</th>
<th>Objectives</th>
</tr>
</thead>
</table>
| **Claimants**           | Users / potential users of ODR            | - Seek out-of-court alternatives, cost and time effective solutions to tackle their disputes  
- Use ODR mechanisms to tackle their problems  
- Become familiar and aware of ODR mechanisms | - To strengthen consumer confidence in ODR mechanisms |
| **Government**          | - Legislates and enforces legislation     | - Produces legislation to safeguard commercial environment  
- Being responsible for implementation, enforcement, harmonization and awareness issues | - To progress effective enforcement of legislation nationally as well as internationally  
- To enhance harmonization in cross-border dispute settlements |
| **ODR Service Providers (ODRSP) or ODR Providers** | Facilitator of the dispute resolution | - Ensure processes (and systems) are capable to operate for an ODR process | - To provide assistance as a neutral third party in negotiating a mutually agreeable settlement to the parties  
- To be receptive to new advancements and adjust them to resolution processes by keeping pace with them |
| **ODR Clearinghouses**  |                                           | - Assist consumer and businesses in initiating electronic dispute resolution processes and provide necessary documentation  
- Recommend the best ODR provider  
- Provide a filtering service of complaints  
- Act as payment intermediary | - To control the quality and specificity of a mechanism for ODR providers  
- To act as central contact point internationally  
- To seek grant ODR providers access to larger markets  
- To assist consumers to overcome cross linguistic and cultural barriers  
- To develop a standard and multilingual on-line intake form for dispute submission (claimants and providers)  
- To develop a “good practice” framework where basic standards for ODR will be included in order to achieve global conformity among providers |
| **ODR Accreditation entities** |                                           |                                                                         |                                                                                                |

4.1.1 **Claimants**

According to the above table, one of the major stakeholders is the users of ODR. These could be claimants, such as individuals, business corporations and e-shops, or even the Government, that face a dispute with another party in the on-line or offline environment and ask for an e-alternative solution.
Claimants are requesting more efficient and creative negotiations instead of asking for compensations without achieving their goals. The central idea of all these mechanisms is to ‘shake the ego’ of the disputing parties and to discern the most appropriate solution for them. The desire of attaining a practical, speedy and cost-effective dispute resolution implies the use of ODR mechanism and requires claimants’ acquaintance and awareness of its function. What discourages them from trying this legal alternative relates to the absence of confidence in using such mechanism, no matter if things have improved since its emergence.

4.1.2 Government

The Government’s intervention is questionable under the concept of ODR, due to fact that the Internet is evolving so rapidly and it is not possible to keep pace with legislation. For that reason, self-regulation is preferable (Schultz, 2004). The e-disputes verify the existence of e-Government (eGov) or electronic governance, another equally important stakeholder that attracts business people, citizens, politicians and policy makers. In Governmental Departments, the ombudsman is the only offline intermediary who is used for administering maladministration (Local Government Ombudsman, 2008). When the Commission of European Communities referred to the role of e-Government (2003), it implied “the use of information and communication technologies in public administrations combined with organisational change and new skills in order to improve public services and democratic processes and strengthen support to public policies”. Leading e-Governmental bodies, in the USA, Australia, New Zealand, Singapore, Canada and the UK are responsible for implementing and enforcing laws, as well as providing funding for researching on ODR mechanism (UNPAN, 2002). Moreover, on-line Governments should accomplish the Government to Consumer (G2C), Government to Business (G2B) and Government to Government (G2G) interactions as well as equality of the digital divide, trust, privacy and security in administrative transactions with its citizens and businesses and the creation of broadband culture. The ODR procedures are controlled by Governmental initiatives. E-com trust’, ‘TRUSTe’, ‘Square Trade’ and ‘BBBO-n-line (Better Business Bureau On-line)’ are some of the initiatives which foster consumers’ believability and encourage them to trust private information. Their role is to check whether the corporations adhere to their privacy policies and consumer protection laws, thus approving their trust mark, their symbol to be attached to their business web-page. According to Consumers Affairs (2005), surveys like BBBO-n-line showed that such trust mark schemes are not even known by consumers. Governments’ presence is important in the progress of such research as it is the ‘ideal host’ (Rule, 2002). They can recommend ODR tools in the negotiation processes, bring the harmonization in cross-border dispute settlements and progress effective enforcement of ODR legislation internationally.

4.1.3 ODR Service Providers

Another key stakeholder is also the ODR Service Providers which consists of lawyers specialized in ODR. They can are either be in the form of non-Governmental or institutional bodies and offer e-dispute settlements. Conley’s (2005) survey presents the overall picture of ODR services and websites. Eighty-two out of one hundred and fifteen dispute resolution services are operational and with caseload over 1.5 million on-line and off-line disputes. Fifty-seven are distributed in North America, thirty-eight in Europe and sixteen in Asia. Some examples of non-Governmental bodies (Katsh, 2000) are Cybersettle (1996), Squaretrade (1999), Settlementonline (2007) and SmartSettle (1997). There are also institutional bodies, the London Court of International Arbitration (ICC, 2008), the American Arbitration Association (AAA, 2007) the World Intellectual Property Organization (WIPO, 1967) and the Chartered Institute of Arbitrators (CIArb, 1915). Their main aim is to assist in the “negotiation of a settlement mutually agreeable to the parties” (Department of Justice, 2009) as third-party neutrals-facilitators and make certain that every mechanism operates properly and is ready for the on-line resolution. The Internet Corporation for Assigned Names and Numbers (ICANN), a non-for-profit public-benefit corporation, approves only four providers, the Asian Domain Name Dispute Resolution Centre (ADNDRC), the National Arbitration Forum (NAF), World Intellectual...
Property Organization (WIPO) and the Czech Arbitration Court (CAC) (ICANN, 2008). For that reason, any e-arbitrator should improve his knowledge and skills, through the attendance of relative workshops and regular (on-line) training, which combines technology issues and on-line dispute resolution processes, in order to be capable to provide high quality and efficient resolutions.

4.1.4 Clearinghouses

The Clearinghouse is a stakeholder, who ensures that consumers have chosen the appropriate ODR provider. Though ODR clearinghouses are not yet well-developed, the following two bodies demonstrate cases on how a clearinghouse can make use of power over ODR providers: These are the ‘European Extra-Judicial Network’ (EEJ-Net) and the project of the International Chamber of Commerce (ICC) of a ‘Dispute Resolution Clearinghouse’. Clearinghouses have strategic role and can provide substantial contribution in the outcome of the above analysis, as they stand in a neuralgic position of the ODR mechanism, by being the (payment) intermediaries between the future claimants and the ODR providers. Not only they look for the most suitable e-arbitrator by launching electronic resolution for the interested disputing party, but also they check the filling of complaint forms and give the necessary documents and appropriate information when users consider whether to have on-line dispute settlement. What it must be achieved by the ODR clearinghouses is to accomplish acting as central contact point internationally. The clearinghouses will assist in handling internationally the linguistic and cultural problems that bother consumers. Furthermore, these accreditation independent entities can similarly control and verify the quality of ODR practitioners “on the basis of predefined and public criteria” (Schultz 2003, p.4). According to that, if practitioners are not in compliance, then these entities can prevent them from participating in the ODR arena. Besides, they can also support and assist policy makers through their practical experience. ICANN has a well-developed accreditation system the Uniform Domain Name Dispute Resolution Policy (UDRP) (ICANN, 2008). It is under consideration the possibility of establishing ODR appellate bodies, capable of reviewing decisions on-line, so as to exercise control over ODR resolvers by the end of the specific processes (Kaufmann-Kobler and Schultz, 2004).

4.2 Secondary ODR Players

ODR is also supported by technical experts and consumer associations, which are the secondary players, in order for the mechanisms to function properly (Table 3).

<table>
<thead>
<tr>
<th>Stakeholder</th>
<th>Role</th>
<th>Actions</th>
<th>Objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical Experts</td>
<td>ODR technical developers</td>
<td>-Develop compatible software and user friendly applications</td>
<td>-To develop a more advanced ODR infrastructure (higher security, faster connection, advanced interactivity and friendlier interface)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-To investigate possibilities of implementing artificial intelligence during dispute resolution process</td>
</tr>
<tr>
<td>Consumer Associations</td>
<td>-Information Providers</td>
<td>-Provide awareness raising-programmes to consumers</td>
<td>-To raise public awareness by informing consumers and claimants about ODR mechanisms and providing them with suggestions on how to tackle problems more effectively.</td>
</tr>
<tr>
<td></td>
<td>-Regulators</td>
<td>-Operate as reporting centres receiving complaints</td>
<td>-To develop a database with ODR cases which can act as an information portal for claimants</td>
</tr>
<tr>
<td></td>
<td>-Educators</td>
<td>-Provide guidance to consumers</td>
<td>-To represent consumers’ in international conferences and forums in ODR</td>
</tr>
</tbody>
</table>

Table 3 Secondary ODR players

Proceedings ECIS 2009
4.2.1 Technical Experts

An important post within ODR is held by the technical experts, who are responsible for developing and maintaining ODR infrastructure. A technical expert, “by virtue of training, education, profession, publication and experience, is believed to have special knowledge of a subject beyond the average person” (IT-Career-Coach.Com, 2007). His role is of great importance and concentration is given on their capabilities to advance the ODR infrastructure by developing more compatible software to the ones that already exist without altering or wiping them out. The main goal is that of building up a scheme for designing an advanced ODR infrastructure, which will offer higher security, faster connection, advanced interactivity and friendlier interface.

4.2.2 Consumer Associations

Consumer Associations’ contribution is noteworthy. Examples of associations are the Consumer International (CI), a federation of consumer groups (Consumers International, 2008), the Transatlantic Consumer Dialogue (TACD), a forum of sixty-five US and EU consumer organisations (2004) and the European Consumer’s Organization (BEUC) which comprises forty-one independent national consumer organisations from thirty European countries (2008). The above associations maintain electronic forums that act as reporting centres, and through that way they control by revealing feasible abuses, displaying irregularities and receiving complaints. They keep interested parties aware of what is happening in the marketplace by providing them with guidelines and consulting them for general issues that bothers them. Consequently, the consumer associations expect raising public awareness about ODR mechanisms and developing a database, as an information portal for claimants, with ODR cases. They can also act as a platform for consumers’ representation in international conferences and forums related to ODR.

5 INTERNATIONAL CO-OPERATION AMONG THE STAKEHOLDERS AND THE NEED FOR A UNITED APPROACH

International co-operation among the stakeholders of ODR is essential and could only be achieved through co-ordinated action. The Governmental bodies need to continue to participate and actively contribute to enhance the learning experience for the ODR players. Clearly one of the biggest problems with the use of ODR is its international character since several disputes may arise from distant locations, a fact that leads to the conclusion that the need for co-operation by ODR experts is essential. This will mean that authorities in Europe as well as internationally should work together and join training initiatives to progress International on-line ODR solutions. This will encourage communication and co-ordination between ODR Providers and technical experts to achieve efficient and effective ODR infrastructures as well as to investigate techniques and ways to address obstacles during an on-line dispute resolution case.

The key components of ODR could fall into three major categories; people, process and technology. Successful ODR sites should consider seriously about these three elements if they want to be successful. More specifically, people develop and power web-sites as well as ODR. No matter how advanced is the technology behind the site, it is important to consider that individuals will finally support and use it. ODR practitioners should develop specific skills before they start handling difficult disputes on-line. The arbitrator should learn how to manage the ODR platform in order to be able to respond and facilitate an on-line arbitration. In addition, he should be able to run through in depth simulations in the ODR platform before the system is ready for a live case. There is a significant value in educating everyone about the tools and techniques of ODR. More specifically, practitioners who are new in the field should be highly encouraged to interact and mentored by experienced neutrals.
regarding what type of language is more appropriate to be used in opening statements and ground rules with other neutrals.

There is no single solution or silver bullet that can be sufficient enough by itself to progress ODR. As a result, all stakeholders should be involved in the effort of developing ODR mechanisms which will be cost effective and easy to use. Even if the ODR technical infrastructure was effective, technology itself is not sufficient to develop an idealistic ODR mechanism. Different stakeholders (ODR Service Providers, Clearing Houses etc) need to co-operate in order to find an integrated solution to progress ODR mechanisms.

6 CONCLUSIONS

The rapid development of e-commerce has generated large numbers of online disputes between parties. These disputes undermine the credibility of online markets discouraging customers and organizations to engage in on-line transactions. The On-line Dispute Resolution mechanisms build trust in transactions, preserve relationships and reduce corporate liabilities. The development of technology plays a significant role for a successful On-line Dispute Resolution mechanism. However, the technology is the medium to an end and not an end in and of itself. There are some arbitrators who are highly skilled and efficient in face to face dispute resolution and they are not equally able to work in an on-line environment. In contrast, there are skillful arbitrators who will emerge as great ODR practitioners. ODR makes customers feel that they can purchase products and services on-line with confidence. In the near future adding links of external ODR services could be an industry standard among major e-commerce web-sites and will act as a means of giving on-line customers confidence. It is crucial for e-commerce to develop a trustful environment to support the building of confidence between parties. On-line Dispute Resolution is an essential component to achieve the above objective. Once consumers and organizations are aware that they could handle their cross differences which arise from their cross-border transactions in a cost effective way, then they will be willing to engage further.

References

ASA (2008) Ethical dilemmas in professional practice in anthropology
   28/10/2008)
Brannigan, C. 2003 On-line Dispute Resolution: Resolving in Cyberspace <http://www.mediate.ca/on-
   lineresolution.htm> (Accessed 3/2/2008)
   Methods in Organizational Research: A Practical Guide, Sage Publications, New Delhi, pp. 187-
   207
   Justice Press, US
Chevalier, J. 2001 Stakeholder Analysis and Natural Resource Management
Di Maio, A. 2006 E-government strategies Highlight key focus areas, Gartner Inc. Stamford.


Local Government Ombudsman (2008) Complained to the Council?


Moustakas, E., Ranganathan, C., Duquenoy, P. 2006 “Email marketing at the crossroads: A stakeholder analysis of unsolicited commercial email (spam).” Internet Research, Vol. 16(1), 38-52.


Rule, C. 2002 Online Dispute Resolution for Business, San Francisco, Jossey-Bass


Schmeer, Kammi. 1999 Stakeholder Analysis Guidelines in Policy Toolkit for


DEVELOPING PUBLIC E-SERVICES FOR SEVERAL STAKEHOLDERS – A MULTIFACETED VIEW OF THE NEEDS FOR AN E-SERVICE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0182.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-government, E-Service, Stakeholder theory, Trust</td>
</tr>
</tbody>
</table>
DEVELOPING PUBLIC E-SERVICES FOR SEVERAL STAKEHOLDERS – A MULTIFACETED VIEW OF THE NEEDS FOR AN E-SERVICE

Axelsson, Karin, Linköping University, Department of Management and Engineering, Information Systems, SE-581 83 Linköping, Sweden & Swedish Business School at Örebro University, Informatics, SE-701 82 Örebro, Sweden, karin.axelsson@liu.se

Melin, Ulf, Linköping University, Department of Management and Engineering, Information Systems, SE-581 83 Linköping, Sweden, ulf.melin@liu.se

Lindgren, Ida, Linköping University, Department of Management and Engineering, Information Systems, SE-581 83 Linköping, Sweden, ida.lindgren@liu.se

Abstract

This paper discusses how several stakeholders in a public e-service development project consider a future e-service and expected changes in administrative processes and working routines. Our findings indicate a much more multifaceted view than the common win-win situation, with increased quality for citizens and increased efficiency for agencies, which is rhetorically put forth as an effect of public e-service implementation. We have studied a development project resulting in an e-service for handling student anonymity when marking written exams in higher education. In this case we have identified five distinct stakeholder groups related to this e-service: students, teachers, course administrators, exam guards, and the university at an agency level. All of them having certain expectations and fears about the new situation. By presenting this diversity in opinions we add further understanding to the notion of e-services as being more or less beneficial for certain stakeholders.

Keywords: Public e-services, e-government, stakeholders
1 INTRODUCTION

Public e-services demand secure information handling in order to build and maintain citizens’ confidence in public administration and in the long run also in democracy. We often see this issue being discussed in technological terms; e.g., how to develop secure IT solutions and how to handle integrity in public e-services (e.g., Argyrakis et al., 2003). In this paper we discuss how different actors apprehend an e-service under development, focusing on their perceived need for this e-service. An e-service, as well as any kind of service, is by definition intended to serve someone (Kelly et al., 2002). 

In this context of public e-services this means that an e-service should fulfill the needs of the citizens that are supposed to use the e-service. This is in accordance with Kelly et al. (2002) who identify services as meeting someone’s relatively enduring need for something.

In many e-government studies there is a focus on the roles of a government agency officer and a citizen who interact through, e.g., an Internet-based public e-service. These two roles are often taken for granted; the agency offers a communication medium to citizens who act as private persons towards the agency. In this paper, we use empirical findings from a case study in order to illustrate that this one-to-one relationship between government and citizen can be too simplified in order to explain the use of public e-services. Instead, we agree with Tranmüller and Wimmer (2000) who state that e-government involves many different stakeholder groups that need to be considered when developing e-services. In the present case we study the development of a public e-service for handling student anonymity during written exams at a Swedish university. The e-service is offered by the university, which has the agency role in this case. The e-service has several distinct user groups, both users working at the agency (teachers, administrators, and exam guards) and users acting in the role of citizens (students). This diversity in user roles implies that users of a public e-service are not a homogenous group. Instead, we argue that the stakeholder concept (e.g., Freeman, 1984), often used when describing and analysing private firms, can be fruitful to use in the e-government context as well.

The research question we focus on in this paper is how we can illustrate and understand different stakeholders’ perceptions of the public e-service during the development process. The perceptions can be expressed as anything from positive expectations to fears. We use empirical examples from the studied e-service development project in order to illustrate how stakeholder groups differ in their perceptions and consequently also in their feelings of relevance and need related to the e-service. By presenting this diversity in opinions we add further understanding to the notion of e-services as being more or less beneficial for certain stakeholders. By adopting a multifaceted perspective on stakeholders, public e-service development can be analyzed and understood in a way that takes several stakeholder groups’ needs into account. The purpose of the paper is thus to explore the effects of using the stakeholder concept when analyzing the development process of a public e-service. Our approach is supported by a study of Flak and Nordheim (2006) which indicates that few e-government studies so far have explicitly addressed the stakeholder complexity and its inherent challenges. However, we believe that the complexity is a potential strength in the development process that, handled in a constructive way, can result in more comprehensive and successful final e-services.

After this introduction, the paper is organized in the following way: In Section Two we discuss the theoretical concept of stakeholders in e-government. The research design is reported in Section Three. The empirical findings from our case study are presented in Section Four. In Section Five the findings are discussed. The paper is concluded in Section Six, in which we also make some statements about the need for further research efforts in this area.
2 STAKEHOLDERS IN E-GOVERNMENT

We use the stakeholder concept in order to discuss different user groups and other actor groups related to the studied public e-service. The stakeholder concept was used by Freeman (1984) in his seminal work with the definition of “any group or individual who can affect or is affected by the achievement of the organization’s objectives” (ibid., p. 46). The stakeholder concept was originally introduced and used in the context of a private firm. There are, however, several scholars who have discussed how the stakeholder concept can be applied to public contexts as well (e.g., Scholl, 2001; Pardo and Scholl, 2002; Chan et al., 2003; Flak and Rose, 2005; Flak and Nordheim, 2006). Scholl (2001) presents a literature review on how the stakeholder concept has been transferred from the private to the public context and used in e-government settings. He distinguishes both benefits and limitations, but concludes that even though the stakeholder theory origins from the private sector the stakeholder concept can be beneficial to use in e-government settings as well (ibid.). By dividing the actor roles into several stakeholders the understanding of citizen and government relations will broaden. Flak and Rose (2005) agree that stakeholder theory can be valuable to use in the e-government field, but they also argue that the stakeholder theory lacks theorization of the relationship between technology and stakeholders, which is important in order to understand e-government.

Mitchell et al. (1997) argue that a stakeholder possesses one or several of the attributes power, legitimacy and urgency. In their study, Mitchell et al. (ibid.) develop a typology of: 1) stakeholders who have power to influence the firm, 2) the legitimacy of the stakeholder’s relationship with the firm, and 3) urgency of the stakeholder’s claim on the firm. These three attributes are intertwined. Mitchell et al. (ibid., pp. 869-870) argue that “power gains authority through legitimacy, and it gains exercise through urgency [...] legitimacy gains rights through power and voice through urgency”. In an e-government context the attributes of stakeholders’ power, legitimacy, and urgency are also present. Chan et al. (2003) discuss how stakeholder theory can be used in order to manage stakeholder relations in e-government projects. They identify a lack of literature within the e-government field concerning strategic management of stakeholder relations, even though several scholars highlight this as an important issue to handle in order to reach success in e-government projects. This is also in line with Flak et al. (2003) who call for more research on how stakeholder theory can be adapted to the e-government field. This is taken a step further by Flak and Rose (2005) who propose a research agenda on stakeholder theory in e-government research, where for example issues such as external and internal stakeholders’ legitimacy as well as government agencies’ ethical duty to respect different stakeholders’ interests are put forth.

We do not intend to apply the entire stakeholder theory in the paper. Instead, we will explore the explanatory force of the stakeholder concept as such, by discussing and comparing the different stakeholders’ perceptions in our case. We follow Freeman’s (1984) definition above, and define a stakeholder as a person or group of persons who can affect or is affected by the public e-service and its surrounding processes. By discussing empirical findings related to different stakeholders the objective is that the analysis should become more comprehensive. This should help us view the studied phenomenon more critically, e.g., by contrasting different stakeholders’ experiences. Flak and Rose (2005) argue that applying stakeholder theory in e-government research could increase a critical stance.

Democracy depends in part on the trust in public institutions (Lauer, 2004). Therefore trust is an essential issue also in e-government research – an issue that can be related to and incorporated into the discussion and understanding of stakeholders. Several studies have explored the role of trust in the e-commerce area and the exploration in the area of e-government has just begun (Bélanger and Carter, 2008). Following Bélanger and Carter’s (ibid., p. 166) reasoning on the concept (based on Rotter, 1971) trust is defined as “an expectancy that the promise of an individual or group can be relied upon”. Social learning theory (ibid.) is the point of departure for the definition, which suggests that experiences of promised negative or positive reinforcements vary for different individuals and between different stakeholders. As a result, people develop expectancies that such reinforcements will occur when promised by other people. Komito (2005) claims that trust rarely is given unconditionally,
especially to governments composed of unknown and unaccountable individuals. Instead trust is earned. Trust has to be earned based on actual interactions that citizens have with particular agencies.

When we develop public e-services there is a risk that we put citizens’ trust in the e-service in the hands of the IT vendors (Lauer, 2004). On the other hand, taking different stakeholders into account during the development of new public e-services can increase the levels of trust and even solidarity and democracy. This means that there would be long-term benefits for civil society and also for, e.g., political participation (Komito, 2005). E-government has the potential to improve government transparency, responsiveness, and accountability, but will only be adopted if citizens deem them trustworthy (Bélanger and Carter, 2008). Trust as a concept has several different and interlinked dimensions, which we will only touch upon in this paper. One can for example discuss social trust, trust in administration and trust in technology. We use this division of trust in the paper in order to understand needs, opinions, etc. from different stakeholders. Bélanger and Carter (2008) provide a model where they divide trust into institution-based trust (e.g., trust in the Internet, such as secure data transmission, as an essential part of e-government) and trust in the government agency providing the e-service (which will highlight aspects such as organization and knowledge in the agency). Labels such as trust of the Internet and trust of the government are also present (ibid.). We label the two dimensions of trust as trust in administration (organization) and trust in technology when we analyze our empirical data.

RESEARCH DESIGN

In order to gain further understanding of how public e-services affect citizen and authority relations, we are conducting a research project in which we are analysing several public e-services from the perspective of trust, organization, and knowledge. The project is based on cooperation between researchers from the fields of information systems and public administration. The project aims to increase our joint understanding of how public e-services can be developed so that users (citizens as well as governmental users) have confidence in and competencies to handle these e-services, in a personal as well as an organizational context. Such e-services should also create democratic legitimacy and efficiency in the society.

2.1 The Case Study

In order to answer the question how we can illustrate and understand different stakeholders’ perceptions of public e-services during the development process, we analyze data from the first case study in the research project, which was performed at a Swedish university during the spring of 2008. The e-service under development is an e-service for handling student anonymity when marking written exams. At this university, 100,000 written exams are administered each year which makes this an extensive process. The e-service consists of several components; one part handling the information transfer from a student administrative IT system to a mobile palm solution that is used on site during the exam events, a web-based interface where students sign up for the exam and another web-based interface that the teachers and administrators use when reporting the results. The studied e-service comprises all these components and is, thus, used by several user groups during the exam process. The studied e-service differs from some public Internet e-services in the sense that it is closely integrated with the back-office IT system. We study all components in this paper.

The origin of the initiative to develop an e-service was student demands for a higher legal security in the marking process of written exams. Students argued that the teachers cannot be totally fair in their marks as long as they know who the student is. Students were afraid that some of them could be “punished” with a lower grade if they had been critical towards the teacher or that some of them would receive a higher grade than appropriate because the teacher liked them. Thus, the student demand for anonymity is in line with a general strive for equal opportunities in higher education; i.e., no one should be discriminated because of his or her sex, age, sexual orientation, ethnicity, religion or other faith, disability or social background. The student demand for anonymity was articulated through the students’ union and resulted in a strategic decision made by the university’s vice-chancellor that an e-
service should be developed to guarantee student anonymity during the marking process of all written exams. A project group was formed consisting of a project leader, systems developers, technical personnel, representatives of the exam guards and central examination administrators. A reference group was also organized consisting of representatives of the teachers, the students' union, and exam guards from all faculties. This implies that all identified stakeholders were represented in these two groups; students, teachers, course administrators, exam guards, and the university (represented by the project leader, systems developers, and technical personnel).

During the case study the authors followed the development project (the project group and the reference group) in their project activities. Data was generated in several different ways. One of the authors observed six project meetings and notes from these observations were taken. During the last project meeting respondent validation (Silverman, 2000) of the findings was accomplished. Data was also collected by observations of three information meetings open for university employees and in one systems training activity for exam guards. 14 representatives from all stakeholder groups have been interviewed during the case study; the project leader (1), course administrators (2), teachers (3), students (3), exam guards (2), and administrators at the students’ office (3). The interviews lasted for 30-60 minutes and were recorded. A focus group was also performed together with parts of the reference group (7 persons). This focus group meeting (approximately two hours) focused on the role of the reference group in the development project. Besides these data generation methods, project documentation as well as e-mails sent from university employees to the project group were also analyzed. Altogether this case study design (Yin, 1989) has resulted in rich empirical material focusing on the development project from several perspectives. The empirical data is of a qualitative nature and has been analyzed with an interpretive approach (e.g., Walsham, 2006). Of course, there are weaknesses in the chosen methods, for example that the presence of a researcher might influence the project meetings or that the choice of interview respondents could be biased. These are weaknesses of qualitative research approaches that always have to be considered. We argue, though, that our method and researcher triangulation (Miles and Huberman, 1994) is a way to confront and confine these weaknesses.

There are several motives for choosing the present case. First of all the development process was about to start when we got in contact with the university, which implied that we were able to follow the process from the start. We also have the possibility to return to the case for evaluations later on. The university also showed great interest and engagement in the research project which gave us valuable access to the case and opportunities to make a critical analysis of the situation. There is also a novelty interest concerning this kind of e-service since this seems to be the first one in Sweden with a direct link to the national IT system handling all information about students’ passed courses and exams (called the Ladok-system). There is a recent trend in Sweden to develop solutions for guaranteeing student anonymity in written exams, so other universities will probably follow this attempt.

2.2 The Process of Anonymous Written Exams

In this section we give a summarised description of the process of anonymous written exams. This includes both the e-service and the administrative routines surrounding the e-service. 1) A couple of times a year each course administrator registers all planned written exams within his or her responsibility area for next term in an IT system called TAL. 2) Based on the information in the TAL-system, a special unit at the university called the Examination Service books rooms and hires exam guards for the planned examinations. 3) Prior to the examination the student signs up for the event in a special (already existing) student web portal. An anonymous ID (called AID) is then created in the TAL-system. The AID is only valid for the specific student at this specific examination, and it is not accessible for the student at this moment. 4) Just before the examination starts, the exam guard downloads information about all students who have signed up for the present examination from the TAL-system to his or her palm. 5) The student gets his or her student identity card scanned in the exam guard’s palm and a list of all present students is created. 6) Then the exam guard gives the AID to each student. The student must write this AID instead of name and social security number on each page he or she completes during the exam. 7) When the student leaves the examination room he or she
gets the student identity card scanned once again. 8) When the examination is finished the exam guard transfers the information in the palm to the TAL-system. 9) Within 24 hours the teacher can create a web-based marking protocol based on this information. 10) When the teacher has completed the marking process he or she or a course administrator will register the grades for each AID in the e-service (which in turn transfers the data to the Ladok-system). 11) Finally, the teacher can print a list of the results where the students’ names are visible; i.e., once the written exam is marked and the results have been transferred to the Ladok-system by the e-service the identity of the student is revealed.

3 DIFFERENT STAKEHOLDERS’ EXPECTATIONS

We have identified five stakeholder groups in the studied case; students, teachers, course administrators, exam guards, and the university. In reference to the traditional view on e-government, the university represents the government agency in the case description below; i.e., the university management as well as project management. Each group expresses their particular expectations regarding the e-service under development. Below we present expectations and fears of each stakeholder group.

3.1 Students

The origin of this public e-service was students’ fear of not being treated in a fair way by the teachers when written exams are marked. Students claimed that they sometimes felt they were punished (by being given lower grades) if they expressed a critical attitude towards the course or the teacher. The students said that they would find it easier to criticize a teacher during a course if they felt that there was no risk for being punished during the examination. The students also stressed fairness as an important feature. It was not just the risk of being marked lower than expected that was seen as a problem, but also to get a higher mark than deserved. The students also saw a risk of being discriminated based on ethnicity or sex. The students we interviewed stated that the e-service, which ensures anonymity during examination, would increase their feeling of being secure from that kind of discrimination. In the interviews they emphasized the importance of perceived security. They made an important distinction between actual and perceived security by saying that maybe the marking process is actually done in a fair way already, but you must also perceive it as legally trustworthy. The e-service was seen as a means to perceive security in the process. The only fear the students mentioned regarding the e-service was that the timetable would not be met and doubts about whether the e-service would function as expected from the beginning. Altogether this stakeholder group was predominantly positive towards the e-service and the re-designed examination process.

3.2 Teachers

Some interviewed teachers did agree with the students’ opinion that this e-service could result in increased legal security – not only for the students, but also for themselves. With the e-service the examination process will become traceable and some of the interviewed teachers saw the traceability as a protection against future discrimination charges. On the other hand, they also expressed a fear to be forced into a certain pedagogical frame if technology (the e-service) was allowed to govern the examination process. The teachers were also afraid that a standardized process would result in decreased freedom of action for each individual teacher regarding their pedagogical design of courses. Another opinion was that this e-service could be seen as a sign of mistrust from the university. The teachers who stated this felt that their competence as fair examiners was questioned by introducing this e-service. One teacher went as far as stating that he would stop giving written exams in his courses due to this re-designed procedure. He objected to the idea that student anonymity would be of any good. Instead, he argued that he could only do a fair examination if he knew who the student was, e.g., in order to observe improvements in cases when a student failed the examination several times. The interviewed teachers were also afraid that this re-designed process would increase their workload and result in technical errors that would lead to chaotic situations and decreased legal justice. On the other hand, we also interviewed teachers who did not see student anonymity as an important issue, but
were positive towards the e-service as a tool for more efficient handling of the examination process (particularly for examinations with a large number of students).

3.3 Course Administrators

The course administrators that we interviewed were the ones with least understanding of the reasons behind the decision to develop the e-service. They stated that there had not been any problems concerning the marking process of written exams. They defended the integrity of “their” teachers as being fair examiners that did not discriminate any students. The administrators did not understand why the examination process could not continue as before. This group had not heard anything about the demands for student anonymity from the students’ union. Depending on how the examination process was performed at the course administrator’s department prior to the e-service development, they expected either an increased or a decreased workload due to the e-service. Administrators who thought that the teachers at their department would not learn how to use the web-based interface for result reporting expected their workload to increase, as they assumed that they would have to do this task instead. At the departments with a huge amount of written exams (e.g., the department of mathematics) the administrators were afraid that the automated routine for result reporting would lead to redundancy among administrators. They feared that this in worst case could lead to unemployment. Administrators who work at the students’ office are responsible for handing out the exams to the students after the marking is done. This group of administrators expected the process to become more stressful as the re-designed process will result in another way of sorting the written exams at the students’ office. The sorting will be based on dates and codes instead of names. This fear made this group of administrators very negative towards the e-service and the changed routines.

3.4 Exam Guards

The exam guards are a stakeholder group that is contracted by the university and temporarily hired for each examination. This group mainly consists of retired women who want to earn some extra money by working a few hours each month. Their responsibility is to supervise the students during the examination event in order to prevent cheating or the use of prohibited aid. It is this stakeholder group that faces the largest changes in their working process due to the e-service. Their work today is totally paper-based and the re-designed process implies that they will use a palm as their main working device. This group expressed fears that they will not be able to learn the new process and how to use the new technology. The degree of IT maturity is low in this group although it differs across individuals. The exam guards were afraid that the re-designed process will lead to increased time pressure during the examination, as the registration of each student in the palm will take some time. Their greatest fear concerned how they are supposed to solve technical problems that might occur, when they are alone in the classroom with a lot of students eager to start their examination. They were not sure what kind of help they can get and from whom. The exam guards also mentioned a positive expectation as they hope to be able to influence the examination process when the e-service is implemented. For example, they hope that the re-designed process will make it easier for them to refuse students who lack a valid student identity card to take part in the examination. These students are not allowed to do the examination, but now they are often difficult to reject when they are begging to participate. In the future process the student identity card must be scanned in the palm in order to get the AID, which means that no students can be permitted to participate if they lack this card.

3.5 The University

The e-service development was initiated by a vice-chancellor decision made after severe pressure from the students’ union. The main motive for the development project, articulated by the university, was to achieve increased legal examination security for students and teachers. Students should not risk to be favoured or discriminated and teachers should not risk to be accused of discrimination. Another expectation that the university has articulated is that the examination process will be more...
standardized thanks to the e-service. Today, the administrative process regarding written exams differs between departments. There are, e.g., many special solutions to get yet another opportunity to pass the exam, to get bonus points, etc. This implies that the regulations surrounding examination are not totally comparable across the university. Such differences are negative, according to the university, and should be removed. There is also some prestige in fulfilling this project expressed by university representatives. It seems as if this university will be the first one in Sweden to develop a solution with automatic information transfer to the national Ladok-system. The fact that the university can assure student anonymity in written exams is seen as a strong argument in the marketing of this university’s courses and educations. Obviously, the university is eager to launch this e-service as quick as possible in order to achieve a so called first mover advantage.

4 DISCUSSION

4.1 Identified Stakeholders’ Perceptions of the E-service

When comparing the expressed expectations from the identified stakeholder groups in the previous section, we find many differences in how our respondents view the future e-service and the changed examination process. In table 1, below, we summarise the main characteristics of each stakeholder group. Our findings obviously indicate differences regarding the stakeholders’ general impression of the development project. We also find differences regarding the stakeholders’ need for the e-service; i.e., whether the e-service corresponds to any expressed problem or not. The stakeholders have had differing influence on the development process and they will also be object for more or less evident changes in their working routines related to the examination process. Finally, we see differences concerning the expressed expectations and fears depending on stakeholder group.

Table 1 Stakeholder groups’ main perceptions in relation to the e-service

<table>
<thead>
<tr>
<th>Stakeholder group</th>
<th>General impression</th>
<th>Need for the e-service</th>
<th>Influence on e-service development</th>
<th>Affected by process changes</th>
<th>Expressed feelings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students</td>
<td>Positive</td>
<td>Demand for student anonymity</td>
<td>Initiated the development, representatives in the reference group</td>
<td>To some extent, but no critical changes</td>
<td>Improved situation, perceived security</td>
</tr>
<tr>
<td>Teachers</td>
<td>Negative, neutral or positive</td>
<td>No expressed need</td>
<td>Representatives in the reference group</td>
<td>Some changes, expect increased workload and technical problems</td>
<td>Questioned competence, a sign of mistrust, decreased freedom of action, improved efficiency and legal justice</td>
</tr>
<tr>
<td>Administrators</td>
<td>Very negative</td>
<td>No expressed need</td>
<td>Representatives in the project and reference groups</td>
<td>Changes that can lead to either increased workload or redundancy</td>
<td>Defend teachers’ competence, unnecessary system change</td>
</tr>
<tr>
<td>Exam guards</td>
<td>Negative or positive</td>
<td>No expressed need</td>
<td>Representatives in the project and reference groups</td>
<td>Severe changes due to introduction of IT</td>
<td>Increased need of technical competence, improved authority towards students</td>
</tr>
<tr>
<td>The university</td>
<td>Very positive</td>
<td>Prestige, marketing argument, standardization, answer to students’ demands</td>
<td>Manages the e-service development (in-house) and owns the development project</td>
<td>Positive standardization of processes, improved legal justice</td>
<td>Focus on rules, standardization and student recruitment – not on internal stakeholders</td>
</tr>
</tbody>
</table>

University representatives in the project group express difficulties regarding how to inform about the project, the future e-service and the process changes. They have arranged several open meetings for all university employees in order to inform and discuss this, but very few people have participated. At many departments there has been a widespread reluctance and skepticism towards the project. Based on our empirical findings there are several conceivable reasons for this situation. First of all, there have been prior attempts to develop an e-service for student anonymity which have failed due to technical problems. There is, thus, a history of failure that dilutes the present project and works as a breeding ground for mistrust among employees. The fact that three major stakeholder groups; i.e., teachers, administrators and exam guards, do not experience any problem that the e-service is supposed to solve is another important factor. Changes that do not correspond to an experienced problem
seem to be harder to accept. In this case, the lack of problem understanding must be focused and discussed in order to gain acceptance among these stakeholder groups. This is especially important since these stakeholder groups expect rather extensive changes in their working routines. Some of these expectations are valid, but others are exaggerated by rumors and lack of information. Since this project is performed in a large public organization (the university has several thousands of employees) the different stakeholder groups are only represented by a few persons in the project group and the reference group. These representatives have been successful to a varying degree in fulfilling their anchoring and informing responsibilities. Most teachers, administrators and exam guards have not been involved or informed about the project before the e-service was to be launched.

When focusing on the positive and negative feelings towards the e-service and the re-designed process that our respondents express, we discover some interesting patterns. The students, who as a group have been the one demanding student anonymity from the very beginning, express mostly positive expectations. They stress that the e-service will result in increased trust in a fair examination process. They distinguish between actual and perceived security; i.e., the examination process might be trustworthy even today but, as long as they do not perceive it to be a hundred percent reliable, they do not trust it. This division between the concepts of actual and perceived is discussed by Oscarson (2007) in relation to information systems security. His conclusion confirms our finding as he states that a high level of actual information systems security is not enough if an actor does not perceive it to be high as well (ibid.). It is also worth noticing that the students discuss increased trust in the process. On the other hand, the students we interviewed seemed to take rather little notice of possible technical problems that the implementation of the e-service could result in. Thus, they seem to rely on technology from the beginning.

On the other hand, the three stakeholder groups which are most negative towards the e-service and process changes (teachers, administrators, and exam guards) all seem to be mistrusting the technology and their own competence in relation to handling the technology (i.e., the e-service in the mobile device as well as the Internet-based part of the e-service). The exam guards express this feeling most distinctly, which of course is natural in relation to this group’s characteristics (consisting of many retired women with little prior IT experience). The administrators fear that the technology will change the conditions for their work in a dramatic way. Either it will result in much more work to do, since the teachers will delegate new tasks to their administrator, or it will make some administrators superfluous since the process will be much more effective. Both these scenarios might come true, depending on how the departments handle the examination process today. This is a result of the standardization that the university intends to create through the e-service. This kind of uncertainty about future effects of organizational changes is inevitable in many cases and do cause worry among employees. Facts and information are, thus, always better than rumours and guessing during change processes.

Another serious finding is the statement from some teachers that they regard the e-service as a sign of mistrust in their competence as examiners. The administrators also defend the teachers’ competence and underpin this feeling of mistrust in competence. There seems to be a mix of general reluctance towards changes – “we have always examined without student anonymity and I don’t see the need for this to change now”, fear to deteriorate pedagogical ideas in the courses leading to decreased quality – “I don’t want to get forced into any pedagogical frame that does not fit my course”, and a feeling that the university wants to decrease teachers’ freedom of action – “don’t they have trust in me as a fair examiner anymore?”. What we see here are signs of changes in competence needs in different stakeholder groups. The exam guards must learn how to handle the mobile device, and the teachers and the administrators must learn to interact with the Internet-based interface. Some tasks will be conducted by another group than before, some new tasks will occur and others will disappear.

When focusing on feelings expressed by representatives for the university at an agency level, it is obvious that these experiences are mainly directed towards external issues. The main aim for developing the e-service was fulfilling the students’ demand for anonymity. The process changes necessary in order to develop the e-service also serve the purpose of reaching a more standardized
examination process at the university level. Standardization is seen as a means to assure legal justice
and high quality in the examination process. Another hope put forth is that the e-service and the re-
designed examination process will lead to a good reputation among presumptive students, which will
be positive for future student recruitment. All together, it is obvious that the university focuses on
external issues related to the e-service development, but seems to have underestimated some internal
issues such as different user groups’ experiences and fears.

Our findings provide a good example of how an aim of increased standardization, accomplished
through adjustment in processes as well as formalization of actions performed through an e-service,
has effects on different stakeholders’ perceptions and expectations. Students experience increased
confidence in a fair marking process and therefore trust for the administration (i.e. the university).
Internal user groups experience mistrust in technology, in their competence and in their employer’s
appreciation of their achievements. The university management experiences increased possibilities to
fulfill the rules of legal justice and hope for competition advantages.

4.2 Stakeholders in Public E-services

By dividing our respondents into several distinct stakeholder groups it becomes apparent that there are
differences between these groups’ apprehensions of the e-service under development. We find that
stakeholders can possess diverse characteristics in relation to each other and the studied e-service. If
we adopt the concepts of urgency, power, and legitimacy, introduced by Mitchell et al. (1997), on our
case we get the following picture: The students raised a demand for student anonymity – this demand
was urgent and as it was supported by the students’ union it was also powerful. The university decided
to meet this demand by initiating the e-service development project – the university obviously has the
power and legitimacy to do this. Since other universities were about to start similar projects this was
also seen as an urgent project to initiate. The teachers did not experience any urgency regarding this
matter and they felt their power as examiners being somewhat weakened by student anonymity as their
legitimacy as fair decision-makers could be questioned. The administrators had similar characteristics
as the teachers (low power, urgency and legitimacy), but they also feared more serious effects such as
redundancy or increased workload. The exam guards did not express any notions of high urgency
related to the e-service, but some exam guards mentioned that the e-service and the process might
strengthen their authority towards students. The e-service might, thus, give them the power to refuse
students without valid student identity cards to do the exam. The exam guards had rather low legitima-
cy in the examination process since they are only hired for a certain examination event and easy to
replace. Without using the full explanation power of Mitchell’s et al. (ibid.) typology, it is obvious that
the concepts of power, legitimacy and urgency can be fruitful to apply to e-government cases as well.

By conducting a stakeholder centered analysis of expectations and fears concerning the e-service
under development we get a more thorough view of the process. If we had adopted the traditional, and
in some sense simplified, view of a government agency and a citizen interacting through an e-service,
we would have focused on the university and the students. Both these actors are positive towards the
e-service and the changed examination process and focusing on them would, thus, not have given us
full understanding of the complexity of the situation. A more detailed analysis divided on all identified
stakeholders seems to be a suitable way to gain a broader and more critical picture of public e-service
development. This finding is also supported by Scholl’s (2001) study. Stakeholder analysis in
combination with identifying different objects for trust (technology and organization, cf. Bélanger and
Carter, 2008) is also promising in order to gain a more multifaceted view of e-services, but needs to be
further elaborated on in future studies.

There is a vital difference in this studied case compared to many other public e-services. Usually a
user can decide not to use a certain e-service. Government agencies cannot exclude other communica-
tion channels as they must assure that all citizens are able to interact with the agency regardless of
access to a certain medium or capability of using the medium due to other circumstances. This is not
the fact in the present case. The identified stakeholder groups cannot choose whether to use the e-
service or not – they have to adapt to this re-designed process and they have to use the developed IT device. The only stakeholder group that might have any potential choice is the teachers who possibly could choose to refrain from having written exams in their courses and, e.g., use other approaches, such as essays, other reports etc., in order to exam students.

5 CONCLUSIONS

This paper’s main contribution is lessons learned from applying the stakeholder concept in an e-government setting. Our intention has been to illustrate and understand different stakeholders’ perceptions of a public e-service during the development process. A crucial conclusion that can be drawn from this study is that an appropriate understanding of an e-service’s internal and external stakeholder groups is important in order to get an adequate view of the complexity often related to the use of a public e-service. There is often a stakeholder group that is supposed to be served by the e-service (c.f. Goldkuhl, 2007), but this group cannot be the only one in focus. The e-service is developed to generate public value (Grimsley and Meehan, 2007) to citizens, but given that several stakeholder groups are related to an e-service, the sometimes complex relations between these groups must also be understood. Chan et al. (2003) show how some stakeholders are dependent on an e-service while other stakeholders are necessary for the e-service. This results in complicated stakeholder relations that need to be identified, understood and managed in a proper way. Stakeholders also differ in their power, legitimacy and urgency towards the e-service (Mitchell et al., 1997). The focus on internal stakeholders without any explicit need for the e-service (i.e., no outspoken need to be served) provided us with a more nuanced and critical understanding of the studied process. A more traditional focus on the citizen (the student) and the government agency (the university) would instead have indicated a more simplified win-win situation. This is an important lesson to learn in order to understand the effects of e-government implementations.

We have identified several, sometimes conflicting, expectations concerning the studied e-service which seem to be possible to explain by a history of previous e-service failures, no correspondence between the e-service and experienced needs among several stakeholder groups, and lack of information and involvement in the development process. A superior goal for this e-service development initiative was to create a more secure and fair marking process; i.e., to improve the legal justice surrounding written exams. Obviously, the feeling of anonymity was apprehended by the students as an important factor – their trust in the administration should increase if they know that their exam would be treated anonymously. Trust in technology (the e-service), on the other hand, was discussed by other stakeholder groups but mainly in terms of mistrust.

This study provides us with some illustrations of how e-service development can create different expectations within the stakeholder groups. The intention has been to add further understanding to the discussion of different stakeholders in e-government. Increased understanding of this complexity can help us develop public e-services that balance different stakeholders’ needs in a successful way. Our intention has not been to give any statistically valid explanations of the studied phenomena. The characteristics we have found in this case (e.g., the notion of need for the e-service) have to be followed up and compared to other cases. We plan to return to the case when the e-service has been used for some time, in order to evaluate whether different stakeholder groups’ expectations and fears have been realized or not. This will add a longitudinal dimension to our analysis and to further results. Another important theme for further studies is trust (in technology and organization); in the analysis above we have not highlighted trust, only used the concept as one aspect of a stakeholder analysis.

Acknowledges

This study has been financially supported by the Swedish Civil Contingencies Agency and is performed within the SAFe project – a multidisciplinary research project at the Department of Management and Engineering at Linköping University in Sweden.
References


## INNOVATION IN VIRTUAL WORLDS: SOCIAL STRUCTURE AND DIFFUSION

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0408.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>Virtual world, Innovation Theory, Diffusion, Social Network Analysis</td>
</tr>
</tbody>
</table>
INNOVATION IN VIRTUAL WORLDS: SOCIAL STRUCTURE AND DIFFUSION

O Riordan, Niamh, University College Cork, Cork, Ireland, noriordan@bismail.ucc.ie
Adam, Frédéric, University College Cork, College Road, Cork, Ireland, fadam@afis.ucc.ie
O’ Reilly, Philip, University College Cork, College Road, Cork, Ireland, poreilly@afis.ucc.ie

Abstract

This research-in-progress paper investigates the impact of social structures in virtual worlds (VWs) on the diffusion of innovations within them. Research has illustrated that innovation diffusion takes place within virtual worlds. VWs facilitate the emergence of purposeful communities which support the (often real-world) activities of their members including the development and adoption of innovations. Thus, VWs alter the social structures in which their users are embedded and the manner in which they communicate, both of which are thought to influence the diffusion of technical innovations amongst individuals. Though technical innovation is at the heart of Information Systems (IS) research, empirical research investigating innovation in the networked age is in its infancy. Thus, this paper presents a framework of propositions in relation to the impact of social structures on the diffusion of innovations within VWs and proposes the use of social network analysis to investigate these propositions.

Keywords: Virtual worlds, innovation, diffusion, social network analysis

1 INTRODUCTION

The Internet is on the cusp of an evolutionary leap and today’s virtual worlds (VWs) are seen as a sign of things to come (Driver 2008). Despite multidisciplinary interest in them, research investigating VWs is in its infancy. A review of extant literature reveals considerable conceptual imprecision surrounding VWs (Boellstorff 2008; Fetscherin, Lattemann et al. 2008); several terms are used to describe them, including virtual environment, synthetic world, and Web 3D; but concise definitions are “hard to find” (Boellstorff 2008, p. 17). As a result, the task of understanding VWs and innovation processes within them is difficult. Nevertheless, the need to do so is clear: approximately nineteen million people now participate in VWs (Jackson and Favier 2008) and more than US$200m is spent annually on virtual goods (Castronova 2005). In Second Life (SL®) alone, residents spent more than 28 million cumulative hours inworld in February 2008, created 200 terabytes of digital goods, with 50,000 businesses claiming profits (CPA 2008). Well known examples of innovative projects in Second Life include the artificial ecosystem Svarga, the audiovisual sculptures at Ramonia and Studio Wikitecture.

This study focuses on the diffusion of innovations in virtual worlds. The importance of innovation for organizational competitiveness is well recognized (Chesbrough 2003) and articulated (e.g. Tushman and Anderson 1986; Henderson and Clark 1990). Innovation diffusion is concerned with the communication of an innovation amongst members of a social system (Rogers 2003, p. 6). The formation, development and operation of VW communities are
influenced by their spatial and graphical properties, functional capabilities, communications and social networking mechanisms. Social Network Analysis (SNA) is deemed an appropriate vehicle to investigate innovation diffusion in VWs. The application of network models represents a “promising avenue” for innovation research (Cowan 2005). Rogers (2003, p. 361) suggests that adopting relationships as the unit of analysis represents the first step in overcoming the individual blame bias within extant diffusion studies and calls for study of the “network turbocharger effects”, defined as “the additional variance… explained by network variables beyond the direct effect of the individual level variables”. SNA recognizes the embeddedness of actors in social systems and holds that “the structure of relations among actors and the location of individual actors in the network have important behavioural, perceptual, and attitudinal consequences” (Knoke and Kuklinski 1982, p. 13). It has already been successfully used to study innovation adoption and diffusion (e.g. Coleman, Katz et al. 1966; Abrahamson and Rosenkopf 1997).

The study makes several contributions to IS research. It contributes a VW definition, derived from a review and extension of existing literature. It addresses a phenomenon (i.e. virtual worlds) of multidisciplinary interest and cross domain potential including simulation and communication (The New Media Consortium 2007; Atlas 2008); education and research (Bailenson 2002; Dickey 2005; Boulos, Hetherington et al. 2007; Duffy 2008); marketing and commerce (Castronova 2005; Hemp 2006; Driver 2008). It contributes to theory by presenting seven propositions regarding the impact of social networks on the diffusion of innovations in VWs. Once tested, these will provide insight into VWs and the diffusion of innovations. The paper is structured as follows. VWs are introduced, characterized and defined. Innovation diffusion and social network analysis are discussed in relation to virtual worlds. Finally, the paper presents a framework of propositions and concludes with a brief overview of how this study is operationalized.

2 VIRTUAL WORLDS

This section provides an overview of VWs and derives a definition of them which utilizes and extends extant literature, revealing them to be shared, interactive, immersive environments where participants can communicate, collaborate, innovate and trade. Today, it is estimated that approximately nineteen million people participate in VWs (Jackson and Favier 2008) and that more than $200 million a year is spent on virtual goods (Castronova 2005). VWs have evolved over three decades (Bartle 2004; Fetscherin, Lattemann et al. 2008) so despite their apparent novelty, they have important histories (Boellstorff 2008). Ever influenced by virtual reality and gaming, virtual worlds have become more graphically sophisticated over time (Sivan 2008), and continue to incorporate media of increasing richness. The origins of VWs can be traced to the release of MUD (multi-user DUNGEN1) in 1978. The first mainstream social world (AberMUD) was released in 1989 and MOO (1990) introduced the first fully functioning scripting language into a socially oriented world (Bartle 2004). It is commonly held that VWs progressed from text based to graphical environments, but several graphical environments were written in the 1960s (for instance PLATO (1960s); Avatar (1979); Island of Kesmai (1981)). The first three dimensional virtual world (Meridian) was released in 1996. As the Web evolved into a more explicit architecture of participation, VW users leveraged the vivid “opportunities for communication, collaboration, and cooperation” afforded by VWs (Fetscherin, Lattemann et al. 2008, p. 232) to deliver user-generated content far beyond that envisaged in Web 2.0 narratives. There was released in 2003 with its own currency (Therebucks), purchasable with ‘real’ money (Brown and Bell 2004). Second Life, released in 2003, incorporated “Linden Dollars” and was

1 DUNGEN referred to the Fortran port used to play the game and is not actually a reference to the game Dungeons and Dragons (Bartle, 2004, p. 5)
billed as a “3D online digital world imaged, created, & owned by its residents” (Second life, 2007). The extent to which these mechanisms influence creativity in VWs is an open question.

Given the multidisciplinary interest in VWs (Fetscherin, Lattemann et al. 2008), the absence of clear definitions and an agreed upon terminology is to be expected. What follows is the articulation of a new definition of VWs as shared, interactive, immersive environments where participants can communicate, collaborate, innovate and trade. VWs are environments and are referred to as places (e.g. Curtis 1992; Bartle 2004; Boellstorff 2008); spaces; environments (e.g. Bartle 2004; Mennecke, Roche et al. 2007); and simulations (Bartle 2004) of real or imaginary environments (Hagsand 1996). Bartle (2004, p. 475) argues that VWs are “a set of locations… People go to places, do things there, and then they go home”. In other words, they are navigable (Hagsand 1996). In defining them as spaces, authors emphasize that these locations are not contiguously bound. Virtual world spaces can therefore include: (1) core public space (2) private space (private access but connected to public core) (3) private spaces (disconnected from public core) requiring teleportation (4) multi-scale spaces that break the rules of realism (e.g. a house inside a grandfather clock) (Dodge and Kitchin cited in Bartle 2004). VWs have their own “rules that enable players to effect changes to them”— in other words, they have their own “physics” (Bartle 2004, p. 3). A common misconception of VWs is that they are media: but as (Bartle 2004, p. 475) points out, a “medium is a channel open for communication with a (large) number of individuals. Although most VWs do contain channels, they are not themselves channels.

VWs are shared (Hagsand 1996), multi-user (Hagsand 1996; Bartle 2004), massively multiplayer (Mennecke, Roche et al. 2007) or distributed (Hagsand 1996). Their users are referred to as users (Curtis 1992; Hagsand 1996) but also as inhabitants (Bartle 2004), residents and participants (Mennecke, Roche et al. 2007). Implicit in the assertion that VWs are shared is the notion that the worlds exist independently of a particular user’s presence within them. In other words, the worlds persist over time. Authors, seeking to emphasize the manner in which virtual world interactions are mediated by avatars, sometimes refer to ‘characters or ‘players’ (Bartle 2004). This has led to the confusion of VWs and online games. VWs contain games. However, just as “the Pasadena Rose Bowl is a stadium, not a game” (Bartle 2004, pp. 473-475), VWs are environments distinguishable from the activities taking place within them.

VWs are interactive (Hagsand 1996; Cagnina and Poian 2007). Users are represented “inworld” as intentionally (and often lavishly) crafted characters known as ‘avatars’. Participants interact with each other (Cagnina and Poian 2007) and engage in “socialization, entertainment, education, and commerce” (Menneke, Roche et al 2007). They also interact with the environment itself (Hagsand 1996; Bartle 2004; Cagnina and Poian 2007; Mennecke, Roche et al. 2007); building objects and embedding new functionalities. This functionality has enabled a new departure in user-generated content and has evolved to varying degrees in different VWs. Increasingly; people interact with ‘bots’ which have been built by other people within virtual worlds. Interactivity contributes to what has been termed social presence, defined as the degree of salience of another person in an interaction and the consequent salience of the interpersonal relationship (Short, Williams et al. 1976). Thus, key facets of VWs include “interconnection… social relationship… [and] network externalities” (Cagnina and Poian 2007).

VWs are immersive, where immersion is defined as that sense of “being there” or more formally as “a psychological state in which the individual perceives himself or herself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli” (Bailenson 2002). In defining VWs as immersive environments, it is necessary to consider the relationship between VWs and virtual reality. Contemporary three dimensional VWs in particular, have been associated with the concept of Virtual Reality (VR) (Boellstorff 2008) and there is considerable overlap between descriptions of the two. For instance, VWs feature “immersion, interaction and imagination” (Zhu, Xiang et al. 2007) and McLellan (2003) defines
virtual reality as “interactive and evoking a feeling of immersion.” VR refers to (typically quite expensive) technological hardware, which is used to allow one to experience “worlds that never were and can never be” (Brooks in Biocca and Levy 1995, p. 6). VR technologies include (1) visual displays (2) graphics rendering systems (3) tracking systems and (4) database construction and maintenance systems (Brooks 1999, p. 16). The focus of VR research has traditionally been the technology itself and in increasing its “immersive and involving properties” rather than on the experience it creates (Steuer in Biocca and Levy 1995, p. 33). The immersion traditionally sought in the development of VR technologies is thus based on perceptual subterfuge or sensory realism leading to a “perceptually based illusion of non-mediation” (Lombard and Ditton 1997) within simulated environments. The immersiveness of these technologies has been defined as the “degree to which... [it] submerges the perceptual system of the user in computer-generated stimuli” (Biocca and Delaney in Biocca and Levy, p. 57). VWs are also immersive. But the technologies used are inexpensive and typically rely on desktop interfaces. The emphasis is less on the technologies used than on the worlds themselves (Bartle 2004; Boellstorff 2008, p. 5) and they “approximate aspects of reality – enough for the purposes of immersion” (Bartle 2004, p. 475). See Table 1 below for a list of prominent virtual worlds.

<table>
<thead>
<tr>
<th>NAME</th>
<th>RELEASED</th>
<th>SIGNIFICANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SecondLife</td>
<td>2003</td>
<td>Linden dollars; new revenue &amp; IP policies for user driven content</td>
</tr>
<tr>
<td>There</td>
<td>2003</td>
<td>Therebucks</td>
</tr>
<tr>
<td>EverQuest</td>
<td>1999</td>
<td>Roaming camera; became a “standard”</td>
</tr>
<tr>
<td>UltimaOnline</td>
<td>1997</td>
<td>Rich environment; Classless role playing</td>
</tr>
<tr>
<td>ActiveWorlds</td>
<td>1997</td>
<td>Influential for educational users</td>
</tr>
<tr>
<td>Meridian 59</td>
<td>1996</td>
<td>First (first person) 3D world</td>
</tr>
<tr>
<td>World of Warcraft</td>
<td>1994</td>
<td>Currently most popular MMORPG</td>
</tr>
<tr>
<td>MOO</td>
<td>1990</td>
<td>Scripting language; Morphed into LambdaMOO</td>
</tr>
<tr>
<td>LambdaMOO</td>
<td>1990</td>
<td>Most widely used distribution of MOO</td>
</tr>
<tr>
<td>Tiny MUD</td>
<td>1989</td>
<td>First social world; allowed user extensibility; became TinyMUSH</td>
</tr>
<tr>
<td>Tiny MUSH</td>
<td>c. 1989</td>
<td>Role playing emphasis; open source</td>
</tr>
<tr>
<td>LPMUD</td>
<td>1989</td>
<td>Enabled user functionality within worlds with a scripting language</td>
</tr>
<tr>
<td>DikuMUD</td>
<td>1991</td>
<td>Sophisticated codebase; Enabled user created worlds</td>
</tr>
<tr>
<td>AberMUD</td>
<td>1987</td>
<td>Successful itself; inspired DikuMUD, LPMud, TinyMUD</td>
</tr>
<tr>
<td>Federation</td>
<td>1988</td>
<td>Non-fantasy</td>
</tr>
<tr>
<td>Avatar</td>
<td>1979</td>
<td>Graphical</td>
</tr>
<tr>
<td>MUD / MUD2</td>
<td>1978</td>
<td>Multi-user; Inspired AberMUD</td>
</tr>
</tbody>
</table>

Table 1. Prominent virtual worlds listed in reverse chronological order

Second Life (SL) is seen as the de facto virtual world for commerce (Kim, Lyons et al. 2008). SL is the brainchild of Philip Rosedale and is managed by Linden Labs. SL is inhabited by millions of users and described as a user-generated content platform (Au 2008). Linden Labs describe it as “an online, 3D virtual world imagined and created by its residents” (Linden Research Inc 2008). Second Life is of particular interest because it embraces “strong economic and legal connections to the real world” in order to maximize the quality and quantity of its user-created content (Ondrejka 2005). Its key feature is the capability of its users to “collectively create” (Au 2008). Second Life residents have built replicas of real world locations, held “inworld” art openings, music concerts, and orchestral performances, and increasingly are using the medium to hold meetings, conferences, and teaching seminars (Jennings et al 2007) using a combination of built in text and voice chat features and streaming media. The infrastructure in Second Life is designed to facilitate a virtual economy, incorporating a Linden Dollar currency system and supports the ability to set objects as “for sale” or as replicable on demand (Jennings et al 2007). The steadily increasing volume of virtual goods and services sold in SL drove mainstream media coverage of SL and prompted the entry of a number of real world businesses into it in 2006 (Jennings et al
2007). Future plans to facilitate the development of its economy include the development of enterprise-safe VWs and universal avatars (which would allow users to “seamlessly travel” between worlds (IBM 2007); cf. Morgado 2009). Unresolved issues include the impact of virtual world economies on real world economies (see Castronova 2005); taxation, intellectual property rights and legal issues. The issue of digital property is also important because property rights are critical to strong markets, businesses, and innovation (Ondrejka 2005).

This study focuses on the potential impact of virtual worlds themselves on the diffusion of innovations within them. The role of virtual worlds in the creation and diffusion of innovations is of particular interest given the increased prevalence of networked innovation and the extent of collaborative and user driven innovation already manifest within them. Research is now needed to investigate the particular forms that innovation in virtual worlds takes and the processes at play as they diffuse through virtual world communities and social networks. To that end, the next section reviews literature on innovation and diffusion of innovation.

3 INNOVATION AND ITS DIFFUSION

This section provides an account of innovation and diffusion research. It introduces the concept of innovation, summarizes extant research streams, and identifies gaps in innovation studies to date. It argues for the adoption of a network perspective to investigate innovation in VWs.

Most broadly, innovation “combines factors in a new way” (Schumpeter 1939, pp. 87-88). It is commonly defined as an idea, practice, object or material artifact that is perceived as new by an individual or other unit of adoption” (e.g. West and Farr 1990; Damanpour and Gopalakrishnan 2001; Rogers 2003, p. 12). The idea is described as creative or meaningfully unique (e.g. Fang 2008); useful; original; or influential (Mayar cited in Paulus and Nijstad 2003). It generates value by solving techno-economic problems (Yayavaram and Ahuja 2008) and results in social change (Rogers 2003, p. xvi) or at least “challenges the present order” (Van de Ven, Angle et al. 2000).

Research has identified innovation types and characteristics (e.g. Zaltman 1973; Rogers 2003); and proposed innovation process models, at organizational (e.g. Utterback and Abernathy 1975), group (e.g. Osborn 1963; West in West and Farr 1990), and individual (Rogers 2003) levels. Authors have focused primarily on the factors affecting the rate of adoption of innovations (e.g. Moore and Benbasat 1991; Taylor and Todd 1995; Rogers 2003) and patterns of innovation diffusion (purposeful or passive) across space and/or time. Organizational level research dominates (King in West and Farr 1990, p. 52). However, results at an organizational level have been inconclusive and inconsistent (Wolfe 1994). Authors (e.g. West in West and Farr 1990; Paulus and Nijstad 2003) note the absence of group perspectives on innovation and creativity. Instead, emphasis is placed on individual reflection in creative accomplishments (Paulus and Nijstad 2003); individual antecedent research is typically cross-sectional; and few place facilitators or inhibitors in theoretical frameworks (King in West and Farr 1990).

The inattention to group level innovation research is understandable in light of phenomena including groupthink (premature consensus leading to suboptimal solutions) (cf. Janis 1972); diminished accountability (Karau and Williams 1993); diminished motivation (Karau and Williams 1993); deindividuation; production blocking; evaluation apprehension; and free riding in interacting groups (Diehl and Stroebe 1987). However, the role of groups in organizational life generally and innovation particularly (Paulus and Nijstad 2003) is increasingly acknowledged. In particular, evidence suggests that computer mediation enhances group performance. Dennis & Valacich (1993) find that large computer-based groups may outperform nominal groups (who tend to outperform groups) for brainstorming activities. Further, computer mediation may diminish the salience of social status variables resulting in greater and more even participation.
Thus, research bridging individual and group levels is needed. The application of network models to this task represents a “promising avenue” (Cowan 2005).

A number of criticisms have been levelled at innovation research. Innovation studies have done little to improve our understanding of innovation in decentralized contexts (Rogers 2003, p. 395). What we think of as single innovations are often the outcomes of lengthy processes involving interrelated innovations (Fagerberg, Mowery et al. 2004). Individual innovations do change (often drastically) over time (Kline and Rosenberg 1986, p. 283) as a result of reinvention, recombination (cf. Van de Ven, Angle et al. 2000; Yayavaram and Ahuja 2008), borrowing and imitation (cf. March and Simon 1958), functional interdependency (Rogers 2003, p. 162), and knowledge spillovers (Cohen and Levin 1989). But research often fails to address these issues (Rogers 2003). A focus on innovation diffusion patterns may shed light on these issues.

Innovation diffusion is defined as the process in which an innovation is communicated through certain channels over time among members of a social system (Rogers 2003, p. 5). Rogers (2003, p. xviii) argues that the diffusion of the Internet itself alters the diffusion of innovations because it changes the interpersonal networks through which information exchange about a new idea takes place. Ryan and Gross (1943) first classified adopters based on time of adoption: classifying adopters as innovators, early adopters, early majority, late majority and laggards. Innovation diffusion studies have investigated eight classes of dependent variable: the rate of adoption of innovations, the innovativeness of social system members (individuals or organizations), the role of opinion leadership, the role of communication channels, diffusion networks, the earliness of knowing about an innovation and the consequences of an innovation (Rogers 2003, pp. 98-99).

In the IS domain, the adoption of technical innovation has received attention resulting in several theories (cf. Davis 1989; Moore and Benbasat 1991; Venkatesh, Morris et al. 2003). These theories demonstrate the importance of adopter perception in adoption decisions, but are “for the most part silent on how users form initial attitudes about technologies” (Melone in Karahanna, Straub et al. 1999). The adoption of the relationship between individuals as the unit of analysis in future studies (cf. Rogers 2003, p. 361) may prove a first step in the understanding of this process. Rogers (2003, p. 361) also calls for the investigation of “network turbocharger effects”, defined as “the additional variance in a dependent variable explained by network variables beyond the direct effect of the individual level variables”. The application of network models represents a “promising avenue” for innovation research (Cowan 2005). Nonetheless, relatively few studies of innovation diffusion have adopted a social network analysis approach. Social network analysis (introduced below) represents a promising approach for the investigation of innovation diffusion in virtual worlds. The next section introduces social network analysis (SNA). It is a necessary prelude to the discussion of propositions which follows.

4 SOCIAL NETWORKS: THEORY AND ANALYSIS

Social networks are patterns of friendship, advice, communication or support which exist among members of a social system (Knoke and Kuklinski 1982; Scott 2000). A social network consists of finite set(s) of actors and relation(s) defined on them; the inclusion of relational information is a defining feature of a social network (Wasserman and Faust 1994, p. 20). Social Network Analysis (SNA) was developed to counteract the overemphasis of individual attributes in sociological research (Knoke and Kuklinski 1982). SNA holds that “the structure of relations among actors and the location of individual actors in the network have important behavioural, perceptual, and attitudinal consequences”; it explicitly recognizes the embeddedness of actors in social systems (Knoke and Kuklinski 1982, p. 13). SNA seeks to explain the behaviour of network elements in part by examining the interconnections amongst elements (Laumann 1994, p. 394). The hallmark of a network based explanation of a process is the inclusion of concepts
and information on relationships among units in a study: theoretical concepts and pertinent data are relational; structures are operationalized in terms of networks of linkages among units (Wasserman and Faust 1994, p. 6). Table 2 summarizes the social network constructs relevant to this study.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centrality</td>
<td><em>Point centrality:</em> local centrality (the relative prominence of a focal point in its neighbourhood) and global centrality (prominence in the context of the whole network). <em>Centralization</em> refers the overall cohesion or integration of the graph rather the relative prominence of points (Scott 2000). <em>Betweenness:</em> the extent to which a point lies “between” the various other points in a graph. It measures the extent to which an agent can act as a broker (Scott 2000, p. 86)</td>
</tr>
<tr>
<td>Density</td>
<td>Density refers to the general level of cohesion in a graph. Centralization describes the extent to which cohesion is organized around particular focal points. Density and Centralization are complementary measures (Scott 2000, p. 89). <em>Individual-level density</em> refers to the degree a respondent’s ties know one another/ proportion of ties among an individual’s nominees <em>Network density</em> is the proportion of ties in a network relative to the total number possible(networks may be sparse or dense)</td>
</tr>
<tr>
<td>Path length</td>
<td>The distances between pairs of nodes in the network</td>
</tr>
<tr>
<td>Social cohesion</td>
<td>A situation where individuals are connected directly to each other into sub-networks where either every actor is connected to all others (maximal connection) or there exists a social circle where each actor is in connection with at least 80% of the clique members (Burt, 1976). A cohesive group is one whose members are strongly attracted to one another where attraction refers to favourable sentiments toward others which find expression in an inclination to engage readily in social intercourse (Blau 1960)</td>
</tr>
<tr>
<td>Structural</td>
<td>Occupants of roles are structurally equivalent if they are to some extent interchangeable (Scott (2000, p. 123) uses the word substitutable), due to having similar linkages to the occupants of other positions (Lorrain and White 1971). Structural equivalence is concerned with the types of social relations maintained by categories of agents and is based on identifying uniformities of action which define social positions. equivalence</td>
</tr>
<tr>
<td>Contagion</td>
<td>Burt (1987) discusses social contagion in the context of innovation diffusion, suggesting that something about the social structural circumstances of ego and alter makes them proximate such that ego's evaluation of an innovation is sensitive to alter's adoption. Both contact and communication make ego and alter proximate (Burt 1987)</td>
</tr>
<tr>
<td>Structural</td>
<td>Static holes that can be strategically filled by connecting one or more links. The concept is related to social capital.</td>
</tr>
<tr>
<td>hole</td>
<td></td>
</tr>
<tr>
<td>Radiality</td>
<td>Degree to which an individual’s network reaches out into the network and provides novel information and influence</td>
</tr>
<tr>
<td>Reachability</td>
<td>The degree any member of a network can reach other members of the network</td>
</tr>
</tbody>
</table>

Table 2. Social Network constructs relevant to the diffusion of innovations in virtual worlds

The study of social networks has contributed to the development of social capital theories. Burt (1999) investigated the social capital of opinion leaders, whose role in the diffusion of innovations has been demonstrated by a number of empirical studies (cf. Rogers 2003). Burt (1992) examined the role of structural holes in social networks, postulating the existence of brokers who bridge those holes. Granovetter (1973) investigated the role of weak ties in the transmission of information across subgroups. Rogers (2003) developed the concept of localities and socialites and hypothesized that more cosmopolitan individuals would be comparatively more innovative. Lorrain and White’s (1971) investigation of the effect of social cohesion on individuals led to Burt’s (1987) analysis of social cohesion and equivalence in the context of social contagion.
In the next section, a number of propositions regarding the factors affecting innovation diffusion in virtual worlds are proposed and presented in the form of a theoretical framework.

5 PROPOSITIONS

This section derives seven propositions regarding the impact of virtual world social networks on the diffusion of innovations within them. These propositions are derived from a combination of virtual world observations (made in Second Life) and extant knowledge of social networks. The propositions are presented initially as a theoretical model (see figure 1 below); and pertain to three key aspects of the process and outcome of innovation diffusion, namely: (1) openness, (2) connectivity and (3) structural equivalence and social cohesion in virtual worlds. The seven propositions are organised under these three headings.

Figure 1. Preliminary model of the impact of VW social structure on the diffusion of innovations

5.1 Openness

People’s perceptions of each other shape their interactions. Indeed the weight they give each other’s opinions are often based on physical and social cues such as race, gender, age, or social standing (Weisband, Schneider et al. 1995). Previous research indicates that higher status group members dominate group activities but in computer mediated contexts, the salience of status effects is diminished, leading to greater participation (Weisband, Schneider et al. 1995). It is theorized that social status exists within VWs and is based on real world (where this information is disclosed) as well as virtual world factors (e.g. avatar appearance; newbie status).

However, the impact of social status on levels of interaction and participation is diminished in VWs leading to greater degrees of direct and equalized interactions.

Proposition 1: There are high levels of participation in the VW innovation diffusion process

Proposition 2: Easy access to detailed (even tacit) knowledge of innovations in VWs, due to direct exposure to opinion leaders and change agents results in rapid innovation diffusion

Computer mediated communication has typically been text-based and asynchronous with limited social presence (Wellman, Salaff et al. 1996). VWs facilitate high levels of presence, facilitating the formation of affective relations and personal contacts, which are thought to “provide more meaningful referents than broad social aggregates” (Gartrell 1987). Furthermore, as a result of high levels of participation in VWs, actors are more likely to have personal contact with high status members, including opinion leaders and change agents.
**Proposition 3:** High status actors, including opinion leaders, will have greater influence on individual adoption decisions in VWs than real worlds

### 5.2 Connectivity

Communication technologies within VWs are diverse and easy to employ. They support the ability to broadcast, locally or otherwise; multiplex communications (direct communications with multiple alters simultaneously using synchronous and asynchronous methods); and teleportation to spatially distant locations inworld. They therefore reduce the costs associated with maintaining a large number of relationships and may increase spatial as well as social proximity relative to the real world. This may facilitate the formation of large, (globally) dense social networks in VWs. Holes in social structure (referred to as structural holes) are thought to create competitive advantages for individuals whose networks span those holes (Burt 1992); to separate nonredundant sources of information; and to afford actors on either side of them the opportunity to increase their own social capital by brokering the flow of information across them (Burt 1999). Low costs associated with adding new linkages, greater network density and radiality as well as publicly available information on social structures in VWs decrease the likelihood that structural holes will be found in VW social networks and increase the ability of actors to circumnavigate them.

**Proposition 4:** Innovation diffusion in VWs is accelerated in VWS due to high availability of (novel) information and the relative ease of accessing it

- **Proposition 4(a):** VWs have high levels of network density and low levels of Betweenness
- **Proposition 4(b):** VW social networks have relatively short path lengths
- **Proposition 4(c):** VW actors demonstrate high levels of reachability and radiality
- **Proposition 4(d):** The existence and significance of structural holes in VWs is minimal
- **Proposition 4(e):** The existence of informational intermediaries in VWs is minimal and their ability to, and benefits associated with, controlling the flow of information in VW networks are reduced

### 5.3 Structural equivalence and social cohesion

Initial observation suggests that virtual world interactions are characterised by high levels of openness and connectivity. Empowered with social networking tools and freed from the spatial constraints of the actual world, users have greater control over the composition of their own social networks. Virtual worlds are theorized to allow structurally equivalent actors to come together to form purposeful and socially cohesive communities resulting in high levels of social contagion and homophily. Burt (1987) discusses social contagion in the context of innovation diffusion, suggesting that “something about the social structural circumstances of ego and alter makes them proximate such that ego’s evaluation of an innovation is sensitive to alter’s adoption”. Both contact and communication have been argued to make ego and alter proximate (Burt 1987). High levels of interaction lead to increased similarity (Gartrell 1987). Homophily is the principle that contact between similar people occurs at a higher rate than among dissimilar people (Rogers 2003). Prior research indicates that homophily increases the rate of innovation diffusion but inhibits the diffusion of innovations across subgroups (Rogers 2003).

**Proposition 5:** VW social networks feature high levels of innovation contagion

**Proposition 6:** VW social networks are homophilious increasing the rate of diffusion

**Proposition 7:** VW subgroup boundaries are permeable facilitating diffusion across subgroups
6 CONCLUSIONS AND FUTURE RESEARCH

The diffusion of innovations in VWs is of scientific and practical interest for several reasons. For many organisations, VWs represent significant commercial and educational opportunities. For these organizations, an understanding of how innovations are diffused within VWs is of paramount importance. In an academic context, the literature focusing on VWs is still immature. Investigation of the propositions outlined in this study will make a significant theoretical contribution in helping researchers understand diffusion of innovation within virtual worlds. Thus, insights gained from this study will be of value to researchers and practitioners alike.

The model (figure 1) is being operationalised using a mixed method approach. The first phase is underway. It involves an initial period of participant observation informing the specification of a survey questionnaire. Jorgensen (1989, p. 12) argues that participant observation is relevant where “the phenomenon is obscured from the view of outsiders”. The questionnaire is being pretested. The aim is to circulate the survey to members of Second Life’s educational community, using data collected by Jennings et al (2007). Survey data will be triangulated using qualitative data from a single case, gathered using both interviews and participant observation. The need to use a combination of methods arises from known (and significant) deviations between self-report and behavioural data in social network studies (cf. Bernard and Killworth 1977). Thus, this combination of methods is deemed appropriate to probe the above propositions and refine the model presented in figure 1.

References

Brown, B. and M. Bell (2004). CSCW at play: 'There' as a collaborative virtual environment. ACM conference on Computer supported cooperative work, Chicago, Illinois, ACM.


A FRAMEWORK FOR ANALYSING SERVICE ECOSYSTEM CAPABILITIES TO INNOVATE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0531.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Open Innovation, E-Service, Case Study, Web services</td>
</tr>
</tbody>
</table>
A FRAMEWORK FOR ANALYSING SERVICE ECOSYSTEM CAPABILITIES TO INNOVATE\(^1\)

Riedl, Christoph, Technische Universität München, Boltzmannstr. 3, 85748 Garching b. München, DE, riedlc@in.tum.de

Böhmann, Tilo, International Business School of Service Management, Hans-Henny-Jahnn-Weg 9, 22085 Hamburg, DE, boehmann@iss-hh.de

Leimeister, Jan Marco, Universität Kassel, Nora-Platiei-Str. 4, 34127 Kassel, DE, leimeister@uni-kassel.de

Krcmar, Helmut, Technische Universität München, Boltzmannstr. 3, 85748 Garching b. München, DE, krcmar@in.tum.de

Abstract

Electronic services delivered over the Internet are gaining importance in the business world. This area has seen an increase in scientific interest over the past years under the labels “Internet of Services” and Web-service ecosystems. The paper develops a conceptual framework of actors and their roles in an open innovation system for a networked ecosystem of Web-services. The framework illustrates how open innovation can be implemented in a Web-service ecosystem to increase innovation performance. Simultaneously this research closes a conceptual gap in current reasoning about Web-service ecosystems that neglects innovation processes. The utility of the framework is demonstrated by two case studies of Web-service ecosystems in which the framework was used to identify gaps in the implementation of open innovation processes. Our research results identify specific functions to support innovation processes.

Keywords: Open innovation, service science, service ecosystems, Internet of services, networked innovation, actors.

---

\(^1\) This research received funding from the German Federal Ministry of Economics and Technology (BMWi) under grant code 10MQ07024. The responsibility for the content of this publication lies with the authors.
1 INTRODUCTION

Service innovation is a primary concern for many businesses. The importance of innovation is increasing through shorter product life cycles, increased competition, changing customer behaviour, and technological progress (Leimeister/Glauner, 2008). Consequently, the management of new service development is a prime concern for companies in the service sector. Innovation is even more important in the area of Web-based services where barriers of entry are particularly low, services can be copied easily, and technological advances are especially rapid (Menor/Tatikonda/Sampson, 2002). Moreover, through the development of service oriented architectures (SOA) and Web-services the process of new service development changes. Through the development of SOAs individual service components become more fine-grained and can be re-used which affects the way new services are developed. This becomes especially apparent when looking at the vision of Web-service ecosystems or the “Internet of Services.” These visions include repositories of services that can be re-used, recombined, and re-purposed to create new, innovative services (Janiesch/Ruggaber/Sure, 2008; Riedl et al., 2009). The emergent Internet of services promises opportunities for new service development. Notably, such an environment provides a fertile ground for open innovation in which Web-service ecosystems act as catalysts. Open innovation proposes principles for the design of innovation systems in which innovation processes are open for external collaboration with a network of customers and suppliers (Ebner/Leimeister/Krcmar, 2010; Leimeister et al., 2009). It has been shown that implementing these principles increases innovation performance (Gassmann, 2006). The inter-organisational networks that are formed by Web-service ecosystems have many links with the idea of open innovation (cf., Vanhaverbeke/Cloodt, 2006). Consequently, Web-service ecosystems provide a promising environment for the implementation of these principles and thus maximising the benefit derived from open innovation. To accomplish this it is necessary to understand the actors involved in service innovation in Web-service ecosystems and how these parties can contribute to and benefit from an open innovation system. The paper develops a conceptual framework of actors and their roles in an open innovation system for Web-service ecosystems; the framework illustrates how open innovation can be implemented in a Web-service ecosystem to increase innovation performance. The framework also shows how the current conceptual thinking about Web-service ecosystems can be evolved to incorporate findings of open innovation research. With the development of this framework it is our objective to close a conceptual gap in current thinking about Web-service ecosystems that neglects innovation processes and one-sidedly focuses on processes for service delivery. The utility of the framework is demonstrated using two case studies of Web-service ecosystems. The framework allows the identification of both white spots in open innovation networks and specific improvements in the implementation of innovation processes. The rest of the paper is structured as follows. Section 2 reviews the theoretical background of Web-service ecosystems and open innovation. Section 3 presents the collaboration framework followed by the framework’s application to two case studies in Section 4. Section 5 concludes the paper with deriving practical implications for implementing open innovation processes in Web-service ecosystems.

2 RELATED WORK

2.1 Web-Service Ecosystems

Web-services have become extremely popular in recent years and the success of Web-service-centred business models such as Amazon.com, Google, and Salesforce.com demonstrate the real commercial success of these models. Building on their wide-spread use new composite services are created that span across business boundaries in order to implement end-to-end business processes. This phenomenon of a large collection of Web services has been described as a service ecosystem and a growing interest in academic research is emerging as a consequence (Barros/Dumas/Bruza, 2005;
Barros/Dumas, 2006; Riedl et al., 2008; Riedl et al., 2009; Sawatani, 2007; Wu/Chang, 2005. Although the terms used may differ, phenomena similar to service ecosystem have been researched in other areas, for example under the label “Service Value Network”, “Business Webs” (Steiner 2006; Tapscott et al., 2000), and “Internet of Services” (e.g., Dorn et al., 2007; Zhang/Chen/Zhou, 2005; Janiesch/Ruggaber/Sure, 2008; Schrot, 2007). The composability of existing services into new and innovative value added services that implement end-to-end processes is a central attribute of these ecosystems whereby services are provided and integrated by different actors of the ecosystem which leads to a division of supply and delivery (Barros/Dumas, 2006). Although the research on Internet of services and service ecosystems is just emerging, existing theories from inter-organisational systems can be drawn on to explain these phenomena. Inter-organisational relationships of business firms are complex phenomena and as such difficult to conceptualise. However, there is a broad consensus that these systems can be best approached by factoring in economic, socio-political, structural, and technological variables (Bensaou/Venkatraman, 1996, Cunningham/Tynan, 1993). Many of these frameworks are modelled on industrial supply processes, such as in the automotive and retail industries, which have now been extended with processes for services delivered over the Internet.

This work analyses the innovation activities from the perspective of the overall ecosystem rather than from the perspective of a single organisation. Barros and Dumas (2006) and later adaptations by Riedl et al. (2008, 2009) propose that the following five actors have stakes in service ecosystems:

- **Provider** - Services are offered by service providers. These organisations provide the service implementation and offer the service by publishing a service description.
- **User/Customer** - Users request and invoke the services provided by service providers. These may be other applications (or other service providers) or the actual end-user of a service.
- **Broker** - Service brokers bring service providers and service consumers closer together. They might also integrate a service with certain delivery functions such as payment and authentication or combine other providers’ services into a new offering.
- **Mediator** - Service mediators offer translations between different service formats and other routine functions to allow service brokers to concentrate on their core competencies by eliminating the need for additional technical transformations.
- **Specialist Intermediaries** - These are providers in the more technical sense as they offer services but distinguish themselves through the nature of the service they offer. Contrary to “normal” providers they do not offer services targeted at end-users but rather offer service delivery components that are used by other providers to create marketable services. Common examples for these kinds of services are payment, authentication, or monitoring services.

Another most obvious role, though not explicitly mentioned by Barros and Dumas is that of the platform provider who builds the overall platform on which the other actors operate. The role might include providing a computing infrastructure (such as Amazon’s Elastic Compute Cloud EC2) and a set of additional services such as a service registry. The main objective of the platform provider is the overall success of the entire platform.

### 2.2 Open Innovation

Open innovation is a phenomenon that is of increasing importance to both theory and practice (Chesbrough, 2003). Three core process archetypes in open innovation have been identified: the outside-in process, the inside-out process, and the coupled process (Gassmann/Enkel, 2004). The outside-in process enriches a company’s knowledge and innovation base through the integration of external knowledge sources, particularly the knowledge sources of customers and suppliers, to increase its innovativeness. The inside-out process exploits a company’s unused inventions in different markets and a managed trade of intellectual property, e.g., through licensing. The coupled process is a

---

A combination of both the outside-in and the inside-out processes intended to maximise the benefits of both approaches.

These three archetypes are achieved through various means of perspectives on opening the innovation including: (1) globalisation of innovation, (2) outsourcing of R&D, (3) early supplier integration, (4) user innovation, and (5) external commercialisation of innovations (Gassmann, 2006). These open innovation processes lead to interfirm cooperation and development of ecosystems of networked firms sharing technology and trading intellectual property (West/Vanhaverbeke/Chesbrough, 2006; Stathel et al., 2008). The successful impact of open innovation processes has been described in several publications (see Gassmann, 2006 for an overview).

We argue that an open innovation paradigm rather than a closed innovation paradigm is necessary for successful innovation development within service ecosystems. This is due to their heavy reliance on re-use, their reliance on new business models, and knowledge leveraging as services are implemented as software (Gassmann, 2006).

The focus of open innovation, however, is a single firm that thus tries to open its own innovation process (West/Vanhaverbeke/Chesbrough, 2006). Furthermore, it says little about what other actors are involved and how they interact and collaborate regarding innovation development (West/Lakhani, 2008). Web-service ecosystems can be seen as a catalyst for open innovation and thus offer an opportunity to extend the firm-centric concept of open innovation developed by Chesbrough and others (Chesbrough, 2006; Chesbrough et al., 2006, Gassmann, 2006; Ogawa/Piller, 2006) by proposing a platform-centred interpretation.

The main aspect of service ecosystems is that of a central platform that brings all actors together. Companies try to extract ideas for service innovation from this central platform and use these ideas to create new or improve existing services. So, instead of a single organisation following the open innovation paradigm, a larger pool of companies bound together through a central platform follows the open innovation paradigm (Figure 1).

![Figure 1](image_url)  

**Figure 1**  
Platform perspective of the open innovation paradigm.

To jointly develop new products and services in an innovation network, different activities need to be performed by different type of roles. These roles characterise the types of activities involved and the type of contribution that are required. For successful innovation projects this is important to understand as the roles define the capabilities that actors need to contribute (Nambisan/Sawhney, 2008). In a general concept of “Network-centric innovation” Nambisan and Sawhney (2008) propose three types of innovation players:

- **Architect** - Architects trigger and catalyse innovation. Furthermore, they envision and direct innovation and attend to the innovation network. Architects are the central members in an innovation network; they provide the initial momentum, and define key elements of the network and the innovations to be carried out.
**Adapter** - Adapters provide specialised knowledge or support services as well as infrastructure services. Nambisan and Sawhney call them adapters because they adapt to the direction given by the architect. Adapters may possess highly specialised knowledge and expertise to solve unique problems during the innovation development.

**Agent** - Agents act as mediators by liaising interactions, mediating knowledge transfer, and mediating innovation.

In a similar approach Steiner (2005) differentiates between two roles. A shaper as an entrepreneur in a central role offering a dominant design or standard, and a multitude of other organisations, called adapters, offer complementary products to that central design.

Tapscott/Ticoll/Lowry (2000) differentiate between the following five classes of network participants:

- Customers who not only receive value but also contribute value through co-creation.
- Context providers play a leading role through facilitating the interface between customers and the other network actors and lead the choreography and value realisation in the network.
- Content providers contribute the main goods, services, or information that constitutes the intrinsic form of value.
- Commerce service providers facilitate trading processes such as financial transaction management, security and privacy, logistics and delivery.
- Infrastructure providers provide the infrastructure on which the platform operates.

While most open innovation studies have focused on the firm level (West/Vanhaverbeke/Chesbrough, 2006) the three works summarised above took a first step at analysing open innovation on an inter-organisational level. However, the resulting roles vary and need to be further conceptualised.

### 3 COLLABORATION FRAMEWORK

In order to answer the questions, who are the actors involved in such an innovation ecosystems and what are their core competencies, this section first presents a consolidated view on the network roles and second an interaction model for innovation in service ecosystems.

From the description of the network roles the considerable overlap in core competencies and contributions that are expected from each role can be identified. Table 1 consolidates the roles proposed for both service ecosystems and networked innovation and groups them under four main paradigms according to their core competencies and their contribution towards the innovation space of service ecosystems. First, the customer judges the value created for her and has requirements for new services (Berkovich et al., 2009). Second, the platform provider pushes an innovation project forward in the role of a leading player and establishes the main environment for the service innovation. Third, service providers offer various support services and specialised knowledge and follow the driver within an innovation project. Lastly, the broker engages in brokering between the providers and customers and engages in transforming ideas within the innovation space without offering services on its own.

<table>
<thead>
<tr>
<th>Barros/Dumas 2006</th>
<th>Customer</th>
<th>Platform Provider</th>
<th>Service Provider</th>
<th>Broker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nambsan/Sawhney 2008</td>
<td>Customer</td>
<td>Provider</td>
<td>Mediator</td>
<td>Specialist Intermediary</td>
</tr>
<tr>
<td>Steiner 2005</td>
<td>Architect</td>
<td>Adapter</td>
<td></td>
<td>Agent</td>
</tr>
<tr>
<td>Tapscott/Ticoll/Lowry 2000</td>
<td>Customer</td>
<td>Context Provider</td>
<td>Content Provider</td>
<td>Commerce Service Provider</td>
</tr>
</tbody>
</table>

**Table 1** Consolidation of network roles.

According to their core competencies service ecosystem actors make different contributions to the innovation space. The innovation space represents possible service designs that may be reached (c.f., Millar/Demaid/Quintas, 1997). In a setting with a central platform, such as service ecosystems described above, the platform forms a collective innovation space that defines the boundaries of trans-organisational, or networked, innovation. We argue that the contributions of the actors to the
innovation space fall into three main areas: services, ideas for new services, and feedback related to service usage. This structure relates to studies of customer roles in product development where customer contributions have been classified as a source of ideas, as a co-creator through participation in product design and development, and testing and supporting products (Nambisan, 2002). Through the heavy reliance on re-using and re-purposing existing services, the variety of existing services strongly influences future service designs. The more services are available on the platform, the larger the innovation space of potential new services becomes. Thus, contributing a new service to the ecosystem may open completely new possibilities. Concrete service ideas or requirements are the most obvious source for service innovations as they directly imply possible design options. Finally, feedback from service users about existing services is a main source for incremental service innovations (Riedl et al., 2008). In addition to contributing to the innovation space, actors may also extract from and expand on knowledge from the innovation space to create new services. Brokers play a special role as they do not necessarily contribute new ideas but transform and refine already existing ideas in the innovation space (Hargadon/Sutton, 2000).

For our collaboration model we used the consolidated roles customer, platform provider, service provider, and broker presented above. Figure 2 shows the actors and their contributions to the innovation space as described above. Table 2 shows each actor’s relationship with the innovation space.

Figure 2 Interaction model for innovation in service ecosystems.

| Customer | Contribute | Customers contribute ideas for completely new services.  
          |            | Customers contribute refinements (e.g., in the form of comments and community evaluation).  
          |            | Contribute requirements and needs (e.g., via innovation communities or lead-user studies).  
          | Co-production | Customers may become providers by developing new services on their own through end-user development (e.g., user generated mash-ups, cf. Dörner et al., 2008).  
          |            | Through providing new services, customers become service providers themselves.  
          | Feedback | Explicit - Customers provide feedback regarding existing services through rating (e.g., five-star rating) or comments left through community tools provided by the platform.  
          |            | Implicit - Through actual service usage (e.g., if a service is used frequently users value the service which allows to derive ideas for service bundling; Riedl et al. 2008). In general, actual service usage indicates user preferences and willingness to pay.  
          | Platform provider | Contribute | Overall environment (i.e., platform APIs etc.)  
          |            | Platform providers contribute ideas and comments about ideas to the innovation space.  
          | Extract | Just like a regular service provider the platform provider extracts ideas to be implemented and new services to be offered. However, the platform provider has a different evaluation function in that it is focused on overall platform success. Hence, the platform provider is likely to fund ideas that benefit the entire platform even if not economically viable on their own.  
          | Service Provider |
Contribute Services to a service repository. These services can be used as building blocks for new services thus shortening time to market and easing implementation. Service providers contribute an idea for which they seek community evaluation or refinement. They may also contribute ideas as a form of requirement communication, thus requesting a new feature and playing the role of a customer. Service providers contribute new/improved service (which might be based on ideas submitted to the platform).

Extract Service providers extract ideas from the ecosystem for implementation. Every service provider rates every idea from its own perspective and decides which idea is valuable. An idea valuable for company \( A \) might not be valuable for company \( B \). This might be due to different business models or available resources. Thus, different actors will have very different views on the same set of ideas, each evaluating ideas according to its own standards.

Brokers

| Transform | Brokers engage in transforming and refining ideas. This translates to a set of four sub-tasks: Capture good ideas, keep ideas alive, imagine new uses for old ideas, and put promising concepts to the test (Hargadon/Sutton, 2000). |

Table 2 Actor relationships with the innovation space.

4 IMPROVING ECOSYSTEM INNOVATION – EVIDENCE OF TWO CASE STUDIES

Two case studies (Yin, 2003) were conducted to explore how the application of our collaboration framework can improve ecosystem innovation. From this case application we were able to identify gaps in the implementation of open innovation processes and propose management guidelines and implications for future tool support.

Since service ecosystems are a rather new phenomenon not many manifestations can be found in business life. Our rationale was to select cases that resemble early stages of service ecosystems. We chose two cases which show early manifestations of Web-service ecosystems. The first case was selected from a state eGovernment initiative, the second from the automobile industry. Both cases involve a network of actors, are concerned with offering electronic services delivered over the Web, and provide these services in an interconnected fashion, i.e., service offerings are interconnected and involve several actors for the service delivery. The case study data was gathered through face-to-face interviews with key informants: the manager for service integration in case 1 and the director for product management in case 2. The interviews were supplemented with publicly available data gathered through extensive desktop research following the methods proposed by Yin (2003) and Miles/Huberman (1994).

4.1 Case 1: The Need for Feedback and Brokerage

The eGovernment portal consists of a central service platform operated by a central government division. This platform is the single point of entry for various agencies offering services for the general public. One of these agencies, our case study partner, offers a search service which makes internal registry data available to the public. Until the start of the eGovernment initiative, registry data was only available on paper.

The portal provider plays the role of platform provider and is used by other departments (such as our case partner) to offer their raw services. A payment service is available which can be used by other agencies to design fee-based services.

The eGovernment portal was launched with an initial set of interfaces and the Web-service-based architecture. The central division providing the eGovernment portal plays a leading role in the service ecosystem by being the main architect. Agencies interested in offering services via the portal platform independently design and develop their services in internal project teams. The project team responsible for the search service was formed after deciding to offer a free interface to their existing database through the central portal. The free search service was chosen as a starting service to gather experience
in working together with the central portal for other, more complex services. Only limited interaction between our case partner and the central division took place during the development phase and this was limited to clarifying the technical interface. In case new requirements towards the portal emerged our case study partner would engage in talks and discuss if the required features could be added to the platform. Prior to launching the new service, the central division operating the portal performed a pre-launch test which included general functionality tests as well as performance stress tests to ensure that all services offered on their platform adhered to a certain quality standard and functioned properly. The search service of our case partner provides its own feedback form for users to comment or report errors in the service. A user community or an innovation community for actual or potential service providers is not offered by the central division.

In summary this case describes the innovation process from the point of view of a service provider offering a service on a central platform.

**Actors:**
- Central government division offering a portal platform open to other agencies to offer services.
- Government agencies such as our case partner, providing services delivered through the central platform.
- End-users using the services offered by individual agencies that are delivered through the portal.

**Case Discussion**

Several observations with potential for improvement can be made in this case. First, the role of a broker could not be observed in the ecosystem. The broker role would be vital in facilitating communication between the government agencies that could potentially offer services on the platform. Thus, valuable cross-fertilising between agencies and an exchange of knowledge and experiences gained through the services already offered could be achieved. Moreover, a broker could capture good ideas from within the agencies, keep ideas alive despite employee turnover, imagine new uses for old ideas, and sponsor and test drive promising concepts.

Second, a shared innovation space or innovation community, through with users, agencies, or the platform provider can communicate and exchange ideas could be offered. Apart from the community aspect, processes could be established to systematically involve service providers in improving the platform or systematically introducing new services on the platform. A central community or other central mechanism for collecting customer feedback could be offered. This mechanism could allow users to rate their satisfaction with the services they have used. This customer feedback could be used as a valuable source for improving and redesigning existing services. The single feedback form currently in existence offers only limited help to involve users in the development of new services. Furthermore, usage information could be forwarded from the central division to the individual agencies to highlight the popularity of specific services among users.

Third, interaction between the agencies regarding the re-use of services and service components is limited although this functionality of agency interaction was considered during the design of the ecosystem. Most services that that are re-used are those provided by the platform itself, thus, potential network effects could be leveraged and potential time and resource savings could be realised.

<table>
<thead>
<tr>
<th>Role</th>
<th>Capability and Contribution</th>
<th>Gaps of Open Innovation Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer</td>
<td>Feedback regarding individual services provided through online form. Demand for new services.</td>
<td>Central collection of explicit feedback through rating mechanisms and customer community. No implicit feedback on customer demand is collected and made available to service providers.</td>
</tr>
<tr>
<td>Platform provider</td>
<td>Overall platform environment Operational infrastructure Trade support processes for service delivery</td>
<td>No central innovation space through which ideas and feedback can be shared.</td>
</tr>
<tr>
<td>Service provider</td>
<td>Specialist expertise in service domain Contribution of various services to be</td>
<td>Limited interaction and exchange between other agencies.</td>
</tr>
</tbody>
</table>
requirements for platform improvements (esp. Web front-end).

| Broker | No broker activities were observed. | Lack of facilitated interaction and exchange between actors. No actor available to capture good ideas, keep ideas alive, imagine new uses for old ideas, and to put promising concepts to the test. |

Table 3  

| Actors’ contributions and gaps in open innovation implementation. |

4.2  Case 2: Breaking up the Strong Platform Provider

Our case study partner is a platform and service provider for the automotive industry. Their platform provides Web-based services for supply chain management in the automotive and production industries. They serve both the source side (i.e., large automotive companies) and the supply side (i.e., suppliers of automotive companies). In addition to the services directly offered by the platform provider two external service providers offer specialised services that extend the functionality of the platform. Thus, our case partner is able to offer highly sophisticated services by relying on specialised knowledge from other actors which increase customer value of the platform.

In addition to providing the general platform our case partner strongly mediates between the source and supply side. Through an established internal innovation process they continually work at improving the functionality offered by their platform through close collaboration with customers from both the source and supply side. This innovation process involves expert workshops and various expert groups with participants from both sides. After gathering and sufficiently refining service ideas a formal selection process that relies on a large set of evaluation criteria is employed and ideas promising the best outcome are selected for implementation. After implementing the new functionality on the platform our case partner closely works with their customers on integrating the new features into their processes. The case partner concentrates on its core competency as an industry insider with special knowledge and experience of understanding both large manufacturers as well as suppliers of these companies. Although they act as platform provider, most of the operation and development tasks have been outsourced and the core competency is seen in industry insight and experience and thus the ability to support and enhance client processes.

Actors:
- Central actor acting as provider of the central platform and mediator between the source and supply side.
- Customers of the central automotive supply platform, both from source and supply side, provide input regarding new features they would like to see implemented on the platform.
- Two external companies offering services that extend the functionality of the central platform.

Case Discussion

The most prominent role in this case study is that of the central platform provider acting as ecosystem leader and the very active broker role. Although all actors of our collaboration framework are present in this case study several observations in the implementation of open innovation processes can be made. Two third parties offer services on the platform: they were invited to the central platform in a strategic partner selection process and the platform is not open to other service providers. The platform provider holds a rather strict guard over their platform. Consequently, there is no shared innovation space through which service providers could interact with each other and customers to drive innovation towards new services. In particular, service ideas that have been rejected by the platform provider’s internal innovation process are not visible to others who could decide that offering a certain service might present a business opportunity. This close control of the platform may limit the potential benefits of the broker role as interaction with external service providers is restricted to selected
partners. As the amount of third party services offered on the platform is very low, the potential for designing new value-added services through the combination of existing services is also very small.

<table>
<thead>
<tr>
<th>Role</th>
<th>Capability and Contribution</th>
<th>Gaps of Open Innovation Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer</td>
<td>Judges perceived service quality through explicit feedback. Directly involved in innovation activities through interviews, expert workshops, and feedback requests.</td>
<td>Customer interaction only through central platform provider. Customer feedback is only limited available to service providers.</td>
</tr>
<tr>
<td>Platform provider</td>
<td>Service platform that defines general ecosystem environment Operational infrastructure</td>
<td>Limited platform functionality with regards to service offerings by third party service providers.</td>
</tr>
<tr>
<td>Service provider</td>
<td>Specialist expertise in service domain Specialised services</td>
<td>Closed selection of service providers. Limited amount of services offered by providers other than platform provider. Limited interaction between service providers, consequently no re-use of existing components by service providers.</td>
</tr>
<tr>
<td>Broker</td>
<td>Extensive industry insights Mediates between source and supply sides, aggregates innovation demands, facilitates communication, acts as catalyst</td>
<td>Broker activity too closely focused on own provision of service. Limited cross fertilisation between service providers.</td>
</tr>
</tbody>
</table>

Table 4  Actors’ contributions and gaps in open innovation implementation.

5 CONCLUSION

The framework shows the capabilities of the individual actors with regard to service innovation and how these capabilities can be exploited by the overall ecosystem to advance service innovation. This highlights the potential advantages that can arise through the constellation of various actors bound together by a single ecosystem platform. Each actor benefits from the contributions of the other participants. End-users contribute knowledge about actual market demand either in the form of ideas or through feedback provided about the services used. The platform provider contributes the overall environment of the ecosystem platform and serves as an architect to drive innovation projects by extracting and implementing ideas that are likely to benefit the whole ecosystem. Service providers contribute services that extend the innovation space and may thus allow new value added services to be composed. Conversely, they extract and implement service ideas that a provider deems valuable. Finally, brokers engage in transforming ideas already present in the innovation space.

This paper proposed a new way of thinking about an innovation ecosystem where each actor contributes to a collective innovation space rather than single companies chasing their individual innovation projects. The framework serves as an interpretative scheme to structure and analyse each actor’s contribution towards the innovation space. Therefore, we believe, this model can serve as a guide in leveraging the combined resources available in service ecosystems and can guide strategies for businesses to successfully participate in service ecosystems. Moreover, it was apparent that the different types of contributions require adequate tool support to facilitate the networked innovation. Furthermore, open innovation has severe consequences on intellectual properties and sharing thereof. Adequate mechanisms to govern the use and sharing of innovation related information are needed.

From the gaps identified in the case applications we derived several practical implications for the implementation of open innovation processes for service and platform providers:

1) Use explicit user-feedback to improve, re-design, and create new services.
2) Use implicit feedback for continuous improvement.
3) Rely on outsourcing, re-use other provider’s services to save development costs, and get to market quickly and cheaply.
4) Cleverly recombine existing services to create value added services for your customers.
5) Look for service ideas outside your current customer base by looking at how you can complement services provided by others.
6) Evaluate service ideas in the service ecosystem’s innovation space according to your competencies to find service ideas that might be profitable to you but not to others.
7) Work with brokers to get new ideas and refine ideas that are “not there yet.”
8) Provide personalized tools that support the capabilities contributed by each actor.

In order to exploit the capabilities in this network of distributed innovation it is important to “find a governance mechanism that strikes a balance between order and chaos” (Sawhney/Prandelli, 2003). Thus, it becomes apparent that across multiple innovation projects a single actor may play different roles. While driving one project as a service provider or architect, in another situation the actor may only contribute end-user feedback about the services consumed from suppliers. This leads to a certain degree of overlap in the roles across different innovation projects.

References


### SOURCES OF IT DYNAMIC CAPABILITY IN THE CONTEXT OF DATA GENESIS CAPABILITY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0120.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>IT capability, Capacity building, Dynamic Capabilities / relationships / perspective, Resource Based View</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
Abstract

Dynamic Capabilities are often considered as the factor justifying the different degrees of success of organizations in turbulent environment. However Dynamic Capability development remains a difficult issue to research, with a paucity of work directly addressing this question, despite its importance. The explanation of the sources of Dynamic Capabilities would give organizations the instruments to rationally improve their chance of success and to more likely sustain their competitive advantage.

We contribute to the emerging literature on Information Technology (IT) Dynamic Capability by proposing a research framework grounded in the three sources of Dynamic Capabilities: organizational processes, firm history and firm’s assets. Our model takes into consideration also the moderating role played by environmental turbulence on Dynamic Capability and on process performance.

In this contribution we lay the theoretical and methodological groundwork and we foresee the test of the model using Data Genesis (DG) capability as the context. DG is the Dynamic Capability of (1) choosing IT to unobtrusively generate and capture data in digital form, (2) integrating the technology in the appropriate business processes, and (3) managing the digital data so captured.

Keywords: IT capability, Dynamic Capability, capability development, Data Genesis.
1 INTRODUCTION

Explaining the variation in the degree of success of business organizations is an evergreen issue in strategic management and organizational studies (Zollo & Winter, 2002). Among the different concepts developed so far, Dynamic Capabilities are one of the most recent ones, employed to justify the different degree of success, particularly in turbulent environment (Pavlou & El Sawy, 2006; Rai, Patnayakuni, & Seth, 2006; Zollo & Winter, 2002).

However Dynamic Capabilities is an unclear and troubling research construct. Dynamic Capabilities risk being a vague and tautological concept as they are the things that enable organizations to sustain competitive advantage, but they can only be inferred when looking at apparently successful organizations over sustained period of time (Priem & Butler, 2001; Zollo & Winter, 2002). Even if a certain consensus on Dynamic Capabilities is growing (Easterby-Smith & Prieto, 2008), the origin of Dynamic Capabilities over time is still a difficult issue. Testament to the difficulty associated with research on the emergence of Dynamic Capabilities is the paucity of research directly addressing this question. If Dynamic Capabilities really impact organizational success, the explanation of the sources of Dynamic Capabilities would give organizations the instruments to rationally improve their chance of success and to durably sustain their competitive advantage. As a consequence, research in this area is clearly important.

The embryonic research on this subject proposes different theories and models on the sources of Dynamic Capabilities (Montealegre, 2002; Pavlou & El Sawy, 2006; Tanriverdi, 2005; Teece, Pisano, & Shuen, 1997; Zollo & Winter, 2002), but an attempt to integrate these different propositions into a comprehensive and testable research model of the sources of Dynamic Capabilities is missing. In this article we seek to contribute to the emerging literature on the sources of Information Technology (IT) Dynamic Capability, in turbulent environments. We will attempt to integrate the different models and theories on the sources of Dynamic Capabilities into a comprehensive and testable research model of Dynamic Capability development and its antecedents: Organizational Processes, Firm history and Firm’s assets. We question the validity of this comprehensive model on a specific IT Dynamic Capability: Data Genesis.

This paper is organized as follows: §2 introduces the theoretical framework, which is based on the resource-based view theory, and formally defines the Data Genesis (DG) Dynamic Capability construct. §3 summarizes the literature on the sources of Dynamic Capability. §4 describes the research model, its variables and hypotheses. §5 presents the research methodology and §6 concludes this research in progress highlighting the future directions and achievements.

2 THEORETICAL FRAMEWORK

2.1 Resource-based view

The resource-based view has been largely introduced in Information Systems research to theoretically ground studies on competitive advantage and its sustainability at the firm level (Wade & Hulland, 2004). This perspective highlights the importance of the firm’s internal resources for the evaluation of the firm’s competitive advantage (Eisenhardt & Martin, 2000).

Resources are the “assets and capabilities that are available and useful in detecting and responding to market opportunities or threats” (Wade & Hulland, 2004). More specifically, “assets are defined as anything the firm can use in its processes for creating, producing, and/or offering its products (and/or services) to the market, whereas capabilities are repeatable patterns of actions in the use of assets to create, produce, and/or offer products (and/or services) to the market” (Sanchez, Heene, & Thomas, 1996; Wade & Hulland, 2004). We label the products and/or services offered to the market by an
organization a bundle (Kohli & Bharadway, 2007) of products/services, in line with the cross-
disciplinary service science movement (Chesbrough & Spohrer, 2006).

The resources that are valuable and rare temporary provide the competitive advantage. The extent to
which these resources are also inimitable, immobile and not substitutable between firms explains the
sustainability over time of the competitive advantage (Barney, 1991).

The attention paid by this perspective to the internal resources of the firm has the weakness of
excluding the socio-economic environment outside the firm. In fact, the environmental conditions
could change and make the firm’s resources far less valuable (Leonard-Barton, 1992). Hence the
resource-based view has been extended to better explain firm performance in turbulent environments
(Eisenhardt & Martin, 2000).

2.2 Dynamic Capabilities

In turbulent environments, organisations need to constantly match or create market changes and
Dynamic Capabilities are “the firm’s processes that use resources – specifically the processes to
integrate, reconfigure, gain and release resources – to match and even create market change”
(Eisenhardt & Martin, 2000). Hence, Dynamic Capabilities have the potential (Prieto & Easterby-
Smith, 2006) to create, to evolve and to recombine internal existing resources to adapt to turbulent
environments (Teece, Pisano, & Shuen, 1997). This adaptability is especially required in fast-paced
technological environments (Banker, Bardhan, Hsihui, & Shu, 2006; Teece, Pisano, & Shuen, 1997;
Zahra, Sapienza, & Davidsson, 2006), as it has been theorized that adaptability can lead to improved
customer value (Sambamurthy, Bharadwaj, & Grover, 2003; Wheeler, 2002).

The Resource-based View foresees a direct relationship between Dynamic Capabilities and the firm’s
process performance (Ray, Barney, & Muhanna, 2004; Ray, Muhanna, & Barney, 2005; Zahra,
Sapienza, & Davidsson, 2006). However the process outperformance by one firm in competing
environments does not automatically imply any firm’s sustained competitive advantage due to the
several mediating and moderating variables interposing between process outperformance and firm’s
sustainable competitive advantage (Ray, Muhanna, & Barney, 2005). The acknowledgement of this
interposition between single process outperformance and sustained competitive advantage stimulates
the study Dynamic Capabilities and its outputs, without employing organizational global performance
as dependent variable.

Hence, in today’s competitive environment characterized by increasing IT intensity (McAfee &
Brynjolfsson, 2008) organizations should be capable of integrating new or established IT. A number of
Dynamic Capabilities have been documented in the literature, in this study we focus on an emerging
Dynamic Capability, Data Genesis (DG) (Piccoli & Watson, 2008).

2.3 Data Genesis as a Dynamic Capability

We define Data Genesis capability as the three-fold process of:

• Choosing Information Technology (IT) (Wheeler, 2002; Williams, 2003) to unobtrusively
  generate and capture data in digital form. Such IT may be emerging IT (Wheeler, 2002): a new
  technology not commercially viable (e.g., multi-touch displays). Otherwise, such IT may be
  enabling IT (Wheeler, 2002): an established technology used in an innovative application by the
  firm (e.g., RFID in gaming chips to track table play);
• Integrating the IT in the appropriate business processes;
• Managing the digital data so produced in order to continuously deliver accessible, accurate,
  complete and current digital data.

Data Genesis is a Dynamic Capability as it is the process that uses IT resources to gain and release
data and as it has the potential to create, to evolve and to recombine internal existing IT and data to
adapt to turbulent environments. This potential is based on the degree of Reconfigurability (Pavlou &
El Sawy, 2006) of the ineffective Data Genesis process into more promising one that better match the environment, better, faster, and cheaper than the competition (Eisenhardt & Martin, 2000).

Note that the DG capability is concerned with the unobtrusively generation and capture of digital data and its management, not with its actual use in, for example, analytical processes. In other words, DG is a prerequisite to being able to compete on analytics, thanks to the provision of accessible, accurate, complete and current digital data. The outperformance in the digital data accessibility, accuracy, completeness and currency are the valuable outcomes of Data Genesis.

Exemplars of DG capability are emerging, such as Harrah’s corporation: this company systematically integrates IT, such as computerized slot machines or RFID chips, to gain unobtrusively valuable digital data on customers’ behaviour at the Harrah’s casinos and it exploits these pieces of data to profile and reward customers (DeLong & Vijayaraghavan, 2003; Piccoli & Watson, 2008). Note that the unobtrusive generation of the data in the above example pertains to DG capability while the exploitation is within the scope of other capabilities (e.g., data analysis).

3 SOURCES OF DYNAMIC CAPABILITY

The relevance of DG capability in fast-paced IT environments motivates its choice as the empirical Dynamic Capability on which we build the model of the sources of Dynamic Capability. If several studies have investigated Dynamic Capabilities and their effects on business performance, there is not an equivalent attention to the sources of Dynamic Capabilities (Montealegre, 2002; Pavlou & El Sawy, 2006; Tanriverdi, 2005; Teece, Pisano, & Shuen, 1997; Zahra, Sapienza, & Davidsson, 2006; Zollo & Winter, 2002) and an integrative model gathering all these sources is missing.

By the consequence, we propose an integrative research model of the sources of Dynamic Capability. Our starting point to understanding the sources of Data Genesis Capability are the three sources of Dynamic Capabilities in rapid technological change environments (Teece, Pisano, & Shuen, 1997):

- The organizational processes of sensing, coordination, integration, learning.
- The firm’s assets, which define the firm-specific strategic position,
- The firm history, which accounts for the path dependent nature of capabilities.

Leveraging these theoretical sources of Dynamic Capabilities, a first case study empirically highlighted the set of actions to develop capabilities (Montealegre, 2002). The organizational processes of sensing, coordination, integration and learning emerged as important in capability development. By contrast, the firm’s assets and the firm history played a marginal supporting role.

Others have theorized that the learning mechanisms would be the main independent variable influencing the development of the Dynamic Capabilities (Zollo & Winter, 2002). Hence theoretically, learning would be the main organizational process for the development of Dynamic Capabilities. The other organizational processes of sensing, coordination and integration as well as the firm’s assets and history have been neglected.

In 2005, firm’s assets and organisational processes were combined in a model of Dynamic Capability development (Tanriverdi, 2005). In particular capability development depended on the IT infrastructure and IT management processes. IT infrastructure is one kind of firm’s assets, while IT management processes are a portion of the organisational processes of sensing, coordination, integration and learning. Nevertheless, this integration in one single model of Dynamic Capability development of these two different sources of Dynamic Capability still excluded firm history as the third source of Dynamic Capability.

An additional theoretical contribution reaffirmed the role of the organizational processes and of the environmental turbulence in the dynamic capability development (Zahra, Sapienza, & Davidsson, 2006). The processes of coordination, selection and combination were proposed as the main organizational processes that enable the firm to build dynamic capabilities. The theoretical model
proposed also that the environmental turbulence decreases the relative performance of the existing capabilities and hence stimulates the development of new dynamic capabilities in replacement of the obsolescing ones. The others sources of dynamic capabilities, as such firm history and firm’s assets, had no impact on dynamic capability development.

The last work proposes and measures the impact of IT leveraging competence on the development of Dynamic Capabilities (Pavlou & El Sawy, 2006). IT leveraging competence is proposed as the only independent variable, is conceived as the ability to effectively use IT functionalities and it can be classified among the firm’s assets. The other firm’s assets and the firm history were excluded from this study. On the contrary, the organizational processes of coordination, integration and learning played the role of mediators between IT leveraging competence and Dynamic Capabilities. Finally, the environmental turbulence moderated the relationships between all these variables.

Even if at different extent, all the three sources of Dynamic Capabilities have been studied separately by different authors. The lack of a comprehensive research model including all the three kinds of sources of Dynamic Capability appears as the main research gap. Therefore, our main contribution is the design of a model which combines the three sources of Dynamic Capabilities in a comprehensive research model of the sources of Dynamic Capability.

4 RESEARCH MODEL

Our research model integrates and organizes previous literature around the three-fold classification of the sources of Dynamic Capabilities: organizational processes, firm’s assets and firm history (Teece, Pisano, & Shuen, 1997). We posit that these processes are responsible for the emergence of DG capability. The outcome variables, Digital Data Accessibility, Accuracy, Completeness and Currency, dependent on DG capability and the variable Environmental Turbulence moderates some relationships. The specification of our hypothesized relationships completes the presentation of the variables (Figure 1).

4.1 Organizational Processes

The Organizational Processes of Sensing, Learning, Coordinating, and Integrating play a pivotal role in developing Dynamic Capabilities when the opportunity or need arise (Kogut & Zander, 1996; Maritan, 2007; Pavlou & El Sawy, 2006; Zahra, Sapienza, & Davidsson, 2006). The Sensing process is related to the understanding of the environment, the identification of market needs and opportunities. The Learning process determines the building of new thinking and the generation of new knowledge to enhance existing resources. The Coordinating process is about the allocation of resources, the assignment of tasks and the synchronization of the activities. The Integrating process concerns the implementation of the new configurations of operational competencies by developing the required patterns of interaction. So:

\( H1: \) the effectiveness of the Sensing has a positive and direct impact on DG capability.

\( H2: \) the effectiveness of the Learning has a positive and direct impact on DG capability.

\( H3: \) the effectiveness of the Coordinating has a positive and direct impact on DG capability.

\( H4: \) the effectiveness of the Integrating has a positive and direct impact on DG capability.

4.2 Firm’s assets

Different kind of assets can positively influence the possibility to develop new capabilities: technological assets, complementary assets, financial assets, reputational assets, structural assets, institutional assets, market structure assets (Teece, Pisano, & Shuen, 1997). The main kind of asset influencing the development of Data Genesis capability are IT assets as on IT assets the IT
Capabilities, and specifically Data Genesis, are built (Tanriverdi, 2005). For example, TER Rhone-Alpes, a French regional railway company chose RFID cards and readers to generate and capture digital data about the passengers’ flows on its trains. Passengers load on this RFID card their train passes and they have to validate the card at a RFID card reader, before getting on any train. These digital data are then used to adapt the frequency and the capacity of the trains based on their actual use, improving, by one hand, the passenger satisfaction and, by the other hand, the optimization of the train fleet.

IT assets are a two-fold category composed of: IT Infrastructure and Information Repositories (King, Grover, & Hufnagel, 1989; Piccoli & Ives, 2005). IT Infrastructure is “the base foundation of the IT portfolio (including both technical and human assets), shared through the firm in the form of reliable services” (Broadbent, Weill, & St. Clair, 1999) or functionalities (Fink & Neumann, 2007; Pavlou & El Sawy, 2006; Zhu & Kraemer, 2005). The IT Infrastructure varies in reach, and range (Piccoli & Ives, 2005). The reach of the IT Infrastructure measures the extent of the connectivity both within and outside of the firm, while the range of the IT Infrastructure sizes the scope of services that it can support. As reach and range of the IT Infrastructure increase, the IT Infrastructure ability to support capability development increases as well. The reach and range of the existing IT Infrastructure influence the possibility and cost of IT integration for gaining unobtrusively valuable digital data hence impacting on the development of DG capability. TER Rhone-Alpes has progressively extended the reach of its IT infrastructure to deploy RFID card readers from the principal traffic railways to the secondary railways, in order to capture digital data on passenger flows.

The second category of IT assets gathers the Information Repositories. An Information Repository is “a collection of logically related data, organized in a structured form, that is accessible and usable for decision-making purposes” (Piccoli & Ives, 2005). As capabilities relying on organized data need Information Repositories to develop (Piccoli & Ives, 2005), also DG capability needs Information Repositories to develop. Given that DG capability includes the use and management of digital data, DG capability needs Information Repositories to organize and access the gained digital data. A lack of Information Repositories would restrain the organization in data availability and by consequence impeding the development of DG capability. TER Rhone-Alpes exploited its information repositories on the sales of paper-based ticket and passes to establish the first kinds of train passes to transfer on the RFID card and the first train stations to deserve with the RFID readers.

Finally, IT Infrastructure and Information Repositories are subjected to asset stock accumulation dynamics (Ingemar & Cool, 1989; Piccoli & Ives, 2005): the IT Infrastructure and the Information Repositories can be accumulated by the organization into IT asset stock over time. The reach and the range of IT Infrastructure as well as the volume of Information Repositories can increase, extending the pre-existing IT asset stock. This IT asset stock accumulation influences, in general, Dynamic Capability development as well as DG capability in particular. An extended stock of IT assets facilitates the development of DG capability. Therefore, we respectively hypothesise that:

\( H5: \) the stock of IT Infrastructure has a positive and direct impact on DG capability.

\( H6: \) the stock of Information Repositories has a positive and direct impact on DG capability.

4.3 Firm history

Firm history explains the existing firm’s position and the same time it influences the firm’s opportunities ahead, framing the path dependencies of organizations (Teece, Pisano, & Shuen, 1997). Present capabilities depend on previous ones and they constrain new ones, because learning tend to be local and related to existing processes (Teece, Pisano, & Shuen, 1997; Zahra, Sapienza, & Davidsson, 2006). By consequence, the development of DG capability depends on historically existing Dynamic Capabilities closely related to DG, as such: IT capability and Information capability.

IT capability is the multi-dimensional and enterprise-wide capability to leverage IT (Bharadwaj, Sambamurthy, & Zmud, 1999). The historical capacity to leverage IT will favours the recognition by
the firm’s IT personnel of the potential of emerging/enabling IT to generate and capture digital data and the good relationships between IT personnel and line management in integrating such IT within appropriate business processes is critical. The lack of IT capability would make unclear the choice of the IT to integrate and would cause the eventual IT integration ineffective. By consequence, digital data would not be accessible or would be of poor quality, impeding any effective use (Culnan, 1983; O’Reilly, 1982; Zimmer, Henry, & Butler, 2007).

Harrah’s IT managers and customer service managers realized very early on that a modern slot machine is a digital computer and they worked together to develop a customer relationship management information system which collect over time digital data on the customers’ behaviours at slot machines (DeLong & Vijayaraghavan, 2003; Piccoli & Watson, 2008).

The concept of Information capability is rooted in Information Theory (Shannon & Warren, 1949) and Information Capability is proposed as the capacity of disseminating (Mathews & Healy, 2007), or applying and managing (Yoon, 2005), or processing (Lin, 2005) information. The historical capacity to manage information will enable the firm to manage digital data and therefore take advantage of its ability to unobtrusively generate the data in digital form. Conversely, the inability to manage digital data would negate the value of data capture and integration.

Harrah’s corporation preferred digital data on guest preferences and transactions coming from slot machines than the unstructured information coming from customer service staff. The same firm processes the collected customer data, from slot machines, to profile gamblers and it disseminates these profiles, throughout the different casinos (DeLong & Vijayaraghavan, 2003; Piccoli & Watson, 2008).

So, we hypothesize that:

**H7: IT capability has a positive and direct impact on DG capability.**

**H8: Information capability has a positive and direct impact on DG capability.**

### 4.4 Digital Data Accessibility, Accuracy, Completeness and Currency

The performance of an organization in Data Genesis capability does not automatically imply any sustained competitive advantage due to the several mediating and moderating variables interposing between Data Genesis outperformance and organizational sustainable competitive advantage (Ray, Muhanna, & Barney, 2005). Data Genesis capability aims at outputting accessible, accurate, complete and current digital data. The use in, for example, analytical processes of the gained digital data will depend on their accessibility, accuracy, completeness and currency (Culnan, 1983; O’Reilly, 1982; Zimmer, Henry, & Butler, 2007).

Information accessibility is the extent to which an individual perceives that any particular source is available for use (Zimmer, Henry, & Butler, 2007). Information accessibility is the most important driver for information source selection for use, with people consistently choosing and using lower-quality sources that are more accessible over higher-quality sources that are less accessible (Culnan, 1983; O’Reilly, 1982; Zimmer, Henry, & Butler, 2007).

Nevertheless, information quality is important because when sources are equally accessible, individuals will consistently choose and use sources that are perceived of higher quality (Hirsch & Dinkelacker, 2004; O’Reilly, 1982). Information Accuracy, Completeness and Currency are dimensions of the quality of the information retrieved from an information system (DeLone & McLean, 1992; Nelson, Todd, & Wixom, 2005). Accuracy refers to the degree to which information is correct, unambiguous, meaningful, believable, and consistent. Completeness is about the degree to which all possible states relevant to the user population are represented in the stored information. Currency concerns the degree to which information is up-to-date and precisely reflecting the current state of the world that it represents.
Harrah’s corporation appreciates the quality and accessibility of the collected data on customers at the slot machines. Basing on the accessibility, accuracy, completeness and currency of the accumulated transactional data from past guests, Harrah’s can quickly estimate the customer’s future value within minutes of the player joining the program. This enables the casino to start treating the customer according to his or her future value rather than having to wait for observed play before starting to provide rewards (Piccoli & Watson, 2008).

Consequently, the hypothesis we propose is that:

H9: the DG capability has a positive and direct impact on Digital Data Accessibility.

H10: the DG capability has a positive and direct impact on Digital Data Accuracy.

H11: the DG capability has a positive and direct impact on Digital Data Completeness.

H12: the DG capability has a positive and direct impact on Digital Data Currency.

4.5 Environmental Turbulence

Environmental Turbulence describes the general conditions of uncertainty and/or unpredictability caused by the changes in customer preferences and technology development (Mendelson & Pillai, 1998). Customer preferences’ turbulence causes unpredictability in market demand, while, technology development’s turbulence causes uncertainty regarding new technological breakthroughs.

On one hand, Environmental Turbulence stimulates the reconfiguration of existing capabilities, increasing the possibility that the Organizational Processes of Sensing, Learning, Coordinating, and Integrating develop new valuable capabilities (Sambamurthy, Bharadwaj, & Grover, 2003; Zahra, Sapienza, & Davidsson, 2006). On the other hand, Environmental Turbulence weakens the process performances depending on the existing Dynamic Capabilities (Teece, Pisano, & Shuen, 1997). So, we propose that Environmental Turbulence reinforces the relationship between Organizational Processes and Data Genesis, while it attenuates the relationships between Data Genesis and Information Accessibility, Accuracy, Completeness and Currency:

H13: Environmental Turbulence positively moderates (i.e. reinforced) the relationship between Sensing and DG capability.

H14: Environmental Turbulence positively moderates (i.e. reinforced) the relationship between Learning and DG capability.

H15: Environmental Turbulence positively moderates (i.e. reinforced) the relationship between Integrating and DG capability.

H16: Environmental Turbulence positively moderates (i.e. reinforced) the relationship between Coordinating and DG capability.

H17: Environmental Turbulence negatively moderates (i.e. attenuates) the relationship between DG capability and Digital Data Accessibility.

H18: Environmental Turbulence negatively moderates (i.e. attenuates) the relationship between DG capability and Digital Data Accuracy.

H19: Environmental Turbulence negatively moderates (i.e. attenuates) the relationship between DG capability and Digital Data Completeness.

H20: Environmental Turbulence negatively moderates (i.e. attenuates) the relationship between DG capability and Digital Data Currency.
We are now engaged in the data collection phase of this research in order to test the model through a double Configurational-Structural approach. Firstly, we will test our theory-based structural model, through Structural Equation Modelling (SEM). As the organizational alignment on the three sources of Dynamic Capabilities facilitates the development of new capabilities (Teece, Pisano, & Shuen, 1997), we think that the internal consistency among the three sources of Data Genesis will enhance its development. Hence, we will cluster the empirical data sample, in order to identify the different organizational configurations (Meyer, Tsui, & Hinings, 1993). Finally, we will test the configuration-based structural model, through SEM.

A questionnaire-based survey will be the main source of empirical data on a sample of key informants in different organizations. The target population is hospitals, as they are data intensive organisations and DG capability could advance the medical care standards (Piccoli & Watson, 2008). 200 responses from distinct organisations are pursued to assure a convenient sample size due to the 100 measured items on the 15 constructs (with 6 items for the most complex construct).

The key informants will be IT managers and line managers in order to reduce common method bias (Straub, Boudreau, & Gefen, 2004). IT managers are likely to be the most informed about IT assets, IT capabilities in general and Data Genesis Capability specifically. Line managers are likely to be the most informed about environmental turbulence, organisational processes, Information Capability and Data Genesis output. A formal check will assess the managers’ understanding on the different topics of the questionnaire (Pavlou & El Sawy, 2006), using a cut-off point of 5 out of 7.

Figure 1 The research model

5 RESEARCH METHODOLOGY
Existing measurement scales operationalize all the constructs of the model.

- **Organizational Processes – Sensing.** The effectiveness in sensing the environment will be reflected by the adaptation of the market orientation measurement scale (Kholi & Jaworski, 1990; Pavlou & El Sawy, 2006).

- **Organizational Processes – Learning.** Effectiveness in learning will be measured by the adaptation of the absorptive capacity measurement scale (Cohen & Levinthal, 1990; Pavlou & El Sawy, 2006).

- **Organizational Processes – Coordinating.** Effectiveness in coordinating will be evaluated by the adaptation of the coordination capability measurement scale (Malone & Crowston, 1994; Pavlou & El Sawy, 2006).

- **Organizational Processes – Integrating.** Effectiveness in integrating will be estimated by the adaptation of the collective mind measurement scale (Pavlou & El Sawy, 2006; Weick & Roberts, 1993).

- **Firm history – Information capability.** The Information capability construct will be measured adapting the Information capability measurement scale (Marchand, Kettinger, & Rollins, 2002) with its two dimensions: Information management and Information behavior.

- **Firm history – IT capability.** The IT capability measurement scale will adapt the Technical Capability, Behavioral Capability and Business Capability dimensions of IT Personnel Capability construct (Fink & Neumann, 2007).

- **Firm’s Assets – IT Infrastructure.** The IT infrastructure measurement scale will adapt the IT infrastructure Capability (Fink & Neumann, 2007).

- **Firm’s Assets – Information Repository.** The Information Repository construct will be reflected by the adaptation and unification of two different Repository scales (Freeze & Kulkarni, 2005).

- **DG Capability.** The DG capability scale will adapt:
  - Choosing New Emerging/Enabling Technologies construct (Wheeler, 2002; Williams, 2003), to measure the ability to choose emerging/enabling IT to gain unobtrusively valuable digital data.
  - IT Business process integration category of the IT Capability construct (Bharadwaj, Sambamurthy, & Zmud, 1999), to measure the ability to integrate in the business processes such IT.
  - Information Management dimension of the Information capability measurement scale (Marchand, Kettinger, & Rollins, 2002), to measure the ability to manage digital data.
  - Reconfigurability. The potential to reconfigure Data Genesis integrating will be estimated by the adaptation of Reconfigurability measurement scale (Pavlou & El Sawy, 2006).

- **Digital Data Accessibility.** The Digital Data Accessibility construct will be assessed by the adaptation of the Information Accessibility measurement scale (Zimmer, Henry, & Butler, 2007).

- **Digital Data Accuracy.** The Digital Data Accuracy construct will be measured through the adaptation of the Information accuracy measurement scale (Nelson, Todd, & Wixom, 2005).

- **Digital Data Completeness.** The Digital Data Completeness construct will be measured through the adaptation of the Information completeness measurement scale (Nelson, Todd, & Wixom, 2005).

- **Digital Data Currency.** The Digital Data Currency construct will be measured through the adaptation of the Information currency measurement scale (Nelson, Todd, & Wixom, 2005).

- **Environmental Turbulence.** The measurement of Environmental Turbulence construct will be based on the Turbulent Environment scale (Pavlou & El Sawy, 2006).

A set of Control Variables complements the measurement scale of the main constructs of the model. Several factors that have been previously shown to be related to Dynamic Capability development will be measured, so that their effects on Data Genesis Capability development will be controlled:

- The functional role of the respondents: management versus non-management positions (Fink & Neumann, 2007).

- The size of the organization through the number of employees (Bhatt & Grover, 2005; Fink & Neumann, 2007).

- The size of the IT department through the number of IT personnel (van der Heijden, 2000).
• The seniority of the respondents among senior, mid-level or junior managers (Fink & Neumann, 2007).

6 CONCLUSION AND FUTURE RESEARCH

Dynamic Capabilities are often considered as the factor justifying the different degrees of success of organizations in turbulent environment. However Dynamic Capabilities development remains partially unexplained. The explanation of the development of Dynamic Capabilities would give organizations the instruments to rationally improve their processes and increase indirectly their chances of success.

We contribute to the emerging literature on IT Dynamic Capability development by proposing and testing a research model on DG: the Dynamic Capability of (1) choosing IT to unobtrusively generate and capture data in digital form, (2) integrating the technology in the appropriate business processes, and (3) managing the digital data so produced.

This research in progress foresees the test of the research model on DG capability before the conference attendance in order to present some preliminary results at that time. Finally, future research includes the validation of the model in others organisations and for other Dynamic Capabilities in order to generalize the findings.

References


Fink, L., & Neumann, S. (2007). Gaining Agility Through It Personnel Capabilities: The Mediating Role Of It Infrastructure Capabilities *Journal Of Ais, 8*(8).


HOW TO SELECT MEASURES FOR DECISION SUPPORT SYSTEMS – AN OPTIMIZATION APPROACH INTEGRATING INFORMATIONAL AND ECONOMIC OBJECTIVES

<table>
<thead>
<tr>
<th><strong>Journal</strong></th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manuscript ID</strong></td>
<td>ECIS2009-0087.R1</td>
</tr>
<tr>
<td><strong>Submission Type</strong></td>
<td>Research Paper</td>
</tr>
<tr>
<td><strong>Keyword</strong></td>
<td>Decision support systems, Information overload, Critical Success Factors (CSFs), Design research</td>
</tr>
</tbody>
</table>
HOW TO SELECT MEASURES FOR DECISION SUPPORT SYSTEMS – AN OPTIMIZATION APPROACH INTEGRATING INFORMATIONAL AND ECONOMIC OBJECTIVES

Röglinger, Maximilian, FIM Research Center Finance & Information Management, Department of Information Systems Engineering & Financial Management, University of Augsburg, Universitätsstraße 16, 86135 Augsburg, Germany, maximilian.roeglinger@wiwi.uni-augsburg.de

Abstract

It is still an open issue of designing and adapting (data-driven) decision support systems and data warehouses to determine relevant content and in particular (performance) measures. In fact, some classic approaches to information requirements determination such as Rockart’s critical success factors method help with structuring decision makers’ information requirements and identifying thematically appropriate measures. In many cases, however, it remains unclear which and how many measures should eventually be used. Therefore, an optimization model is presented that integrates informational and economic objectives. The model incorporates (statistic) interdependencies among measures – i.e. the information they provide about one another –, decision makers’ and reporting tools’ ability of coping with information complexity as well as negative economic effects due to measure selection and usage. We show that in general the selection policies of all-or-none or the-more-the-better are not reasonable although they are often conducted in business practice. Finally, the model’s application is illustrated by the German business-to-business sales organization of a global electronics and electrical engineering company as example.

Keywords: decision support systems, information overload, critical success factors, design research.
MOTIVATION AND OBJECT OF RESEARCH

Due to the complexity of intra- and extraorganizational structures, it is impossible for decision makers in general – and executives in particular – to continuously monitor all fields of action that possibly require intervention. With reports containing in average up to 15,000 data points based on measures, i.e. key figures or (performance) indicators, information proliferation makes it even harder to focus on decision-relevant information (Axson 2007). Some measures more or less significantly influence the complexity of reports and the amount of time needed to understand them. The number of measures also drives the costs for customizing and maintaining reports. Hence, a central problem in the design and adaptation of (data-driven) decision support systems (DSS) and data warehouses (Alter 1980, Inmon 2005) still is to determine relevant fields of action and to select appropriate measures (Eccles 1991, Watson et al. 1993). Particularly the latter requires formal research (Evans 2004).

Some classic approaches to information requirements determination (IRD), such as Rockart’s critical success factors (CSFs) method (1979), provide valuable assistance with structuring decision makers’ information requirements (IR) and identifying thematically appropriate measures. However, these measures are often too many and it is unclear which should eventually be used. In this respect, decision makers’ cognitive restrictions (Browne et al. 2002, Davis 1982), limitations of reporting tools such as management cockpits and dashboards (Sisfontes-Monge 2007), and negative monetary implications need to be considered. As for measure selection in particular, there are additional deficiencies with respect to whether the selection process is intersubjectively comprehensible, decision makers can participate systematically, and (statistic) interdependencies among measures (e.g. quantifiable by means of correlation or contingency coefficients) are considered. In business practice, these deficiencies can result in that measure selection is based on “gut instinct”, that many time-consuming interviews are conducted, and that the utility of selected measures remains doubtful. Therefore, the research question is: Which and how many measures shall be selected from a preselected set of thematically appropriate measures in order to provide decision makers with optimal information as regards informational and economic objectives?

The paper relies on a design-oriented, formal, and deductive approach (Hevner et al. 2004). Section 2 compares existing approaches with respect to general requirements and identifies the research gap. Section 3 proposes an optimization model as artifact. Section 4 evaluates the optimization model by illustrating its application in business practice and by assessing how it meets the general requirements outlined above. Section 5 summarizes the results and points out future research.

RELATED WORK

Currently, measures are often contained in performance measurement systems (PMS). In management accounting and operations management literature, there is a range of requirements on PMS (e.g. Artley et al. 2001, Caplice et al. 1995, Neely et al. 1995). Accordingly, PMS are expected to capture all relevant constituencies of a specific field of action (completeness, R.1), to encompass a manageable amount of measures (clarity, R.2), and to transfer the overall business strategy to decision makers (vertical integration, R.3). The process of measure selection should be intersubjectively comprehensible (intersubjectivity, R.4), consider (statistic) interdependencies among measures (interdependencies, R.5), and involve domain experts (participation, R.6). Although the requirements are somehow vague due to prosaic formulation, the author considers that they provide basic assistance with comparing existing approaches and with identifying the research gap (see Table 1 where completeness is omitted as it is not addressed by any approach).

In the following, selected approaches from international journals and textbooks are presented. Due to space restrictions, this is done briefly. Giorgini et al. (2008) present a goal-oriented approach to determine IR for data warehouses that considers the organizational environment and decision makers’
needs. Neely et al. (2000) advocate a selection of measures in terms of a cost-benefit-analysis. Liebetruth et al. (2006) present a linear optimization model with which a utility-optimal subset of measures can be chosen from a set of preselected and thematically appropriate measures. Rockart (1979) shows how decision makers’ IR can be structured and reduced to a few essential fields of action, the so-called CSFs, each of which is monitored by several measures (see also Leidecker et al. 1984). Axson (2007) extends CSFs analysis by incorporating additional interactive elements, distinguishing primary and supporting measures as well as vaguely postulating “minimal confusion”.

The following findings are noteworthy: Almost all approaches neglect clarity (R.2) as they do not specify how many measures are to be selected. One approach allows to set a maximum number of measures. This is arbitrary and considers neither the decision makers’ information processing capacity nor economic implications. All approaches are vertically integrated by linking measures with CSFs, business strategy, or goals (R.3). Moreover, measure selection is (at least) partially subjective (R.4). Interdependencies among measures are not considered (R.5). Most approaches involve decision makers by means of explorative elements (e.g. interviews or games) (R.6). Summing up, there is a primary research gap with respect to clarity (R.2) and interdependencies (R.5). Furthermore, there still is potential for improvement with respect to intersubjectivity (R.4) and participation (R.6).

We focus on the primary research gap. In the end, this will also ameliorate the other requirements. We adopt the ideas of optimization and preselection of thematically appropriate measures from Liebetruth et al. (2006), the structuring momentum of CSFs from Rockart (1979), and the idea of explicitly incorporating negative economic implications from Neely et al. (2000). Our contribution is that we formally address the trade-off between provided information, information complexity, and negative economic implications to determine which and how many measures should be selected optimally.

### 3. AN OPTIMIZATION MODEL FOR MEASURE SELECTION

Consider a company where the reporting has historically grown and multiple (data-driven) DSS and data warehouses are in use. In order to react on its decision makers’ demand for clear information, the company launches a project for implementing a consolidated DSS. Two essential steps in this project are: (1) structuring the decision makers’ IR into relevant fields of action and (2) (pre-)selecting thematically appropriate measures from the existing systems – assuming no new measures will be added. In most cases, it will not be reasonable to integrate all preselected measures – nor even only those desired by the decision makers (e.g. Ackoff 1967, Davis 1982). This is for several reasons: some measures may (partially) “overlap” due to (statistic) interdependencies, decision makers can only cope with restricted information complexity, and customizing as well as maintaining reporting tools is expensive. It is advisable to analyze in advance which fraction of the preselected measures the consolidated DSS should contain. Whereas above thematically appropriate measures had to be (pre-)

<table>
<thead>
<tr>
<th>Clarity (R.2)</th>
<th>Vertical integration (R.3)</th>
<th>Intersubjectivity (R.4)</th>
<th>Interdependencies (R.5)</th>
<th>Participation (R.6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Giorgini et al. (2008)</td>
<td>No maximum of measures</td>
<td>By goal analysis</td>
<td>Subjective mapping of measures to goals</td>
<td>Isolated consideration</td>
</tr>
<tr>
<td>Neely et al. (2000)</td>
<td>No maximum of measures</td>
<td>By business strategy</td>
<td>Partial intersubjectivity via checklists</td>
<td>Postulated, but not elaborated</td>
</tr>
<tr>
<td>Liebetruth et al. (2006)</td>
<td>Arbitrary maximum of measures</td>
<td>By CSFs</td>
<td>Partial intersubjectivity via optimization model</td>
<td>Postulated, but not elaborated</td>
</tr>
<tr>
<td>Rockart (1979)</td>
<td>No maximum of measures</td>
<td>By CSFs</td>
<td>Subjective mapping of measures to CSFs</td>
<td>Isolated consideration</td>
</tr>
<tr>
<td>Axson (2007)</td>
<td>No maximum of measures</td>
<td>By CSFs and business strategy</td>
<td>Subjective mapping of measures to CSFs</td>
<td>Isolated consideration</td>
</tr>
</tbody>
</table>

Table 1. Comparison of existing approaches to measure selection
selected, here measures are of interest that together provide much information about other measures. As indicated, two perspectives are important here: the economic and the informational perspective. While the former is indispensable when investing in IT, the latter is necessary as DSS primarily aim at supporting decision processes by supplying decision-relevant information (e. g. Power 2002).

In order to determine the optimal fraction of measures, we propose an optimization model. Though being inherently discrete, the problem of measure selection can be interpreted as approx. continuous for sufficiently many measures. This allows to determine algebraic solutions and to gain general insights. We maintain the affiliation with the original problem setting and make reasoning about functions more illustrative by using discrete examples. A basic model for the informational perspective is proposed in section 3.1 and extended by the economic perspective in section 3.2.

3.1 A basic model for the informational perspective

Let us first consider the informational perspective where information has no price. Selecting one of the preselected measures provides information about the measure itself – as it becomes known – and about non-selected measures – due to (statistic) interdependencies. This creates informational utility. The more strongly a measure interdepends with non-selected measures, the more informational utility it creates. This is because a stronger interdependency allows to estimate values more reliably. There are also negative informational effects of selecting measures. Due to increasing information complexity, each additional measure makes it harder to cognitively process the entire amount of information. This creates informational disutility. Thus, there is an informational trade-off. The question is: Up to which optimal fraction of measures does the utility due to more information justify the disutility due to higher information complexity? The optimization model relies on the following assumptions:

A.1: There is a given finite set of measures that have been preselected ex ante with respect to thematic appropriateness. Between some measures there are meaningfully interpretable pairwise (statistic) interdependencies, that is, selected measures provide information about (the values of) non-selected measures. All measures together satisfy the decision makers’ information requirements and provide complete information. Moreover, all measures together cause highest complexity.

A.2: The fraction of the preselected measures that will be integrated into the consolidated DSS, \( x \in [0;1] \), is infinitely divisible (see discussion above). With \( x = 0 \), no measures are selected. With \( x = 1 \), all measures are selected.

A.3: \( U_{\text{info}}(x) \) represents the informational utility due to the information that a fraction of selected measures provides about itself and non-selected measures. \( D_{\text{info}}(x) \) represents the informational disutility due to information complexity. Both are functions of \( x \) and can be forecast ex ante.

On these assumptions, the informationally optimal fraction of measures \( x^{\text{opt, info}} \) can be determined by optimizing the difference between \( U_{\text{info}}(x) \) with \( D_{\text{info}}(x) \). This difference is also called informational net utility \( U_{\text{info, net}}(x) \). The corresponding objective function is given by:

\[
U_{\text{info, net}}(x) = U_{\text{info}}(x) - D_{\text{info}}(x) = \max! \tag{1}
\]

In order to formalize the optimization model, \( U_{\text{info}}(x) \) and \( D_{\text{info}}(x) \) are examined. We start with \( U_{\text{info}}(x) \). If a (rational) decision maker were restricted to select only one measure, he would select the measure with the highest individual informational utility – say \( m_1 \) –, i. e. the measure that in sum interdepends most strongly with the non-selected measures. If the decision maker were allowed to select two measures, he would take those that create the highest joint informational utility – say \( m_2 \) and \( m_3 \). In general, this joint informational utility is higher than the individual utility of \( m_1 \). This is because either \( m_1 \) is kept (as \( m_2 \) or \( m_3 \)) and another measure is added or \( m_1 \) is discarded and two other measures with higher joint utility are chosen. The only exception is if all measures interdepend perfectly. In this case already one measure alone – no matter which – provides complete information. In general, the joint informational utility of \( m_2 \) and \( m_3 \) is smaller than the sum of both individual utility values. This is
because interdependencies cause “informational overlap”. To put it more precisely: If we only consider \( m_2 \) and \( m_3 \), the joint utility of knowing both \( m_2 \) and \( m_3 \) is smaller than the sum of the individual utility values due to knowing \( m_2 \) (or \( m_3 \)) and its interdependency with \( m_1 \) (or \( m_2 \)). If \( m_2 \) and \( m_3 \) interdepend, \( m_2 \) provides information about \( m_3 \) – and vice versa. The only exception is if \( m_2 \) and \( m_3 \) are (statistically) independent of each other. In this case, the joint utility equals the sum of both individual utility values. If we consider all non-selected measures, the joint interdependency-induced utility of \( m_2 \) and \( m_3 \) is smaller than the sum of the individual interdependency-induced utility values of \( m_2 \) and \( m_3 \). For each non-selected measure, the strongest interdependency will be used to estimate its value. The only exception is if \( m_2 \) and \( m_3 \) are independent of all non-selected measures or interdepend with disjoint subsets of non-selected measures. With many measures at hand, this is rather unlikely. Hence, the marginal utility of selecting \( m_2 \) and \( m_3 \) (two measures) compared to selecting \( m_1 \) (one measure) is smaller than (and exceptionally equal to) the marginal utility of selecting \( m_1 \) (one measure) compared to selecting zero measures. This holds for any number of measures. Hence, the more measures are selected – i.e. the higher \( x \) is –, the higher is the joint informational utility and the less is the marginal utility. In mathematical terms, \( U_{\text{info}}(x) \) increases \((\partial(U_{\text{info}}(x))/\partial x \geq 0)\) and strictly convex \((\partial^2(U_{\text{info}}(x))/\partial x^2 \leq 0)\). If we neglect the discussed exceptions and treat \( U_{\text{info}}(x) \) as strictly increasing and concave, it may be formalized in a simplifying manner as follows:

\[
U_{\text{info}}(x) = x^\alpha A \text{ with } \alpha \in [0;1] \text{ and } A \in IR^+
\] (2)

Selecting no measures provides no information \((U_{\text{info}}(0) = 0)\), whereas – according to \( A.I \) – selecting all measures provides complete information \((U_{\text{info}}(1) = A)\). The constant \( A \) represents the decision makers’ present-value monetary equivalent of complete information, that is, the amount of money they are willing to pay at the moment of measure selection for complete information during the planning horizon, i.e. as long as the selected measures are in use. Reasoning from an informational perspective, \( A \) represents the value of information by itself. It does not incorporate payments e.g. for data collection. The transformation into monetary units enables to integrate the economic perspective later. The diminishing marginal utility, which was introduced above and is caused by a higher fraction of measures, is formalized by the fact that the exponent \( \alpha \) is restricted to \([0;1]\). This also excludes the case where all measures interdepend perfectly, which would lead to a non-realistic course of \( U_{\text{info}}(x) \). A value of \( \alpha \) close to 0 is appropriate if all preselected measures interdepend rather strongly. Therefore, very few measures already create almost complete information. A value of \( \alpha \) close to 1 is appropriate if all measures are rather independent, that is, the marginal utility is rather constant. A mean value of \( \alpha \) indicates that the measures split into several groups with strong intra-group and weak inter-group interdependencies. The higher the value of \( \alpha \), the more (and the smaller) groups tend to exist.

The objective function’s second component represents the disutility created by information complexity \( D_{\text{info}}(x) \). It intuitively holds that the more measures are selected, the more complex is it to cognitively process them. Mathematically spoken, \( D_{\text{info}}(x) \) increases with \( x \). According to cognitive sciences (e.g. Duncan 1980, Miller 1956), the amount of information becomes overproportionally more complex when the fraction of measures increases. Hence, a higher fraction \( x \) is also characterized by an increasing marginal disutility with respect to \( D_{\text{info}}(x) \). In summary, \( D_{\text{info}}(x) \) is strictly increasing \((\partial(D_{\text{info}}(x))/\partial x > 0)\) and strictly convex \((\partial^2(D_{\text{info}}(x))/\partial x^2 > 0)\). This can be formalized as follows:

\[
D_{\text{info}}(x) = x^\beta B \text{ with } \beta \in [1;\infty[ \text{ and } B \in IR^+
\] (3)

Selecting no measures does not lead to complexity \((D_{\text{info}}(0) = 0)\), whereas – according to \( A.I \) – selecting all measures leads to highest complexity \((D_{\text{info}}(1) = B)\). The constant \( B \) represents the decision makers’ present-value monetary equivalent of understanding complete information during the planning horizon. The increasing marginal disutility, which was introduced above and is caused by a higher fraction of measures, is formalized by the fact that the exponent \( \beta \) is restricted to \([1;\infty[\). Its value depends on the decision makers’ and employed reporting tools’ ability of coping with
information complexity\(^1\). A value close to 1 is appropriate if the decision makers already have serious problems with processing few measures and/or the employed reporting tools are restricted to a few measures. The higher the value of \(\beta\), the less decision makers are susceptible to information complexity and/or the more powerful are the employed reporting tools. Based on (1) to (3), the optimization model is as follows:

\[
\text{Maximize } U_{\text{info,net}}(x) = U_{\text{info}}(x) - D_{\text{info}}(x) = x^\alpha A - x^\beta B \\
\text{w. r. t. } x \in [0;1]
\]

(4)

A mathematical analysis shows that \(U_{\text{info,net}}(x)\) strictly increases until \(x^*_{\text{info}} = \frac{(A \cdot \alpha)/(B \cdot \beta)}{1/(\beta - \alpha)}\). Up to that fraction, each additional measure provides more additional information than it causes additional complexity. Beyond, \(U_{\text{info,net}}(x)\) strictly decreases. Each additional measure then causes more additional complexity than it provides additional information. As \(x\) is restricted to [0;1], the optimal fraction is \(x^*_{\text{info}} = \min\{x^*_{\text{info}}, 1\}\). Due to the concave course of \(U_{\text{info,net}}(x)\), a border solution such as \(x^*_{\text{info}} = 1\) only occurs on rare occasions (see below).

Two interesting questions are: How is the decision makers’ attitude towards complete information and highest complexity reflected in \(A\) and \(B\)? How do both parameters ceteris paribus affect the course of \(U_{\text{info,net}}(x)\) and the position of \(x^*_{\text{info}}\)? The following case differentiation is also depicted in Figure 1. If \(A = B\), complete information creates as much utility as highest complexity creates disutility. Decision makers then are indifferent between making decisions based on zero measures or based on all preselected measures. The optimal fraction is \(x^*_{\text{info}} = (\alpha/\beta)^{1/(\beta - \alpha)}\) and only depends on \(\alpha\) and \(\beta\). If \(A < B\), complete information creates less utility than highest complexity creates disutility. Decision makers prefer making decisions based on zero measures to making decisions based on all measures.

The optimal fraction \(x^*_{\text{info}}\) is ceteris paribus smaller than in the first case. If \(A > B\), complete information creates more utility than highest complexity creates disutility. Decision makers prefer making decisions based on all measures to making decisions based on zero measures. The optimal fraction is \(x^*_{\text{info}}\) is ceteris paribus higher than in the first case. For certain constellations of \(\alpha\) and \(\beta\) (see Figure 1 on the right), \(U_{\text{info,net}}(x)\) could have its maximum \(x^*_{\text{info}}\) outside the interval [0;1]. With \(x\) being restricted to this interval, the optimal fraction then is \(x^*_{\text{info}} = 1\). This is the only case where it may be informationally optimal to select all measures.

### 3.2 An extended model for the informational and the economic perspective

In reality, information is not for free. Hence, it is necessary to integrate an economic perspective. In order to support decision makers, measures need to be compiled into reporting tools (e. g. management cockpits and dashboards). These need to be customized and maintained during their time in use. Abstracting from fixed costs, this leads to one-time and continuous payments. Both create economic disutility and influence measure selection. Thus, there is a joint informational and economic trade-off.

The question is: Up to which optimal fraction of measures does the additional informational net utility justify the additional economic disutility due to higher present-value payments for customization and maintenance? The extended model additionally relies on the following assumptions:

**A.4:** All preselected measures are implemented and their values can be extracted automatically from the respective application systems. The consolidated DSS will be connected to the existing application systems.

**A.5:** \(D_{\text{eco}}(x)\) is the economic disutility due to the present-value payments for customizing and maintaining reporting tools. It is a function of \(x\) and can be forecast ex ante.

---

\(^1\) To simplify matters, \(\beta\) is viewed as average value of how well decision makers/reporting tools are able to cope with information complexity. Of course, it can be further refined with respect to different groups/types or even individual decision makers/reporting tools.
Mathematically spoken, is, the higher according to Axson (2007), in average 15 to 20% of the selected measures will have to be changed. These arise e.g. from updating ETL procedures, assuring data quality, or changing selected measures.

...the reporting tool. This also applies to the present-value payments for maintaining reporting tools.

...systems integration and data collection need not be considered. The more measures are selected, the more time-consuming – and expensive – is it to initially customize reporting tools. Imagine the... customers first year, 10 to 15% in the following years. Hence, the more measures are selected – that is, the higher x is –, the higher is the economic disutility and the higher is the marginal disutility.

Mathematically spoken, $D_{\text{econ}}(x)$ is strictly increasing ($\partial(D_{\text{econ}}(x))/\partial x > 0$) and strictly convex ($\partial^2(D_{\text{econ}}(x))/\partial x^2 > 0$). This may be formalized as follows:

$$D_{\text{econ}}(x) = x^\gamma C \quad \gamma \in [1;\infty] \quad \text{and} \quad C \in IR^+$$

Selecting no measures does not lead to payments ($D_{\text{econ}}(0) = 0$), whereas selecting all measures leads to highest payments ($D_{\text{econ}}(1) = C$). The constant $C$ represents the highest amount of present-value payments due to customization and maintenance of reporting tools. The increasing marginal disutility, which was introduced above and is caused by a higher fraction of measures, is formalized by the fact that the exponent $\gamma$ is restricted to $[1;\infty]$. A value close to 1 is appropriate if a small fraction of measures already leads to high payment and each measure causes approx. the same marginal disutility. The higher $\gamma$ is, the less payments and marginal disutility causes a small fraction of measures and the higher is the marginal disutility of higher fractions. Based on (4) to (6), the extended optimization model is as follows:

Maximize $U_{\text{info+econ,net}}(x) = U_{\text{info}}(x) - D_{\text{info}}(x) - D_{\text{econ}}(x) = x^\alpha A - x^\beta B - x^\gamma C$

w. r. t. $x \in [0;1]$  

Although there is no general algebraic solution, the course of $U_{\text{info+econ,net}}(x)$ and the position of $x_{\text{info+econ}}^{\text{opt}}$ can be discussed with respect to the component functions (see Figure 2). As $U_{\text{info,net}}(x)$ is concave and $D_{\text{econ}}(x)$ is convex, $U_{\text{info+econ,net}}(x)$ is concave with one global maximum at $x_{\text{info+econ}}^{\text{opt}}$. As $U_{\text{info+econ,net}}(x)$ equals $U_{\text{info,net}}(x)$ diminished by $D_{\text{econ}}(x)$, the joint informational and economic optimum $x_{\text{info+econ}}^{\text{opt}}$ is smaller than or equal to $x_{\text{info}}^{\text{opt}}$, that is, $x_{\text{info+econ}}^{\text{opt}} \in [0;\text{x}_{\text{info}}^{\text{opt}}]$. This is reasonable because $x_{\text{info}}^{\text{opt}}$ is determined on the assumption that information has no price. If $D_{\text{econ}}(x)$ is close to zero – e.g. for large $\gamma$ and/or very small $C$ –, measures can be selected almost negligible of negative economic

Figure 1. Exemplary courses of $U_{\text{info}}(x)$, $D_{\text{info}}(x)$, and $U_{\text{info,net}}(x)$

On the assumptions $A.1$ to $A.5$, the informationally and economically optimal fraction $x_{\text{info+econ}}^{\text{opt}}$ can be determined by optimizing the difference between $U_{\text{info,net}}(x)$ and $D_{\text{econ}}(x)$. This difference is also called joint informational and economic net utility $U_{\text{info+econ,net}}(x)$. The objective function is given by:

$$U_{\text{info+econ,net}}(x) = U_{\text{info,net}}(x) - D_{\text{econ}}(x) = U_{\text{info}}(x) - D_{\text{info}}(x) - D_{\text{econ}}(x) = \max!$$

(5)

To formulate the extended optimization model, $D_{\text{econ}}(x)$ is examined. According to $A.4$, payments for systems integration and data collection need not be considered. The more measures are selected, the more time consuming – and expensive – is it to initially customize reporting tools. Imagine the selected measures had to be integrated into a dashboard. If only one measure is selected, the dashboard can be customized easily. If two measures are selected, an overall layout is more difficult (but still easy) to find. The more measures are selected, the overproportionally more time-consuming – and expensive – is it to find an adequate overall layout. This includes choosing among different visualization elements, adapting their size, trying different layouts, or – in the worst case – changing the reporting tool. This also applies to the present-value payments for maintaining reporting tools. These arise e.g. from updating ETL procedures, assuring data quality, or changing selected measures. According to Axson (2007), in average 15 to 20% of the selected measures will have to be changed during the first year, 10 to 15% in the following years. Hence, the more measures are selected – that is, the higher $x$ is –, the higher is the economic disutility and the higher is the marginal disutility.
effects. Then $x_{\text{opt info+econ}}$ and $x_{\text{opt info}}$ are approx. equal. If $D_{\text{info}}(x)$ is close to zero – e. g. for large $\beta$ and/or very small $B$ –, decision makers are hardly susceptible to information complexity and/or powerful reporting tools are employed. Then an approx. solution is $x_{\text{opt info+econ}} \approx \left( \frac{A \cdot \alpha}{C \cdot \gamma} \right)^{1/(\gamma-\alpha)}$. In this case, analogous to above, the relationship between the decision makers’ subjective attitude towards complete information and highest (present-value) payments can be analyzed with respect to $A$ and $C$.

Concluding, the optimization model allows determining the optimal fraction of measures to be chosen from a preselected set of thematically appropriate measures. Accordingly, those measures are selected that together create the highest informational utility. The model integrates an informational and an economic perspective. The former reflects the decision makers’ attitude towards information and information complexity. The latter considers present-value payments for customizing and maintaining reporting tools. It could be shown that, in general, the selection policies of all-or-none or the-more-the-better, which are often implemented in business practice, are reasonable neither from an informational nor from a joint informational and economic perspective. What makes sense instead is a differentiated and balanced selection of measures.

4 EVALUATION

4.1 Applying the optimization model in business practice

The optimization model was developed in the context of a project at the German business-to-business sales organization of a global electronics and electrical engineering company. As there were only few measures, the model could be applied in a discretized form. If there had been very much measures, we would have had to evaluate a manageable subset in order to infer the continuous functions introduced above. Due to confidentiality, all data is anonymized and modified. Yet the principal results still hold.

The project’s overall goals were to better support the sales force, to reduce IT operation costs, and to modernize sales reporting. As for the first two goals, the company decided to introduce a single CRM system and to harmonize the application landscape, which consisted of more than one hundred division-specific legacy systems. The reporting mainly consisted of financial and lagging measures such as volume of sales and price margin. It was to be modernized with respect to non-monetary and leading measures. Our task was to structure the salespeople’s IR into CSFs and to select appropriate measures. At first, candidate CSFs were identified by explorative interviews with sales managers and senior members of the CRM board. Sales managers had usually worked as sales representatives for several years and were supposed to provide valuable hints with respect to IR and sales reporting. They were selected by reputational methods (Knoke 1993). For each candidate CSF, several items were identified and compiled into a five-point Likert scale-based questionnaire. After a pretest, the questionnaire was presented to 25 sales managers (the amount was restricted by the project budget).

All in all, CSFs were identified for 3 perspectives, namely organizational structures and processes (e. g. long-term customer care, cross-divisional cooperation), salespeople’s skills and knowledge (e. g. with respect to installed base and competitors’ portfolios), and IT functionality (e. g. integration with office communication software, IT-based planning of sales calls).

Figure 2. Exemplary course of $U_{\text{info+econ,net}}(x)$ for $A = B = C$
The CSF “cross-divisional cooperation” will serve as example. Together with the sales managers, we retrieved 8 thematically appropriate measures. These were: fraction of converted leads from other divisions (\% \_leads), average overall time spent on creating leads for other divisions (\_T \_leads), average time spent on creating one lead (\_T \_lead), number of trainings on other divisions’ portfolios (\# \_trainings), number of meetings with colleagues from other divisions (\# \_meetings), number of sales calls with colleagues from other divisions (\# \_calls), number of shared customers (\# \_customers), number of bids for customers of other divisions (\# \_bids). All measures had existed for several years and were reported monthly on a sales manager’s granularity.

First, we assessed informational utility $U_{\text{info}}(x)$. The exponent $\alpha$ – which indicates how the measures interdepend – was operationalized based on Pearson’s correlation coefficient. That is, we treated the interdependencies as pairwise, symmetric, and linear. We accepted this simplification because the correlation coefficient is an intuitive, widely used, and relatively easy-to-compute measure. Moreover, linear interdependencies are often considered as sufficiently good approximations for many economic settings (Edwards 1976). This turned out to be useful because, due to missing hierarchical and logic structures, the existence and strength of interdependencies among non-monetary and leading measures often need to be assessed empirically e. g. by interviewing domain experts and analyzing historical data (Küpper 2005). Let $M = \{m_1, m_2, \ldots, m_n\}$ comprise $n$ preselected and thematically appropriate (metrically scaled) measures between some of which there are meaningfully interpretable interdependencies/correlations. We considered absolute values as the correlation coefficient’s algebraic sign only indicates direction, not strength. The values are represented as $n \times n$-matrix $CM$ where $c_{ij}$ indicates how strong $m_i$ and $m_j$ correlate and where $c_{ij} = 0$ if $m_i$ and $m_j$ are statistically independent or if their interdependency/correlation cannot be meaningfully interpreted ($0 \leq i, j \leq n$, $i \neq j$). The individual correlations of a measure $m_i$ equal the $i$-th column vector of $CM$.

$$c_i = (c_{i1}, c_{i2}, \ldots, c_{in})^T$$

The joint correlations of multiple selected measures $m_1, \ldots, m_n$ (without loss of generality) are also represented as vector $c_{1,\ldots,n}$. The elements of all selected measures are $1$ (perfect autocorrelation). The element of each non-selected measure indicates the strongest correlation with any measure selected so far. This is reasonable because if decision makers want to estimate the value of a non-selected measure $m_j$, they will reasonably revert to the selected measure that correlates most strongly with $m_j$.

$$c_{1,\ldots,n} = (\max\{c_{11}, \ldots, c_{1n}\}, \max\{c_{12}, \ldots, c_{2n}\}, \ldots, \max\{c_{1n}, \ldots, c_{nn}\})^T$$

The concept of joint correlations enables to formalize a discretized informational utility as function of $x$. We need the highest joint correlation of $x \cdot n$ measures. It is determined by calculating the highest scalar product value $\langle l, c' \rangle$ where $l$ is an $n$-vector $(1, 1, \ldots, 1)^T$ and $c'$ is the joint correlations vector of $x \cdot n$ arbitrary measures. Dividing the scalar by $n$ normalizes it to $[0;1]$. This operationalization can be interpreted as a monetized mean absolute correlation.

$$U_{\text{info}}(x) = \left[ \max \{ \langle l, c' \rangle \mid x \cdot n \text{ measures are selected} \} / n \right] \cdot A$$

After CSF analysis, some sales managers from the CRM board were asked to (subjectively) judge which interdependencies between measures are meaningfully interpretable as regards the sales domain. The strength of these interdependencies was calculated by means of absolute correlation values based on historical data. Table 2 shows the results with the light grey cells marking excluded interdependencies/correlations. Concerning the value of $A$ – which represents the sales managers’ present-value monetary equivalent of complete information – we asked each sales manager how many daily rates he would pay for having complete information on cross-divisional cooperation during the planning horizon. We obtained an average of 10 daily rates, which we multiplied with the sales

---

2 In the CRM context, a lead represents a hint with low degree of maturity from inside or outside one’s division that refers to a potential customer or project opportunity.
Table 2. Absolute correlation coefficients values of the preselected measures

<table>
<thead>
<tr>
<th>%_leads</th>
<th>Ø_T_leads</th>
<th>Ø_T_lead</th>
<th>%_trainings</th>
<th>%_meetings</th>
<th>%_calls</th>
<th>%_customers</th>
<th>%_bids</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>0.43</td>
<td>0.83</td>
<td>0.67</td>
<td>0.00</td>
<td>0.42</td>
<td>0.34</td>
<td>0.96</td>
</tr>
<tr>
<td>0.43</td>
<td>1.00</td>
<td>0.34</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.12</td>
<td>0.54</td>
</tr>
<tr>
<td>0.83</td>
<td>0.34</td>
<td>1.00</td>
<td>0.67</td>
<td>0.00</td>
<td>0.00</td>
<td>0.36</td>
<td>0.24</td>
</tr>
<tr>
<td>0.56</td>
<td>0.00</td>
<td>0.67</td>
<td>1.00</td>
<td>0.00</td>
<td>0.25</td>
<td>0.41</td>
<td>0.21</td>
</tr>
<tr>
<td>0.54</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.38</td>
<td>0.25</td>
<td>0.12</td>
</tr>
<tr>
<td>0.42</td>
<td>0.00</td>
<td>0.00</td>
<td>0.25</td>
<td>0.38</td>
<td>1.00</td>
<td>0.74</td>
<td>0.73</td>
</tr>
<tr>
<td>0.34</td>
<td>0.12</td>
<td>0.36</td>
<td>0.41</td>
<td>0.25</td>
<td>0.74</td>
<td>1.00</td>
<td>0.58</td>
</tr>
<tr>
<td>0.96</td>
<td>0.54</td>
<td>0.24</td>
<td>0.21</td>
<td>0.12</td>
<td>0.73</td>
<td>0.58</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 3. (Dis-)Utility values for the CSF “cross-divisional cooperation”

<table>
<thead>
<tr>
<th>#_leads</th>
<th>Ø_T_leads</th>
<th>#_trainings</th>
<th>#_meetings</th>
<th>#_calls</th>
<th>#_customers</th>
<th>#_bids</th>
</tr>
</thead>
<tbody>
<tr>
<td>176</td>
<td>119</td>
<td>134</td>
<td>102</td>
<td>124</td>
<td>172</td>
<td>186</td>
</tr>
<tr>
<td>0.21</td>
<td>0.52</td>
<td>0.78</td>
<td>0.62</td>
<td>0.74</td>
<td>0.85</td>
<td>0.88</td>
</tr>
<tr>
<td>0.12</td>
<td>0.36</td>
<td>0.41</td>
<td>0.25</td>
<td>0.74</td>
<td>1.00</td>
<td>0.58</td>
</tr>
<tr>
<td>0.96</td>
<td>0.54</td>
<td>0.24</td>
<td>0.21</td>
<td>0.12</td>
<td>0.73</td>
<td>0.58</td>
</tr>
</tbody>
</table>

managers’ average daily rate of 750 € and their overall number – there were 50. We finally obtained A = 375,000 €. As for informational disutility D_{info}(x), the sales managers received sample reports. Each contained a different amount of measures, but had exactly the same layout as the reports that were planned to be finally used. The sales managers’ task was to entirely understand the reports. For each amount of measures, we logged the time. In order to determine the value of B, we used the average value for 8 measures – which was 2.0 hours. We normalized it with respect to the sales managers’ average daily working time – which was 9 hours. Then, we multiplied it with the sales managers’ average daily rate and their overall number. As the report was planned to be presented monthly and the planning horizon was 3 years, we calculated the annual payments and the corresponding present value with an interest rate of 10%. Assuming that the managers had to try to understand the report completely anew each time they received it, we obtained B = 273,554 €. We used the other time values for approximating β. The sales managers coped well with a low number of measures, but had problems with more than approx. 4–5 measures. So we obtained β = 2.8. The economic disutility D_{econ}(x) was calculated based on Boehm’s widespread cost estimation model CoCoMo (1981). Together with the company’s DSS experts, we parameterized the estimation model as PM=2.94-0.20-LOC^{1.2} where PM and LOC denote person months and thousand lines of code respectively. The present-value effort for customizing a report with one measure and maintaining it during the planning horizon was estimated equivalent to 1.250 LOCs. With the DSS experts’ average daily rate of 400 € and 20 working days per month, we obtained C = 74,553 € and γ = 1.2.

On this basis, we determined the optimal fraction of measures x_{opt \_info+econ} by computing the joint informational and economic net utility U_{info+econ}(x) (see Table 3 and Figure 3). If we had only considered the informational perspective, the highest informational net utility would have resulted from x_{opt \_info} = 0.5. We would have selected the 4 measures with the highest informational utility, i.e. %_leads, Ø_T_leads, #_meetings, and #_customers. As we also took on an economic perspective, the highest joint net utility resulted from x_{opt \_info+econ} = 0.375. We selected the 3 measures with the highest informational utility, i.e. %_leads, #_customers, and #_bids. Here, it becomes obvious that measures can be discarded if the number of selected measures changes.

We applied the same procedure to the other CSFs. Advantageously, this caused considerably less effort because D_{info}(x) and D_{econ}(x) needed not to be determined anew. U_{info}(x) could be calculated based on historical data so that the sales managers’ expertise was only required for preselecting thematically appropriate measures and identifying meaningfully interpretable interdependencies. In sum, we modernized the company’s sales reporting by identifying CSFs, by integrating non-monetary and leading measures, and by significantly reducing the overall number of measures.

<table>
<thead>
<tr>
<th>No. of selected measures</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction (x)</td>
<td>0</td>
<td>0.125</td>
<td>0.25</td>
<td>0.375</td>
<td>0.5</td>
<td>0.625</td>
<td>0.75</td>
<td>0.875</td>
<td>1</td>
</tr>
<tr>
<td>(U_{info}(x)) [€]</td>
<td>0</td>
<td>217,969</td>
<td>281,719</td>
<td>310,781</td>
<td>337,500</td>
<td>352,969</td>
<td>365,156</td>
<td>373,125</td>
<td>375,000</td>
</tr>
<tr>
<td>(D_{info}(x)) [€]</td>
<td>0</td>
<td>810</td>
<td>5,640</td>
<td>17,552</td>
<td>39,279</td>
<td>73,368</td>
<td>122,240</td>
<td>188,220</td>
<td>273,554</td>
</tr>
<tr>
<td>(U_{info+econ}(x)) [€]</td>
<td>0</td>
<td>217,159</td>
<td>276,079</td>
<td>293,229</td>
<td>298,221</td>
<td>279,601</td>
<td>242,916</td>
<td>184,905</td>
<td>101,446</td>
</tr>
<tr>
<td>(D_{econ}(x)) [€]</td>
<td>0</td>
<td>6,148</td>
<td>14,125</td>
<td>22,978</td>
<td>32,451</td>
<td>42,415</td>
<td>52,789</td>
<td>63,515</td>
<td>74,553</td>
</tr>
<tr>
<td>(U_{info+econ+meta}(x)) [€]</td>
<td>0</td>
<td>211,011</td>
<td>261,954</td>
<td>270,251</td>
<td>265,770</td>
<td>237,185</td>
<td>190,127</td>
<td>121,390</td>
<td>26,893</td>
</tr>
</tbody>
</table>
4.2 Checking the optimization model against the PMS requirements

The optimization model particularly aims at closing the primary research gap with respect to clarity (R.2) and interdependencies (R.5) (see section 2). By requiring to select a manageable amount of measures, the model addresses clarity in an explicit manner. Informational disutility expresses the decision makers’ and reporting tools’ ability of coping with informational complexity and is contrasted to informational utility. Thereby, we make sure that the whole amount of information does not become too complex and remains manageable for the decision makers. Interdependencies are also addressed explicitly. This is because informational utility uses interdependencies in order to express how “much” information measures provide about one another. The example showed – albeit in a simplifying manner – how an interdependency-based informational utility can be operationalized for non-monetary and leading measures with Pearson’s correlation coefficient. Moreover, the model makes the process of measure selection more intersubjectively comprehensible (R.4). Although most parameters cannot be determined without subjective influences, just the fact that it is clear how they are formally linked increases intersubjectivity. Decision makers do not only participate by means of explorative interviews, but also in a structured manner by validating interdependencies and estimating model parameters (R.6). Concluding, the model does not only address the primary research gap, but also ameliorates the other PMS requirements.

5 SUMMARY AND FUTURE RESEARCH

An optimization model has been proposed that helps to determine which and how many measures should be selected from a set of thematically appropriate measures in order to monitor specific fields of action. Informational and economic objectives are considered. That is (statistic) interdependencies among measures, decision makers’ and reporting tools’ ability of coping with information complexity as well as payments for customizing and maintaining reporting tools influence measure selection. The model’s principle applicability was shown with a real-world example. Admittedly, business practice entails problems (e. g. estimation of costs, data collection in complex social contexts, decision makers’ partial inability of unambiguously specifying IR) that make it hard to achieve truly optimal solutions. In order to cope with some of these problems, it may be useful for companies to implement the model stepwise and to involve operating staff in data collection. Nevertheless, the proposed model is a first step towards a more well-founded measure selection. It will be subject to future research:

1. The optimization model is applied to one field of action a time. Several fields of action can only be addressed successively and isolated. The fact that measures may be thematically appropriate for more than one field of action is not considered. Hence, an integrated perspective is desirable and should be added.
2. So far, only measures from existing application systems are considered. On the one hand, this is reasonable as in many companies more measures exist than any decision maker can ever analyse. On the other hand, positive effects of innovative measures are neglected and need to be integrated.

3. (Data-driven) DSS and data warehouses do not only comprise measures, but also other master data for evaluation. In order to deal with their full scope, the model needs to be complemented e.g. by an approach which assesses relevant dimensions and dimension elements.

4. Although the model has been employed successfully with real-world data, empirical evidence is missing with respect to whether its recommendations actually improve decision quality. It would be insightful and strengthen evaluation to conduct respective empirical studies.

References

Miller, G. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. The Psychological Review, 63 (2), 81-97.
Justifying Design Decisions with Theory-based Design Principles

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0290.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Design Science, Design research, Theory Building, Conceptual modelling</td>
</tr>
</tbody>
</table>
JUSTIFYING DESIGN DECISIONS WITH THEORY-BASED DESIGN PRINCIPLES

Michael Schermann, Chair for Information Systems, Technische Universität München, Boltzmannstrasse 3, 85748 Garching, Germany, michael.schermann@in.tum.de
Andreas Gehlert, Software Systems Engineering, University of Duisburg-Essen, Schützenbahn 70, 45117 Essen, Germany, andreas.gehlert@sse.uni-due.de
Helmut Krcmar, Chair for Information Systems, Technische Universität München, Boltzmannstrasse 3, 85748 Garching, krcmar@in.tum.de
Klaus Pohl, Software Systems Engineering, University of Duisburg-Essen, Schützenbahn 70, 45117 Essen, Germany, klaus.pohl@sse.uni-due.de

Abstract

Although the role of theories in design research is recognized, we show that little attention has been paid on how to use theories when designing new artifacts. We introduce design principles as a new methodological approach to address this problem. Design principles extend the notion of design rationales that document how a design decision emerged. We extend the concept of design rationales by using theoretical hypotheses to support or object to design decisions. At the example of developing a new conceptual modeling grammar we demonstrate two main benefits of using design principles. First, the link between theory and design decision enables the design researcher to reason about the resulting behavior of the IT artifact prior to instantiation. Second, design principles allow deducing empirically testable hypotheses to foster the rigorous evaluation of IT artifacts.

Keywords: design principles, design rationale, design theories, theories, design research
INTRODUCTION

Often, design research results are criticized as being subjective and anecdotal in nature. Design researchers need to argue, how the IT artifact relates to hypotheses about the nature and characteristics of the artifact and its proposed benefits. Thus, theories are the cornerstone of improving the intersubjectivity of design artifacts (Walls et al., 1992). Figure 1 shows a bilateral relation between theory and design (Nunamaker and Chen, 1991). According to relation 1 “Contribution” in Figure 1, theory may follow design. Here, design research contributes to the IS knowledge base by initiating theory-building based on existing design artifacts. Walls et al. (1992) and others introduced the concept of design theories to specify the theoretical contribution of design research to the IS knowledge base (e.g. Markus et al., 2002; Walls et al., 1992).

Figure 1: The elements of IS design theories based on Nunamaker and Chen (1991) and March and Smith (1995)

Theories can also be used to justify design decisions and subsequently anticipate the behavior of the IT artifact (relation 2 “Justification” in Figure 1). This relation is based on the insight that any design artifact cannot work if it does not comply with the laws embodied in the underlying theory (Walls et al., 1992). In addition, the artifact may be significantly improved if the underlying theory is explicit (van Aken, 2004, p. 228).

Although this second relation is widely recognized on a general level (Gregor and Jones, 2007; Venable, 2006; Walls et al., 1992), we show that little attention has been paid to the methodological guidance on how to use theories when designing artifacts. This paper aims to fill this gap by answering the following research question: How can theoretical knowledge be used in the design process? To answer this question, we argue that the iterative nature of design research requires a particular focus on the specific decisions made during the design research process and the underlying arguments which finally led to the design artifact (Hevner et al., 2004). We apply and extend the concept of design rationale to structure, justify, and document the decisions made in the iterations of the design research process.

Design rationales can be used to justify design decisions during the construction of the artifact by documenting design alternatives and the reasons for selecting or rejecting them (Regli et al., 2000). Because of the transparency of the decision making process, design rationales are valuable sources to stimulate theory-building, i.e., they facilitate design researchers in specifying the first relation between design and theory (see 1 in Figure 1). However, to support the theoretical justification of design decisions (see 2 in Figure 1), we extend the concept of design rationale by introducing theory-based arguments. In doing so, the design decision can be traced back to empirically testable hypotheses, which enable the design researcher to reason about the effects of the IT artifact as well as to reason about the artifact’s possible side-effects (Weber, 1987). We argue that such theory-based design decisions form a new type of IS design knowledge, which we call design principles. A design principle is a coherent claim of utility (Markus et al., 2002) and consists of a specific design problem, a solution to the problem and the trace path of the discussion leading from theory to the solution.
The link between theory and design decision enables design researchers to reason about the resulting behavior of the IT artifact prior to instantiation. Furthermore, design principles allow deducing empirically testable hypotheses to foster the rigorous evaluation of IT artifacts.

The remainder of the paper is organized as follows: In the next section, we review existing work on the role of theories in design research. We show that little attention has been paid on how to incorporate theories when designing new artifacts. Next, we analyze the notion of theory in IS research and the concept of design rationales as the basic elements of our approach. Then, we show how to develop design principles by applying theoretical hypotheses as justificatory knowledge in design rationales. At the example of developing a new conceptual modeling grammar we demonstrate the utility of our approach. Overall, we contribute to the methodological foundations of IS design research by explaining how theories can be used to justify design decisions. The paper closes with a discussion of the implication of this approach and an outline for further research.

This paper is of exploratory and conceptual nature. Hence, we provide argumentative support when answering our research questions. However, we base our arguments upon available empirical and conceptual research results as well as on the results of a previous paper (Gehlert et al., 2009). This paper extends the results of the previous one by introducing the notion of design principles and by an in-depth analysis of the relation between these design principles and theories.

2 THEORIZING IN DESIGN RESEARCH

In this section, we summarize the current debate on the role of theories in IS design research. The important role of theories in design research has been addressed by a number of authors (Gregor and Jones, 2007; Markus et al., 2002; Schermann et al., 2007; Verschuren and Hartog, 2005; Walls et al., 1992). Gregor and Jones (2007) distinguish two roles of theories in IS design research: First, theories can be the output of design research: “The distinguishing attribute of theories for design and action is that they focus on ‘how to do something’. They give explicit prescriptions on how to design and develop an artifact, whether it is a technological product or a managerial intervention” (Gregor and Jones, 2007, p. 313). Second, theories provide the foundation of justifying decisions during the design process and the subsequent features of the artifact: “[..] we argue that these theories are a linking mechanism for a number, or all, of the other aspects of the design theory. […] Theories might come from natural science, social science […], other design theories, practitioner-in-use theories […], or evidence-based justification such as seen in medical research and action research.” (Gregor and Jones, 2007, p. 327)

Hevner et al. (2004) propose guidelines to determine the contribution of design research results to the IS knowledge base. Their guideline 3 requires that the “[…] utility, quality, and efficacy of [the] design artifact must be rigorously demonstrated” (Hevner et al., 2004, p. 83). In the line with this argument, Walls et al. (1992) argue, that design theories should propose empirically testable hypotheses to provide means for testing the utility of the design theory. Hence, the role of theories in IS design research is to facilitate the evaluation of design research results.

Furthermore, in their guideline 6, Hevner et al. (2004) characterize the design process as a search process: “The search process for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment” (Hevner et al., 2004). Hence, during the design process, the design researcher probably tries several means to solve the problem, receives unsatisfactory results, and consequently reshapes his or her approach. Hence, the design process can be characterized as an evolutionary process which should be governed by theories from the problem environment.

In sum, the role of theories in IS design research is threefold: theories are the output for the purpose of communicating and evaluating design knowledge, theories are the input for the construction of design artifacts, and theories shape the evolution of the IT artifact along the design process. Table 1 distinguishes two groups of elements. The first group consists of goals and requirements and the
subsequent design proposals that describe a solution, which matches the requirements and fulfils the goals. Overall, this group of elements describes “how to do something” (Gregor, 2006, p. 628). The design proposal can be characterized as the claim of the design researcher to achieve utility. Hence, the second group consists of elements which provide information that substantiate this claim in an auditable way (Walls et al., 1992). The approaches presented in Table 1 provide the necessary structure to document the results of design research.

Table 1: Approaches to design theories in IS design research

From a methodological point of view, the subsequent questions are how to proceed when incorporating theories in the design process and deducing testable hypotheses from the design results. Table 1 shows that the evolutionary aspect of design research has not been recognized until recently (Gregor and Jones, 2007; Schermann et al., 2007). The search process as described by Hevner et al. (2004) requires design researchers to document and trace their decisions and the corresponding justification across several iterations. Hence, design researchers need to track their decisions and the underlying linkages between informing theories, aspects of the design proposal, and resulting hypotheses (Gregor and Jones, 2007).

Surprisingly, little attention has been paid to developing methodological guidance on the use of theories in design research (Venable, 2006). We use the design research section on isworld.org (Vaishnavi and Kuechler, 2004) as example to clarify this position. In their extensive review of design research methodology, Vaishnavi and Kuechler (2004) characterize the role of theory in design research as informative: “[…] The design process, when interrupted and forced back […] contributes valuable constraint knowledge to the understanding of the always-incomplete-theories that abductively motivated the original design” (Vaishnavi and Kuechler, 2004). However, no further advice is given on how to proceed and document the learning process. How does the researcher document which design decision was grounded on which theoretical advice? Clearly, “more detailed and clear concepts would be helpful” to describe the use of theories in design research (Venable, 2006, p.1).

Overall, the general role of theory and its importance for rigorous design research has been widely recognized. However, existing approaches do not show how to proceed when making design decisions based on theory or deducing hypotheses from design proposals. In sum, our analysis shows that the required “linking mechanism” (Gregor and Jones, 2007) between goals, requirements, design proposals, underlying theories, and deduced hypotheses is largely unspecified.

3 FOUNDATIONS OF IS THEORIES AND DESIGN RATIONALES

In the previous section, we discussed that little methodological guidance exists on how to apply theoretical knowledge in design research. We suggest extending the concept of design rationales by enabling design researchers to develop a line of arguments leading from theory to an abstract design solution. In the following, we first clarify our notion of theories based on the typology provided by Gregor (2006). Second, we review the design rationale literature and select an appropriate approach for this paper. Both results are summarized as conceptual models, which will be then integrated in the next section.
3.1 The Notion of Theory in IS Research

The word theory is currently used in IS research to denote at least five different types of theoretical knowledge (Gregor, 2006). To integrate theories with design rationale approaches, we need to clarify the structure of such theories. A theory consists of a set of hypotheses (Popper, 2002). In general, a hypothesis is a correlation relationship based on two constructs (Balzer et al., 1987; March and Smith, 1995). The independent construct of the hypothesis is called cause while the dependent construct is called effect. According to the Structural Equation Modeling (SEM) approach, each construct is either represented by a set of indicators, which is called reflexive construct or is causing these indicators, which is then called a formative construct (Freeze and Raschke, 2007). The indicators operationalize the construct by providing the necessary means to measure the construct during theory testing. This notion of theory is summarized in the right area of Figure 2.

From our definition of theory we can draw three conclusions:

C1 Based on the definition of a construct, we conclude that two constructs with different sets of indicators are different; two constructs with the same sets of indicators are equivalent.

C2 Based on our definition of a hypothesis, it follows that two hypotheses are equivalent if they share the same constructs; otherwise the hypotheses are different.

C3 Based on our notion of theories, we conclude that theories with different sets of hypotheses are different and theories with the same sets of hypotheses are equivalent.

In sum, theories provide important information on cause-effect relations. However, we have shown that existing approaches to theorizing in IS research pay only little attention to the specific steps necessary to intertwine theorizing and design activities. We argue that the iterative nature of design research requires a particular focus on the specific decisions made during the design research process and the underlying arguments which led to the design artifact (Hevner et al., 2004). However, the challenge of documenting and justifying design decisions has been faced in the domains of software engineering and systems engineering, too. Thus, in the following section, we show that the notion of design rationales provide the foundation for developing theory-based design arguments, i.e., design principles.

3.2 Design Rationale Approaches

Research on design rationales is concerned with documenting and justifying particular decisions made during the design process (Rossi et al., 2004). The main goal of design rationale management is to document the relationship between the artifact, the underlying goals, the design proposal, and potential constraints (Louridas and Loucopoulos, 2000). This relationship can then be used to justify the design decisions which have led to the artifact. In doing so, the design rationale enables the design researcher to reason on his or her decisions at any given stage in the design process (Regli et al., 2000). For an extensive overview on existing approaches to rationale management see (Regli et al., 2000) and (Louridas and Loucopoulos, 2000).

Most of the approaches to design rationale management are based on the issue-based information systems (IBIS) approach (Kunz and Rittel, 1970). An issue in IBIS defines a certain problem that has occurred during the design phase. The design researcher can now establish various positions that may address the issue adequately. Then, different arguments can be made to either reject or substantiate a certain position. Finally, a position is being chosen and subsequently implemented in the design proposal (see the left area in Figure 2).

The following example explains IBIS: Let us assume we are in the process of constructing a new conceptual modeling grammar to resolve some issues with existing grammars, e.g. the Unified Modeling Language in a special problem domain (Object Management Group, 2005). The goal of the design project results in the requirement that the modeling grammar should enable its user to
understand the resulting models easily. During the design process the design researcher has to decide how to represent the concepts of the modeling grammar. An issue in this discussion could be: “Should we assign graphical symbols to the grammatical concepts?”

The design researcher is now facing two positions: “yes” and “no”. A supporting argument for the “yes” position can be found in the experience, that graphical models are commonly easier to understand. An argument objecting the use of graphical symbols could arise from the costs of developing and introducing an appropriate graphical user interface when the modeling grammar is implemented. After the arguments for all the issue’s position have been discussed, the designer needs to select a position to resolve the issue. As the experience of the design researcher favors a graphical notation, the position “yes” is chosen. This results not only in the benefits of achieving better model comprehension by having a graphical representation (desired effect) but also in the potential constraint to the design process of implementing a costly graphical user interface (undesired side-effect).

### 4 DEVELOPING THEORY-BASED DESIGN PRINCIPLES

The goal of using theories in design research is to justify design decisions. The link between theory and design decision enables the design researcher to reason about the resulting behavior of the IT artifact prior to instantiation. Furthermore, this link allows deducing empirically testable hypotheses to foster the evaluation of the IT artifact. So far, we have discussed that design rationale approaches provide the means for establishing a traceable link between design decisions and experiences of the designer. Before that, we have analyzed the structure of theoretical arguments, i.e., cause-effect-relations. In the following we show how to develop design principles by grounding design rationales in theoretical arguments. The main idea of developing theory-based arguments is to reinterpret the cause-effect relation in a hypothesis as a goal-means relation (March and Smith, 1995; Weber, 1987).

![Figure 2: Integration of Theories into IBIS](image)

To illustrate this reinterpretation, assume the following example: Larkin and Simon (1987) show that a diagrammatic notation (cause) results in a higher comprehension (effect) than textual notations. Inverting this relation and applying it to the construction of a modeling grammar allows the design researcher to formulate a theory-based argument: If the goal of a modeling grammar is to produce easily comprehensible models, one way to achieve it is to use a diagrammatic representation. However, this reinterpretation is possible only if the design researcher can actually influence the cause construct of the hypothesis, i.e., the design researcher is able to choose graphical symbols when developing the modeling grammar. In addition, there may be other mean–based on other hypotheses–which have the same effect.

In the context of the IBIS model, the starting point for developing theory-based arguments is the construct **Argument**. It provides the reasons for supporting a position or objecting to the position. The given set of arguments provides the foundation for making a decision with regard to an issue.
However, the argument itself does not explain why supporting or objecting a position resolves the issue, as long as the argument is not grounded in theoretical knowledge. In other words, if arguments are grounded in theoretical knowledge we do not only know, which arguments lead to the decisions made but also why these arguments contribute to the requirements of a design principle (\textcircled{1} in Figure 2).

When assigning a hypothesis to an argument, we implicitly build upon two assumptions (\textcircled{2} and \textcircled{3} in Figure 2):

A1 The cause construct of the hypothesis represents the position the argument refers to.

A2 The effect construct of the hypothesis represents the requirement the argument refers to via the position and issue.

These assumptions result from the reinterpretation of a hypothesis (cause-effect relations) as goal-means relations as described above. In this sense a requirement can be interpreted as (operational) goal to achieve. A position together with an issue represents possible solutions for this requirement (means). Consequently, the reinterpretation of a hypothesis must be represented by the position and requirements construct. Based on our notion of theory we know that the hypothesis’ constructs are uniquely defined by its indicators (see Conclusions C1). In the light of A1 and A2, this means that the position and the requirement must be expressed and measured by the same indicators as the respective cause and effects constructs from the theory. Otherwise the mapping between the theoretical constructs (cause and effect) and the design constructs (requirement and position) is inconsistent.

Whenever the design researcher grounds an argument in a hypothesis, i.e., whenever the link between argument and hypothesis is established (\textcircled{4} in Figure 2), the design researcher must map the hypothesis’ cause construct (\textcircled{2} in Figure 2) to the respective position and the hypothesis’ effect construct to the requirement (\textcircled{3} in Figure 2). The links \textcircled{2} and \textcircled{3} represent how the position and the requirement is interpreted, i.e., how the requirement and the position must be measured when evaluating the design principle.

Now we can deduce whether a theory-based argument supports a given position:

D1 The argument supports a position, if the selected position implements the cause construct and the hypothesis’ correlation is positive.

D2 The argument supports a position, if this position is not based on the cause construct and the hypothesis’ correlation is negative.

D3 In any other case, the argument objects to a position.

Let us return to the example of developing a modeling grammar and assume that the design researcher identified the following two hypotheses:

H1 If a graphical representation is used, the resulting models are easier to understand.

H2 If a modeling grammar has a graphical representation, then the automatic analysis of the models is more difficult.

Clearly, H1 has a positive correlation, whereas H2 has a negative correlation. As already explained above, the initial goal and thus the requirement of the modeling grammar is to enable its users to produce comprehensible models. The issue raised was whether to implement a graphical representation of the grammar. The available positions consist of “implementing a graphical representation” and “not implementing a graphical representation”.

Both positions refer to the cause construct of H1 and H2, the graphical representation. The positive correlation between the graphical representation and comprehension of H1 supports the position of “implementing a graphical representation”, because it advises the design researcher to use a graphical representation for the modeling grammar (D1). The negative correlation between graphical
representation and automatic analysis of H2 supports the position of “not implementing a graphical representation”, because it advises against implementing a graphical representation (D2).

After all arguments were put forward the design researcher needs to select one of the positions to resolve the issue. This resulting decision manifests how the requirement should be interpreted, because the effect constructs of the hypotheses, are now related to the initial requirement (\[ \in Figure 2 \]). In our example following H1 means to refine the requirement “model comprehension” to “model comprehension by persons” while using H2 means to refine the requirement to “model comprehension by machines”.

Two important implications for the design researcher can be drawn from this analysis:

1) We discussed above that the constructs of a hypothesis are defined by a set of indicators. Consequently, selecting a position results in the decision of interpreting the requirement in the way the effect construct of the underlying hypotheses is being interpreted. By choosing the position of “implementing a graphical representation” we select to interpret and measure comprehension in accordance to hypothesis H1.

2) Selecting a position is rarely straightforward. In realistic cases more than one position is supported by theory-based arguments. Consequently, choosing a particular position means rejecting potentially important theoretical arguments. These arguments represent known side-effects of the particular design decision. In the example of the modeling grammar, the design researcher now knows that implementing a graphical representation will impede any automatic analysis.

After linking hypotheses to IBIS, we can now refine our notation of a design principle. A design principle is a set of requirements, issues, positions, selected positions, arguments and hypotheses. Grounding the design principle’s arguments in hypotheses provides not only the theoretical rationale for choosing among potentially useful positions but also explains how the requirements and the selected positions should be measured. In addition, the design principle shows possible drawbacks of the selected position as these drawbacks are represented by supporting arguments for all positions, which have not been chosen.

5 THEORY-BASED DESIGN PRINCIPLES FOR DESIGNING CONCEPTUAL MODELING GRAMMARS

To demonstrate the utility of our approach, we extend the example of constructing a new modeling grammar. The overall goal is to “construct a conceptual modeling grammar for inexperienced modelers”. In particular, the modeling grammar should not be a mere implementation of an existing framework such as the BWW model (Gehlert et al., 2005) but should be based on empirical evidence (Gehlert and Pfeiffer, 2007).

Before any issue can be raised, we need to decompose the goal into requirements. According to Wand and Weber (2003) a modeling grammar should enable users to interpret their models efficiently, i.e., to understand them easily. Consequently, our requirement discussed here is: “models should be easy to understand”. From a theoretical perspective, we seek theories that explain when and why models are easily comprehensible. To explain model comprehension we apply the work of Bodart et al. (2001). They distinguish between surface level comprehension and deep-level comprehension. Surface level comprehension is measured as proportion of the model to be recalled from memory after the model has been studied by a model user for a certain amount of time. Deep level comprehension is measured as the ability of a model user to answer questions regarding the model’s content (see the right area of Figure 3).

The authors studied the influence of optional properties on surface level comprehension and deep level comprehension where optional properties refer to attributes and relationships model elements may or may not possess. Bodart et al. (2001) found a significant positive effect between avoiding optional
properties and deep level comprehension. In addition, they found a significant negative effect between
avoiding optional properties and surface level comprehension. In sum, the authors showed that
optional properties hinder deep level comprehension (negative correlation) but support surface level
comprehension (positive correlation).

Figure 3: Discussion of Optional Properties

To describe the hypotheses used by Bodart et al. (2001) we need to know which indicators were used
to measure them. The construct “surface level comprehension” was measured by the indicators
“proportions of entities recalled correctly”, “proportion of relationships recalled correctly”,
“proportion of attributes recalled and named correctly”, “proportion of attributes recalled and typed
correctly” and “proportion of relationships recalled correctly with correct cardinalities” (Bodart et al.,
2001, pp. 391). The construct “deep level comprehension” was measured by the indicators “accuracy”,
“time” and “normalized accuracy” (Bodart et al., 2001, pp. 397). The “optional properties” construct
was measured binary – optional properties were either included in the model or were not included. It is
important to note that the comprehension constructs “surface level comprehension” and “deep level
comprehension” are different because they were measured with different sets of indicators (see Figure
3). As this theoretical setting explains the consequences of including optional properties, the central
design issue for the new modeling grammar might be: Should the modeling grammar provide optional
properties? Figure 3 summarizes the discussion about optional properties.

The issue whether to implement optional properties is related to the requirement “model
comprehension”. There are two possible positions to resolve this issue, either to allow or to prohibit
optional properties in the modeling grammar. Both positions are supported by the respective
arguments and each argument is based on a hypothesis. For instance, the position “prohibit optional
properties” is based on the hypothesis “using optional properties in models hinder deep level
comprehension” (negative correlation). Since the position “prohibit optional properties” does not
implement the hypothesis’ cause construct and the hypothesis’ correlation is negative, the hypothesis
supports this position (see D2). This hypothesis, however, can only be applied in this argumentation
process when its constructs are measured in the same way as the position and the requirement
(assumptions A1 and A2). Committing to this assumption leads to specializing the meaning of the
requirement “model comprehension” to “deep level comprehension”. In the same way, the assignment
of the hypothesis to the arguments implies that the position is interpreted as “diagrams without
optional properties”.

Proceedings ECIS 2009
The same argumentation can be formulated for the position “allow optional properties”. It is assigned to the respective argument, which in turn is based on the hypothesis “allowing optional properties supports surface level comprehension”. By applying assumption A1 and A2, the meaning of the requirement “model comprehension” is specialized as “surface level comprehension”.

To resolve the issue, the design researcher needs to decide whether to allow optional properties in the modeling grammar. Figure 3 shows that the design researcher has decided to prohibit optional properties in the modeling grammar. Consequently, the design researcher chooses to interpret model comprehension as deep level comprehension rather than surface level comprehension. Based on this decision, we can derive possible constraints from our theory-supported arguments. Because of the choice to prohibit optional properties we know from the underlying theory, that the models of the modeling grammar will hinder surface level comprehension. In our case, was have inferred a side-effect from the design rational. This side-effect is expressed as constraint “the modeling grammar hinders surface level comprehension”.

Hence, the design principle can be summarized as follows: If a design researcher’s goal of a modeling grammar is to achieve comprehension, one way to realize it, is by prohibiting optional properties in the modeling grammar. This decision particularly enables deep-level comprehension of the models. At the same time it hinders the model’s surface-level comprehension.

Applying the hypothesis relating to optional properties in the design principle results in the following four effects:

- When using the design principle, the original requirement “model comprehension” is refined to the concept of “deep level comprehension” which is represented by the indicators “accuracy”, “time” and “normalized accuracy”.

- Since the argument is based on a generally applicable hypothesis, the design researcher now knows that the design principle supports deep-level comprehension whenever it is used. This claim is based on the logic of deductive reasoning. Whenever the design rationale is applied to a more concrete design situation, the hypothesis is also deductively applied in a more specific context. Because of the truth preserving nature of deductive reasoning (Balzer et al., 1987; Popper, 2002), the hypothesis will still hold in this specific context.

- Design principles enable the design researcher to keep track of the development of IS theories more easily. Assume that due to advances in theoretical IS research, the relation between optional properties and deep-level comprehension is falsified. Since the design researcher knows the link between the design principle and theory, the design decision of avoiding optional properties in the modeling grammar may be renegotiated.

- The proposed approach has also implications for theory testing. If the design principle developed above is used in different modeling grammars and the effect of deep-level comprehension cannot be achieved by this grammar, this indicates that the underlying hypothesis “avoiding optional properties supports deep level comprehension” may not be correct. Confronting theory experts with this situation may stipulate further theory testing (Weber, 1987).

6 CONCLUSION AND OUTLOOK

In line with Walls et al. (1992), we have argued that theories should be used when designing IT-artifacts. In particular, we discussed the justificatory role of theories in design research by integrating theoretical arguments in design rationales. We have introduced the concept of design principles to denote a set of requirements, issues, positions, selected positions, arguments and hypotheses. On the example of developing a modeling grammar we demonstrated the utility of design principles to capture theory-based design knowledge, which can be reused in similar design research projects.
Besides justifying design decisions, design principles have two additional benefits with regard to cumulative research activities in IS design research:

1) When progress in developing and testing IS theories falsifies a theory, the design researcher knows which design principle needs to be modified. If the design principle was not based on theory, this theoretical knowledge was incorporated into the experience of the design researcher and could not be easily identified in the design principle.

2) If a theory-based design principle fails to be useful when implementing an IT artifact, it indicates that the hypothesis used to back up the design principle’s arguments may not be correct. This situation gives raise to extensive theory testing and may lead to a falsification of this theory.

Consequently, using design principles helps to bridge the gap between design research and theoretical research and intertwine both disciplines. However, we do not argue that every design principle should be backed up by theory. Since our approach is an extension to IBIS, it is still possible to derive design principles based on experience only. The main reason for deriving design principles based on experience is that IS theories are not yet available for every situation a design researcher may face. Furthermore, theories that are available may be not mature enough or too coarse or too fine grained to be used in design principles. For instance, the well accepted and tested Technology Acceptance Model (Davis, 1989) was too coarse grained to be usable for our example of designing a modeling grammar. Instead we needed to justify our design decision with the relatively new theory by Bodart et al. (2001), which was tested only once. The flexibility to develop design principles based on both theory and experience opens up two interesting avenues of future research to extend our concept of design principles:

1) Existing design principles may be complemented with theoretical knowledge, which explains the effects and side-effects of the design principle. Because a theory’s hypothesis is always bound to its construct’s indicators, applying the hypothesis means to explain how the positions and the requirements should be measured. Consequently, adding theoretical support to existing design principles leads to more precise understanding of the requirements and the solution of the design principle.

2) If a designed artifact proves to fulfill its requirements, its design principles based on experience may be used to initiate theory building. Hypotheses can be derived from existing design principles by interpreting the decision’s position as a cause construct and the requirement as an effect construct. Consequently, each design principle exhibits at least one hypothesis. These hypotheses can then be used to construct a theory for predicting (theory type IV in Gregor, 2006).

From a methodological perspective, our immediate subsequent research will focus on exploring the benefits of using theories of the types I, III, and V in design research and incorporate other argumentation theories, e.g. from Toulmin (1958).

7 REFERENCES


APPLYING THEORY-BUILDING TECHNIQUES TO THE DESIGN OF MODELLING LANGUAGES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0364.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Conceptual modelling, Grounded Theory, Design Science, Research methodology</td>
</tr>
</tbody>
</table>
APPLYING THEORY-BUILDING TECHNIQUES TO THE
DESIGN OF MODELING LANGUAGES

Becker, Jörg, University of Münster, Leonardo Campus 3, 4149 Münster, Germany,
joerg.becker@ercis.uni-muenster.de
Karow, Milan, University of Münster, Leonardo Campus 3, 4149 Münster, Germany,
milan.karow@ercis.uni-muenster.de
Müller-Wienbergen, Felix, University of Münster, Leonardo Campus 3, 4149 Münster, Germany,
felix.mueller-wienbergen@ercis.uni-muenster.de
Pfeiffer, Daniel, University of Münster, Leonardo Campus 3, 4149 Münster, Germany,
daniel.pfeiffer@ercis.uni-muenster.de
Seidel, Stefan, University of Liechtenstein, Fürst-Franz-Josef-Strasse, 9490 Vaduz, Principality of
Liechtenstein, stefan.seidel@hochschule.li

Abstract

In their 2004 paper Hevner et. al proposed a set of guidelines for conducting design science research
projects in the IS discipline. While useful, these guidelines have a relatively high level of abstraction.
However, various IT artifacts such as models, methods, techniques and implementations require IS
researchers to apply differing methods in order to construct and evaluate purposeful artifacts
respectively. In this paper we discuss a particular class of IT artifacts: conceptual modeling
languages. As constituent parts of software development methods, a multitude of such languages has
been proposed and discussed. Yet, in the related literature on method design only little guidance is
provided on how to derive appropriate conceptual modeling languages from empirical data. We
believe that “good methods” need to be rigorously grounded in empirical findings. Taking a look at
the related literature on inductive theory building reveals that at there are prominent similarities
between the elements that constitute theories and those that constitute conceptual modeling
languages: whereas theories comprise of constructs and relationships between these, conceptual
modeling languages comprise of language constructs and relationships among these. We draw from
the body of literature on grounded theory building and propose a new approach to designing
conceptual modeling languages.

Keywords: Method Engineering, Conceptual Modeling Language, Theory Building, Grounded Theory
Method
1 INTRODUCTION

Design Science Research has emerged as a popular research area in the IS discipline. Recently, there has been an increasing number of design science studies (cf. March & Storey, 2008; Winter, 2008) and also several discursive papers report on the usage of design science research in the IS discipline (e.g. Niehaves, 2007). Hevner et al. (2004) propose a set of guidelines for conducting design science research projects in the IS discipline. These guidelines are widely accepted as being feasible and providing guidance to an area of IS research that was often accused to not be rigorous and lack evaluation. While useful, the guidelines have a relatively high level of abstraction. However, particular IT artifacts such as constructs, models, methods, and instantiations require IS researchers to apply differing methods in order to construct and evaluate purposeful IT artifacts.

One such type of artifacts subsumes conceptual modeling languages, as being part of the “methods applied on the development and use of information systems” (Hevner et al. 2004, p. 82). The construction of conceptual modeling languages mainly originates from the field of information systems development (ISD). In the last two decades, a great number of modeling languages has been developed which left software engineers facing the major problem of method evaluation and selection (Harmsen, 1997). This problem has been addressed in a two-fold manner. First, unification efforts have been made in order to merge the mutual aspects of prevalent methods (Jacobson et al., 1999). Second, much effort has been devoted to the adaptation of development methodology by assembling specific aspects of different methods in order to meet project-specific requirements. This area of research became well-known in the ISD discipline under the notion of (Situational) Method Engineering (Brinkkemper, 1996; Harmsen, 1997; Ralyté & Rolland, 2001).

However, these approaches tend to not consider the actual domains, or contexts, methods are eventually applied to. The modeling languages used in this context mostly describe formal (software) systems, i.e. their constructs are anchored with formal semantics (such as programming language, c.f. Harmsen, 1997), thus having no denotation towards concepts of a material (“real-world”) domain. Nevertheless these languages are increasingly used for describing material contexts, e.g. in business process modeling (Rosemann et al., 2008) or requirements engineering (Mylopoulos et al., 1999).

In this paper we advance the construction of conceptual modeling languages by introducing a new approach to ground the development of such methods in empirical data. The motivation for this approach rests in the awareness that existent literature to a great extent does not address inductive development of methods based on empirical data. Yet, the related literature on theory building reveals that there are prominent similarities between the elements that constitute theories and those that constitute conceptual modeling languages: whereas theories comprise of constructs and relationships between these, conceptual modeling languages comprise of language constructs and relationships among these. Thus, we draw from the body of literature on theory building and propose a new approach to designing conceptual modeling languages.

The paper is structured as follows. In the next section we introduce related work on method design, design science research and theory building. We then introduce a new approach to designing conceptual modeling languages by transferring existent knowledge on theory building to the domain of designing methods. We then discuss the proposed approach. The paper concludes with a brief discussion of contributions and limitations and provides an outlook onto our future research agenda.

2 BACKGROUND

2.1 Design Science Research

Design science research aims at solving practical and theoretical problems "by creating new and innovative artifacts.” (Hevner et al., 2004 p. 75) The basic principles of design science research (DSR) can be traced back to engineering and Simon’s (1996) sciences of the artificial. In contrast to
behavioral science, DSR does not seek to understand the world as it is and how it works. Rather, it strives to develop solutions to improve the current state of affairs. DSR intends to provide IT artifacts that are novel and useful. These IT artifacts must exceed the current state of the art and have to serve a human purpose.

There have been several attempts in the IS community to define the IT artifact (Orlikowski & Iacono, 2001; Weber, 2003; Benbasat & Zmud, 2003; Venable, 2006). March and Smith (1995) differentiate between four types of IT artifacts: constructs that provide language concepts in which problems are described, methods that explicate the process of how to solve a problem, models that utilize the constructs to represent an application domain and express the problem and solution space, and instantiations that constitute the technical realization of constructs, models, and methods. In the understanding of March and Smith (1995) design science research must ultimately lead to one of these artifacts. A design science research method seeks to systematically guide the development of an artifact.

Until now, no widely accepted research method for design science has been established in the IS community. Even more so, there exist concerns that a general design method cannot be defined (Hooker, 2004). It is argued that design is a creative process that cannot be fully formalized. Nonetheless, various procedures have been suggested to methodically support the design activities (Peffers et al., 2007; Takeda et al., 1990; Nunamaker et al., 1990; Walls et al., 1992; Cole et al., 2005). More specifically, various approaches for the design of conceptual modeling methods and languages have been proposed, the most prominent of which we discuss in the subsequent section.

2.2 Design of Conceptual Modeling Languages

As has been indicated, conceptual modeling languages are applied in order to represent the relevant knowledge of a domain (Wand et al., 1995). In this paper we focus on modeling languages that are designed in order to represent facts about material domains, meaning aspects of the physical and social world or – more precisely – perceptions thereof, for means of communication and understanding (Mylopoulos, 1992). Note that such methods are not ought to be used in order to specify formal systems.

Conceptual modeling languages comprise of fundamental modeling constructs, that is, language primitives which are called the vocabulary. For example, in the Business Process Modeling Notation (BPMN), constructs represent activities, events or sequence flows. In addition to that, conceptual modeling languages provide a collection of rules that describe how the constructs can be combined to create statements about the domain of discourse. Such rules specify what constructs may be connected to each other. For example, it can be defined that activities can be linked by a sequence flow. Usually, rules are specified in a language’s meta-model and complementing contextual conditions (Earwig, 1999).

In the existent literature, two major approaches to the design of conceptual modeling languages can be found: language design in method engineering and ontology-based language development.

The first approach views conceptual modeling languages as artifacts that are created as part of a method engineering process. Thus, it is concerned with the selection, adaptation and design of (situation-specific) conceptual modeling languages as well as their corresponding modeling procedures. A system development method is assumed as to consist of a set of reusable fragments (Brinkkemper et al., 1998; Harmsen, 1997). Ralyté et al. (2004) describe different strategies for method engineering projects that differ regarding the degree of fragment reuse. However project specifics may, require the method engineer to derive novel language constructs from the problem domain at hand. While method engineering literature provides comprehensive directions on how to prepare method fragments for tool-supported integration and assembly, little guidance is given to the challenge of approaching a problem domain in order to derive feasible modeling language constructs. Furthermore, the concept of domain has primarily been perceived as the formal target systems in
earlier ISD modeling language development, such as programming languages or paradigms. Consequently, the resulting languages are mostly anchored to formal semantics.

The second approach aims to overcome the deficiencies of those languages to describe real-world-phenomena: to anchor modeling languages to material domains, it draws on the concept of ontologies as a theoretical foundation. Wand and Weber (1993) utilized a *top-level ontology* (Bunge, 1977) to evaluate existing modeling languages with regard to precision and completeness by matching the language elements with ontological concepts. Guizzardi (2005) developed an own ontological foundation for structural conceptual modeling languages and suggested an approach to derive language constructs based on this ontology. Although ontologies represent feasible anchoring systems for modeling language constructs (Harmsen, 1997), relying on this concept merely shifts the problem of how to identify useful constructs from language design to ontology design. Although one can find the notion of *ontology engineering* (Devedžić, 2002) and examples for the construction of particular domain ontologies (e.g. Fernández-López et al., 1999), the body of research work on this approach lacks generalized guidance on how to derive a conceptualization from empirical data.

Taking off from this discussion, in this paper we suggest an empirically-based approach to designing conceptual modeling languages. It is hoped that the inductive development of such languages based on empirical data can contribute to the languages’ usability and adequateness.

### 3 ON THE APPLICABILITY OF THEORY BUILDING PROCEDURES TO DESIGN SCIENCE RESEARCH

In the IS discipline, there have been attempts to classify theories and develop a more narrow understanding of what theory is. Gregor (2006) distinguishes between theories for analysis, explanation, prediction, explanation and prediction and design and action. Theory for analysis “does not extend beyond analysis and description. No causal relationships among phenomena are specified and no predictions are made” (p. 620). Similarly, conceptual modeling languages provide an analysis and description of a problem domain. Thus, we argue that conceptual modeling languages can be compared to analytic theory: they provide clear definitions of constructs that are relevant in a certain problem domain and describe relations among these.

The most general term that describes the building blocks of a theory is called a concept (Strauss & Corbin, 1998). *Concepts* represent phenomena and can be grouped into more abstract concepts that are then referred to as *categories* (Strauss & Corbin, 1998). Categories have *properties*, which describe certain characteristics that objects of the same category share. For example, if one category was “actor” objects belonging to this category could share the property of “position”. Thus, different actors could be placed on a dimensional range describing various positions, such as producer or director.

Generally, the scientific process comprises of the stages of observation, induction, and deduction (Eisenhardt, 1989; Handfield & Melnyk, 1998; Wallace, 1971). Thus, it can be argued that the scientific process starts with the inductive development of theory that is then deductively applied to incoming data and thus validated. This process of validation may lead to new or revised theory. There are various approaches of how to inductively develop theories, for example case study research (Eisenhardt, 1989), or Grounded Theory Method (GTM) (Glaser & Strauss, 1967; Strauss & Corbin, 1998). In this paper we particularly draw on the literature on GTM. As indicated, GTM aims at inductively develop theory based on empirical data. GTM is thought to ground the emergent theory in the data. It is not preconceived or forced upon the data but rather emerges from it (Glaser & Strauss, 1967).

We argue that GTM offers the researcher a set of procedures that can be beneficial in order to inductively develop conceptual modeling languages. Our argument rests in the following observations:

- The process of building grounded theories is highly iterative. Theory and data are constantly compared (Glaser & Strauss, 1967). This process can be referred to as *comparative analysis*. Similarly, DSR processes are highly iterative and constantly compare the evolving artifact with its purpose (Hevner et al., 2004).
Glaser and Strauss (1967) further introduce the term theoretical sampling as a process of “data collection for generating theory whereby the analyst jointly collects, codes, and analyzes his data and decides what data to collect next and where to find them, in order to develop his theory as it emerges” (p. 45). When designing a modeling method, the method designer will start to identify what concepts may be relevant in a certain context. In order to advance the construction, he or she will further investigate the domain at hand by decisively choosing locations and respondents he or she talks to.

Grounded theory studies typically start with a stage called open coding (Glaser & Strauss, 1967). In open coding the researcher identifies a set of themes or categories that appear to be relevant in order to describe and explain a phenomenon under investigation. Similarly, when designing a modeling method, the method designer has to identify those language constructs that are relevant and applicable to a particular domain.

Grounded theory provides procedures that support the researcher in identifying relationships between concepts. For example, Strauss and Corbin (Strauss & Corbin, 1998) suggest to classify emergent categories by whether they represent (a) phenomena, (b) conditions, (c) actions/interactions, or (d) consequences. Thus, conditional structure is identified. Likewise, the method designer seeks to identify relationships among language constructs.

Grounded theory relies on a technique called memoing (Miles & Huberman, 1994). Memos are used to document the researcher’s conceptual thoughts that eventually lead to the generation of theories. Memos are constantly written, re-written, and integrated (Strauss & Corbin, 1998). Thus, the process of memoing is conducive to the iterative nature of DSR projects, such as the development of methods. This concept is similar to what is referred to as a method rationale (Rossi et al., 2000).

In the following section we compare the basic elements of conceptual modeling languages to the basic elements of grounded theories. We then describe how the above outlined procedures of building theory can be applied to generate conceptual modeling languages.

## 4 APPLYING GTM PROCEDURES TO INDUCTIVELY DEVELOP CONCEPTUAL MODELING LANGUAGES

### 4.1 Concepts of Conceptual modeling languages and of GTM

Table 1 provides an overview of the comparison and matches the terminology of conceptual modeling language design and grounded theory method.

<table>
<thead>
<tr>
<th>Modeling Language Design</th>
<th>Description</th>
<th>Grounded Theory Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language Construct</td>
<td>Concepts of a domain can be translated to constructs of the domain-specific language</td>
<td>Code / Concept</td>
<td>building blocks of a theory, abstracts descriptions of real world phenomena</td>
</tr>
<tr>
<td>Candidate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language Construct</td>
<td>Categories indicate a core concept of the domain, thus a language for describing instances of that domain should provide a dedicated representation</td>
<td>Category</td>
<td>Derived from concepts, aggregated and structured, constituent part of a theory’s statement</td>
</tr>
</tbody>
</table>

*Table 1 Relationship between concepts from GTM and conceptual modeling language design*
Language construct candidates are early abstractions from phenomena that a method designer perceives as relevant. These elements are part of the individual conceptualization (Guizzardi, 2005) of the context at hand. In grounded theory development, codes represent first cognate incidents in the data, which a researcher assigns to more abstract terms or themes. In further iterations, these codes are assembled to concepts by further clarifying the context’s structure and terminology.

The language construct is the central element of modeling language design. In a semiotical sense, language constructs are types of particular signs (Genova et al., 2005), that is, they have a syntactic, a semantic and a pragmatic dimension (Morris, 1970). The syntactic dimension of modeling languages can be split into abstract and concrete syntax (Earwig, 1999). While the former describes what discriminate constructs are available and how they may be combined, the latter assigns a graphical representation to each construct, so as to create a language primitive (Guizzardi, 2005). Techniques utilized in Grounded Theory Method can contribute to outline the semantic component of language constructs. The concepts derived from codified data yield a promising starting point for the material backdrop of a domain-specific modeling language. The process of refinement and abstraction to develop early sketched concepts to structured and well-defined categories is analogous to the definition of a language’s vocabulary.

Properties are a special class of constructs that denote existential dependency on other language constructs. We distinguish between different types of properties. So-called intrinsic properties are property types of language constructs that obtain a definite value when instantiated as model elements. For instance, if we defined a language construct “task” an intrinsic property could be “duration”. Mutual properties describe property types that are shared by instances of language constructs, such as being in a relationship or being part of a composite concept (Shanks et al., 2008). In the process of theory building, properties emerge from concepts that give further explanation to particular categories and are therefore existentially dependent on them.

The language rules constrain the possible combinations of language constructs and are part of the abstract syntax of a modeling language (Guizzardi, 2005). As language constructs denote concepts grounded in the domain, these combinations denote meaningful statements that must also be grounded in the empirical world. One major component of theory building is the exposition of such basic statements by revealing the relevant relationships among the concepts.

Based on the identified analogies, we propose a process that guides the development of a special type of IT artifacts, namely domain-specific conceptual modeling languages. Modeling languages provide clear definitions of constructs and (potential) relationships between constructs. The process draws on techniques that stem from the literature on theory building and results in what can be referred to as analytic theory. The approach to building theory we consider generates substantive theory, that is, theory that is applicable to a certain domain. Similarly, any language developed according to the scheme we are presenting will depend on the context it was developed in.

4.2 Applying GTM procedures in order to inductively design conceptual modeling languages

In the following we describe how conceptual modeling language constructs and their relationships can be inductively developed based on empirical data. To illustrate the process, we have chosen examples
from a grounded theory study that was conducted based on data from the film industry in order to study business processes in creative environments (Seidel et al., 2008). Thus, the language to be designed would be a business process modeling language tailored to that specific material domain. The process comprises of the following steps (cf. Fig. 1): Data Collection, Identification of concepts, further developing concepts, relating concepts, and concluding the design process. As has been indicated, alike the generation of theory the design of conceptual modeling methods is a highly iterative and interwoven process, which becomes particularly evident through the use of constant comparative analysis and theoretical sampling.

**Figure 1: GTM-based Language Design Process**

**Stage 1: Data collection**

At the outset, the modeling language designer must decide upon the data the language development is based on. Examples are the analysis of existent documentation, interviews, or observational data. Generally, a multitude of data sources can be considered, a process that is often referred to as triangulation across methods (Orlikowski, 1993). As triangulation across methods is typical for GTM studies (Glaser & Strauss, 1967), we suggest method designers to consider different data sources so as to allow for multiple vantage points for identifying what is relevant in a particular domain of interest. The result of this stage is a clear outline on what data sources will be used according to the intended scope of the modeling language to be designed.

**Stage 2: Identifying concepts**

The method designer starts with the identification of concepts being relevant in a certain context and for a certain modeling purpose. This identification of concepts draws on the process of open coding (Glaser & Strauss, 1967). Thus, the researcher opens up the text in order to identify what may be relevant in the data. By comparing incident to incident, the method designer comes up with first concepts which share certain characteristics and comes up with various concepts. Even though much of what will be needed may be found in the interview or observational data, the method designer may want to work with other techniques than simple comparisons. One such strategy that has also been proposed by Strauss & Corbin (1998) is that of making theoretical comparisons. Thus, the method designer enhances her “theoretical sensitivity” (Glaser, 1978), e. g. by evaluating existent modeling languages for reusable conceptualizations. The result of this stage is a quantity of domain-relevant concepts that are candidate language constructs.

**Stage 3: Further developing concepts**

It is not uncommon that the researcher ends up with generating a large number of concepts (Strauss & Corbin, 1998). To further integrate concepts, they are grouped under more abstract concepts called categories. By using categories the method designer reduces the number of items she works with. This process depends on the modeling purpose as well as of the individual perspective of the method designer. The result of this stage is a reduced list of categories, which comprises the elements of the first draft on the modeling language model (the meta model).

**Stage 4: Relating Concepts**
When starting to analyze the data, the researcher will recognize first relationships between concepts. Eventually, these relationships result in the formation of hypotheses or propositions. Similarly, when designing conceptual modeling languages, the method designer identifies potential relationships between language constructs. As has been indicated, depending on the type of language that is constructed it may be possible to distinguish different types of categories. By grouping categories accordingly, relationships emerge. Thus, the result of this stage is a first model of the domain-specific modeling language (the meta model) that comprises the quantity of language constructs and the typed relationships among them. These relationships represent allowed connections between instances of the involved constructs.

Stage 5: Concluding the design process, development of the concrete syntax

The iteration between analyzing data and generating language constructs and relationships can be concluded when additional data analysis does not provide any further insight. In GTM, this stage in the process is referred to as “theoretical saturation” (Glaser & Strauss, 1967). This highly iterative process rests in the application of procedures that were discussed earlier, namely the making of comparisons and theoretical sampling. During the conclusion of the theoretical development, the meta model will reach a state where no more substantially changes to the language core will be made. At this point, the language designer will assign a visual representation to each component the abstract language model (Guizzardi et al., 2002). Conclusively, the result of this stage is a domain-specific language prototype. This prototype can be used to further evaluate the underlying conceptualization, as well as to analyze the lucidity of the chosen representation (concrete syntax).

It is vital for the claim of traceability of language design and thus for the feasibility of language evaluation to rigorously document all decisions and their basis in the empirical data. As indicated, we propose to make extensive use of memoing (Miles & Huberman, 1994). There are different types of memos that can be used in order to provide a comprehensive method rationale: Codes notes, for example, accompany the process of conceptualizing based on constant comparison and theoretical sampling, whereas operational notes help to guide the researcher in deciding on what data to collect next, etc. (Strauss & Corbin, 1998).

4.3 Example Case

In the following, we illustrate the application of the proposed approach by developing exemplary language constructs based on data collected in an exploratory study on organizational creative processes. (c.f. Seidel, 2009).

Data collection: Data has been collected in three organizations with over 30 interviewees using semi-structured interview and process modeling techniques.

Concept identification: While studying the concept of CIP, the people conducting the tasks within these processes emerged as important context concepts. Example codes identified in the data were “visual effects artist”, “editor” or “sound editor”.

Further development of concepts: The roles identified within the CIP context were further investigated and mutual properties could be identified in the data. All these individuals share a certain process expertise that is necessary to carry out creative tasks, e.g. the ability to break down a creative problem in order to find a solution strategy to it. Furthermore they share the property of creative skills, i.e. the ability to generate novel artifacts and to judge solution on aesthetic aspects. As another important property with influence to CIP, the working location has been identified. The concepts have accordingly been generalized into the category artist.

The modeling language aims to provide for means to describe the processes within the domain. The category artist is codified as an element type within the language.

Relating concepts: Artists represent a specialized type of task owner in CIPs. Thus, they can be associated to creativity-intensive (sub-)processes. The property location has significant influence to
the collaboration with supervisors and clients, thus it will be modeled as an attribute of artist (c.f. Figure 2).

<table>
<thead>
<tr>
<th>Artist</th>
<th>CIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>jobTitle: String</td>
<td>label: String</td>
</tr>
<tr>
<td>description: Text</td>
<td>creativeSupervisor</td>
</tr>
<tr>
<td>location: Location</td>
<td>artist[]: artist</td>
</tr>
<tr>
<td>assignedTo[]: CIP</td>
<td>parent: CIP</td>
</tr>
</tbody>
</table>

Figure 2: Meta model detail

Concluding design process: In this last step, a representation of the element type has to be developed. The CIP might be described in a form-based model where assigned artists can be added in a list and implemented into a modeling tool. To conclude the DSR process, the resulting language must be evaluated with appropriate measures (e.g. Recker, 2008).

5 DISCUSSION

In order to evaluate the approach we have suggested in this paper, we consider the guidelines proposed by Hevner (2004). Table 2 provides an overview.

Table 2: Evaluation with Guidelines as proposed by Hevner (2004)

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Our Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guideline 1: Design as an artifact</td>
<td>The process we proposed in this paper aims at developing conceptual modeling languages. Thus, the process results in what is considered to be purposeful IT artifacts.</td>
</tr>
<tr>
<td>Guideline 2: Problem relevance</td>
<td>It lies in the responsibility of the method designer that the problem that is to be targeted by the conceptual modeling language is relevant.</td>
</tr>
<tr>
<td>Guideline 3: Design evaluation</td>
<td>Evaluation is an integral part of the suggested process. By applying the principles of constant comparative analysis and theoretical sampling, the researcher constantly compares concepts and relationships to incoming data.</td>
</tr>
<tr>
<td>Guideline 4: Research contributions</td>
<td>It is hoped that methods designed based on the approach described in this paper are both “clear and verifiable” as Hevner (2004) states. It is suggested to accomplish this by constantly iterating between inductively generating categories and relationships that are then deductively applied.</td>
</tr>
<tr>
<td>Guideline 5: Research rigor</td>
<td>The process we introduced aims at providing a set of procedures that can be applied in order to design conceptual modeling languages. It is hoped that by following and documenting these procedures, the method designer makes the process of method development transparent and traceable. It cannot be claimed that a particular method is complete or correct – however, it can be argued that the process of method development is plausible.</td>
</tr>
<tr>
<td>Guideline 6: Design as a search process</td>
<td>The development of conceptual modeling languages based on constant comparative analysis and theoretical sampling is a highly iterative search process that in every stage is highly dependent on the substantive area in which it is grounded.</td>
</tr>
<tr>
<td>Guideline 7: Communication of research</td>
<td>Applying rigorous procedures of documenting the research process can contribute meaningfully to successful and appropriate communication of results.</td>
</tr>
</tbody>
</table>
Note that the perspective we presented in this paper largely differs from that of those approaches to method engineering that stem from the discipline of software engineering. Whereas these streams of literature seek to technically sound developing conceptual modeling languages and focus on syntactic integrity, our approach focuses on the identification of relevant language constructs and relationships.

It is our belief that the appropriateness of modeling languages is determined by the context in which they are used. Of course, the development of software systems requires methods that enable to construct syntactically correct models. Thus, in many cases it will be necessary to combine formal procedures with empirically-based, inductive methods to identifying concepts and relationships that are relevant and applicable to a particular domain.

6 CONCLUSION

This research contributes to the IS body of knowledge by proposing a rigorous, empirically-based approach to inductively develop conceptual modeling languages methods based on well-established techniques known from the domain of theory building. It is our belief that there is need for detailed approaches supporting both researchers and practitioners in developing purposeful IT artifacts. Our argument rests in the observation that (a) there are similarities between basic elements of conceptual modeling languages and the elements that constitute theory, and (b) that the GTM offers researchers a set of procedures that can also be applicable to the development of such languages.

6.1 Limitations

Inductively developing theory or modeling methods limits the scope of the artifact to a so-called substantive area (Urquhart, 2001; Strauss & Corbin, 1998). Thus, they may be very practical but are also limited to a particular scope. Thus, the approach we have advanced in this paper is limited to developing modeling languages that are applicable to particular domains. The study so far limits the discussion on the development of the language elements as constituent parts of a modeling method. However, a functional method also has to provide for a modeling process that guides modelers on how to use the language constructed. Although we believe that an empirically grounded language will assist the efficient elicitation of information in its aspired domain, one can argue that the process of modeling is also dependent on the context of application.

Furthermore, it must be noted that the result of any method design process is highly dependent on both the method designer as well as the intended purpose of the language. A possible strategy to achieve a more independent view on the domain is triangulation (Eisenhardt, 1989). For instance a researcher might employ a colleague to develop an own conceptualization based on the same data.

Conclusively, any new guideline, theory, method, or approach must be tested in practice. Thus, we motivate researchers and method designers to applying those principles we presented in this paper.

Acknowledgements

This paper was written in the context of the research project ManKIP (Manangement of Creativity-Intensive Processes). The project is funded by the German Federal Ministry of Education and Research (BMBF) and by the European Social Fund of the European Union, promotional reference 01FM07061. We gratefully acknowledge the support of the Project Management Agency as part of the German Aerospace Center (PT-DLR).

References


Miles, M. B. and Huberman, M. A. (1994) Qualitative data analysis: An expanded sourcebook. Sage, California.


GREEN IT: EVERYTHING STARTS FROM THE SOFTWARE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0014.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Power, Software Metric, IT performance management, IT cost evaluation</td>
</tr>
</tbody>
</table>
GREEN IT: EVERYTHING STARTS FROM THE SOFTWARE

Eugenio Capra, Politecnico di Milano, Dipartimento di Elettronica e Informazione, via Ponzio 34/5, 20133, Milano, Italy, eugenio.capra@polimi.it

Francesco Merlo, Politecnico di Milano, Dipartimento di Elettronica e Informazione, via Ponzio 34/5, 20133, Milano, Italy, merlo@elet.polimi.it

Abstract

In this position paper we discuss the importance of Green IT as a new research field that investigates all the environmental and energy issues related to IT and information systems in general. In particular we focus on the energy consumption of software applications, which is amplified by all the above IT layers in a data center and thus is worth a greater attention. By adopting a top-down approach, we address the problem from a logical perspective and try to identify the original cause that leads to energy consumption, i.e. the elaboration of information. We propose a research roadmap to identify a set of software complexity and quality metrics that can be used to estimate energy consumption and to compare specific software applications.

Keywords: Green IT; energy efficiency.
1 INTRODUCTION

“Green IT” is an expression that indicates a new research field that investigates all the environmental and energy issues related to IT (Murugesan, 2008). More specifically, we think that Green IT may refer to three different research areas:

1. Energy efficiency of IT;
2. Eco-compatible management of the lifecycle of IT;
3. IT as an enabler of green governance.

The first research area aims at designing energy-efficient IT architectures and data centers, covering also all the effects that utilization practices have on energy consumption. As we will explain in this paper, energy consumption impact on operating cost has grown in the last years and is now very significant. The second research area proposes to study new methodologies and technologies for eco-compatible manufacturing of IT components, to optimize packaging, and to minimize the environmental impacts of the whole lifecycle of IT. This includes eco-labeling and eco-compatible management and storage of waste and dismissed IT components. Finally, the third research area aims at leveraging IT as a means for measuring and monitoring the green parameters (e.g., energy consumption, temperature, toxic waste produced) related to all business processes, not limited to the IT area. This includes the design of monitoring devices as well as decisional support systems and dashboards to store, analyze, and compare green KPIs.

In this position paper we focus on the first research area and propose a research plan to analyze software energy efficiency. We illustrate as quantum physics theory offers an overall interpretation for the energy consumption of IT: elaborating information requires a minimum energy related to the quantum nature of the world. Actual consumptions are by far higher than the minimal theoretical level because current IT systems introduce a number of inefficiencies related to many different layers, which go from logical gates level to servers architecture and data centers infrastructure level. From another standpoint, this means that there are many possibilities for increasing the energy efficiency of IT, and this is a great challenge that the scientific community needs to face.

Software plays a crucial role in this scenario. Although software does not directly consume energy, it deeply affects the consumption of hardware equipment. Software applications, ranging from operating systems and drivers for hardware devices to decision support systems and ERP suites, indicate how information should be elaborated and to some extent guide the functioning of hardware. Consequently, they are indirectly responsible of energy consumption. In this paper we propose a research plan to measure and compare energy consumption of different applications and to correlate these data with traditional software design quality metrics.

The paper is organized as follows. Section 2 describes why Green IT is gaining more and more importance due to the devastating effects of energy consumption on the environment, on operative costs, and on scalability. Section 3 describes how energy consumption is distributed in a data center and illustrates how the energy actually used for chip-level computation is amplified by all the above IT layers. Section 4 presents the theoretical hypotheses at the base of our research, and Section 5 proposes our research plan. Finally, Section 6 concludes the work.

2 WHY GREEN IT IS IMPORTANT

Green IT is attracting more and more attention both in the scientific and business communities. In the past decades research and innovation have focused on increasing clock frequency and on
miniaturization (Schaller, 1997), with only a marginal focus on power consumption, mainly associated with battery autonomy of laptop devices. This has resulted in extremely fast IT systems, but which consume a lot of energy that is very often inefficiently employed.

Energy consumption has devastating effects on:

1. Equivalent CO\(_2\) emissions;
2. Operating costs;

As consumptions rise, the attention on Green IT gains momentum.

According to recent researches (Murugesan, 2008, Brown and Lee, 2007, Kumar, 2007), IT is responsible for more than 2% of global CO\(_2\) emissions, and its environmental footprint is comparable to that of the aeronautic industry. The average amount of energy consumed by a PC in 1 year corresponds to the emission of 1 ton of CO\(_2\), and a server has roughly the same annual carbon footprint as an SUV doing 5 miles-per-gallon (Restorik, 2007). In addition to that, 70% of the landfills of lead, cadmium and mercury derives from the IT industry (Brown and Lee, 2007).

From an economical perspective, whereas the cost of hardware has only slightly grown in the last 12 years, the cost of power and cooling has grown four times. Figure 1 shows data on the global spending for servers in the last years and estimates for next years.

![Figure 1 - Global spending for server (Billion dollar, Source: Josselyin et al., 2006).](image)

Nowadays, power and cooling operating costs represent 60% of the total spending for new infrastructures, and consequently have a great impact on TCO (see Figure 2). This proportion is expected to rise even more, also because of the continuous growth of energy unit cost. As in most
companies energy costs are not charged to the IT budget, the importance of this phenomenon is not yet fully perceived, but it is likely that accounting rules will change as the impact of energy costs on overall IT costs rises more and more.

In addition to that, energy consumption is a limit to the scalability of data centers. New IT equipment requires an extremely high quantity of energy per square meter (e.g., a rack with 5 blade servers of 8 units consume more than 20KW, as much as an apartment complex) and also the energy required by personal computers rises at a rate of 8-10% per year. When data centers are located in areas with high population density, as it often happens in Europe, it may be difficult for power distributors to bring the required energy in the same building. As power infrastructure modifications are difficult and expensive, data centers that are not energy efficient cannot expand their capabilities. According to Forrester Research (Brown and Lee, 2007), in the next few years 60% of data centers will be limited by power consumption, cooling, and space issues.

![Figure 2 – Spending for energy and cooling/ spending for new servers (Percent, Source: Josselyin et al., 2006).](image)

Recent surveys show that there are growing concerns about Green IT in corporate contexts: according to Forrester Research (Forrester Research, 2007), 33% of North American and 48% of European IT procurement and operations professionals think that environmental and energy-related issues are very important in planning their company’s IT operations, whereas only 15% in North America and 6% in Europe think that Green IT is not a problem at all.

### 3 A MULTILAYER APPROACH

An average data center usually consumes at least 300 KW, whereas a large data center may consume more than 10MW. However, it is important to note that this significant quantity of energy is consumed at different layers, i.e. from different parts of the data center with different logical functions.
According to (Renzi, 2007), 40% of the energy consumed by a data center is absorbed by HVAC (cooling) and UPS (back-up batteries) systems, and another 42% is absorbed by fans, AD/DC transformers, and storage, whereas only 18% is consumed by the processors. In addition to that, as some processors stay idle for some time, the energy really used for computation may be as low as 3% of the total. Hence, researches and actions aimed at optimizing the energy consumption of data centers should address all the IT layers involved.

The reduction of power consumption should obviously focus on the optimization of the layers that consume the biggest part of energy, i.e. power and cooling and peripherals systems. This requires research also on non-IT items, such as UPS, air conditioning and other equipment. In addition to that, virtualization can greatly reduce the idle time of processors thus optimizing the energy consumptions. All these researches are specific to particular contexts and typologies of infrastructure.

In this paper we take a different and innovative perspective and we focus on the cause of energy consumption by information systems, independently from their infrastructural implementation. In the next section we illustrate how elaborating information per se requires energy, according to recent quantum physics researches. Quantum physics also quantifies the minimum theoretical amount of energy needed to commute a bit of information, which according to the current state of our knowledge could be optimally represented by the spin of an electron. Of course this minimal amount of energy is much lower than current consumptions. This gap between theoretical and actual consumptions is due to all the inefficiencies introduced by the different architectural layers of a computational system, e.g. because we use transistors rather than atoms to store and elaborate bits.

We posit that all these layers amplify the unitary amount of energy required to elaborate information, as all the hardware supporting a processor and the infrastructure in a data center are sized according to the amount of elementary computations required. This hypothesis will be verified during our research.

When a processor spends 1W to elaborate information the total energy consumed by the system may be as much as 28 times higher, due to drivers, memory, cooling, back-up batteries and all the other auxiliary components needed by the processor to work (see Figure 3). Thus, the benefits obtained by optimizing the energy consumed for computation are amplified by the above IT layers and have a great impact on the total consumption.
Accordingly, our research will focus on analyzing the energy efficiency of software algorithms, i.e. how efficiently information is elaborated, thus laying down the foundations for future optimizations.

Recent researches (Bruschi, 2007) have shown that current energy efficiency of algorithms and applications is on average 20%, whereas energy efficiency related to data quality (low quality data requires more operations) is no higher than 60% (Restorick, 2007).

4 THEORETICAL FOUNDATIONS

This section introduces the theoretical foundations of our research roadmap. We will assume both a physical and logical perspective. The former is essential to understand how energy is actually consumed by the technological infrastructure that is at the base of an information system; the latter will help us understand why energy is required to manage information.

From a physical perspective it is well known that the average power consumed by a microprocessor while running an application is $P = I \cdot V_{CC}$, where $I$ is the average current and $V_{CC}$ is the supply voltage. Since power is the rate at which energy is consumed, the energy consumption of a given application is the integral of the power consumption $P$ over time $t$, that is:

$$EC_{\text{physical}} = \int I \cdot V_{CC} \cdot dt$$

(1)
Measuring energy consumption of an application by means of Expression (1) requires to measure $I$ and $V_{CC}$ on the hardware system actually used. Consequently, this kind of measures always refer to a specific microprocessor architecture. Moreover, from Expression (1) it is impossible to analyze why energy is consumed.

In order to solve this problem, we need to link the physical domain (i.e., electric energy consumption) to the logical domain. Energy consumption can be assessed by analyzing why a given application requires a certain amount of energy to produce the desired output. The following paragraphs introduce some theoretical definitions that are required to consider the problem of energy consumption from a logical perspective.

The Margolus-Levitin theorem (Margolus and Levitin, 1998) posits that the maximum frequency for the status commutation of a physical system is directly proportional to the total energy of the system itself. As a consequence, the minimum commutation energy required by a system to operate at a given frequency can be computed as:

$$E_{\text{min}}(f) = f \cdot \frac{h}{4}$$

(2)

where $f$ is the frequency, and $h$ is the Planck’s constant. For example, if we represent a bit by means of the direction of an electron’s spin, the commutation energy required at the frequency of 1 GHz (thus comparable to that of current desktop computers, considered that, for an average particle like an electron, the maximum commutation frequency is $f_e \approx 3 \cdot 10^{13}$ Hz) is $E_e \approx 5 \cdot 10^{-21}$ J.

The thermodynamic depth (Lloyd, 2006) is a property of each physical system: it is essentially a measure of the information required to describe, and consequently build, the system itself. This metric is related to the concept of entropy. It is well known that entropy is the measure of the level of disorder in a given system (Haddad et al., 2005). Assuming that a system can always be described by a string of bits (e.g., by describing initial speed and position of all its atoms), entropy is the number of bits of the system that are disordered and unavailable to produce work. Conversely, negentropy is the measure that quantifies the number of bits that are ordered and structured. For example, a human being has an high degree of negentropy, whereas a balloon full of helium is completely lacking of negentropy. If we want to describe a table we need a certain number of negentropic bits, but we do not need to describe the positions of all the billions of atoms of the table: these bits can stay entropic without affecting our description.

Based on these definitions, the thermodynamic depth is defined as the number of negentropic bits that have been used to build the system.

The logical depth (Lloyd, 2006) of a generic string of bits, that can be interpreted as the representation of a generic system (as well as the output of a computer application), is defined as the computational complexity of the most efficient program that is able to produce that output. In other words, it is the smallest number of elementary logical operations required to perform the computation that produces the desired string of bits.

A software application executes a certain number of computations on a defined number of bits in order to obtain a result. Applying the theoretical definitions discussed above, the energy consumption of a software application can be estimated from a logical perspective as:

$$EC_{\text{logical}}(f) = E(f) \cdot C_c \cdot T_d$$

(3)

where $E(f)$ is the energy required by a single bit status commutation at frequency $f$, $C_c$ is the computational complexity of the application that is executed and $T_d$ is the thermodynamic depth of the computation that is performed onto the problem representation. In other words, Expression (3) estimates the energy consumption by considering how much energy is required by a single bit status commutation ($E(f)$) that is applied on a given number of bits ($T_d$) for a given number of operations ($C_c$). Expression (3) allows to analyze the causes that lead an application to consume energy because it is elaborating information, without focusing on the physical and electrical mechanisms of consumptions.
First of all, we note that there is an unavoidable trade-off between energy and frequency: a faster system requires more energy. The minimization of energy consumption can be achieved by optimizing each of the three terms of Expression (3).

As discussed before, the minimum energy required for the commutation of a bit status at a given frequency is given by the Margolus-Levitin theorem. This is only a theoretical lower bound, which is valid if bits are represented by electrons’ spins. As a matter of fact, several attempts of building a computing machine that uses the electron’s spin to represent a bit have been made (e.g., Isaac Chuang at MIT has factorized the number 15 with a 7 qubit computer (Vandersypen et al., 20010). It should be considered that the energy $E_{\text{min}}$ is by far lower than the actual energy that is consumed to switch a bit in current computers based on transistors (modern architectures require approximately $10^{-14}$ Joules to commute a bit, and research are being carried on to reduce this energy to $10^{-16}$ Joules1). However, this is an hardware-related research area, and goes beyond the purposes of this paper.

The remaining two terms, namely computational complexity $C_c$ and thermodynamic depth $T_d$ can instead be faced from an information system perspective.

The minimization of the computational complexity required to produce a desired output can be obtained if the generic application $A$ that is executed has the minimum possible computational complexity, that is, exactly the logical depth $L_d$ of the required output.

The thermodynamic depth can be minimized by adopting the most efficient way of representing the problem and the data required to produce the desired output, that is the minimum thermodynamic depth $T_{d\text{-min}}$.

As a consequence, the lower bound of Expression (3) at a given frequency is given by:

$$EC_{\text{min}}(f) = E_{\text{min}}(f) \cdot L_d \cdot T_{d\text{-min}}$$

Just as the minimum energy of commutation given by the Margolus-Levitin theorem, also Expression (4) is only an ideal theoretical lower bound. In particular, the problems of writing an application with the minimum computational complexity required to obtain a desired result or stating which is the most efficient representation of a given problem are not trivial problems. For example, there exist problems for which we do not know whether the algorithms used to compute their solution are the most efficient ones (e.g., sorting algorithms). Furthermore, there exist whole classes of problems for which we do not even know if an efficient solution exist (e.g., the NP-complete problem class, if $P≠NP$). On the contrary, it is possible to design a methodology that allows the comparison of different applications from the efficiency of energy consumption point of view. By considering Expressions (1) and (3), we posit that:

$$EC_{\text{physical}}(f) \sim EC_{\text{logical}}(f)$$

That is, the energy consumption described from the physical perspective can be considered as a measurement proxy of energy consumption defined from the logical perspective. Accordingly, the comprehension and the optimization of application on the logical level should directly impact on the physical level, i.e. on the actual power absorption.

The following section presents in detail the research roadmap for the definition of such methodology.

---

1 http://www.itrs.net.
5 A RESEARCH ROADMAP

Our research roadmap focuses on the definition of a methodology that allows the comparison of different applications from the efficiency of energy consumption perspective. Our research roadmap includes the following steps:

1. Comprehension of the problem and study of the state of the art.
2. Identification and operationalization of proxy metrics for computational complexity and thermodynamic depth of specific software applications.
3. Implementation of a tool that measures these metrics.
5. Analysis of data and identification of the most representative proxy metrics.
6. Integration of the results in a software tool to support IT managers in assessing software energy efficiency.
7. Evaluation of the impact of energy costs on the Total Cost of Ownership of an application.

After a first step focusing on getting a more detailed comprehension of the problem, we plan to define benchmarking methodologies to compare different applications and, finally, to propose optimization approaches.

The definition of our methodology requires a thorough analysis of the boundary conditions for the execution of the applications that should be analyzed. Therefore we will perform our analyses on a number of different configurations of hardware infrastructures.

We assume that the commutation energy $E(f)$ of Expression (2) is constant for a given hardware setup (please note that also the frequency $f$ can be made constant for current hardware setup by disabling dynamic frequency adaptation mechanisms such as Intel SpeedStep or AMD PowerNow!). As a consequence, we will focus our analyses on the assessment of energy consumption inefficiencies caused by computational complexity $C_c$ and thermodynamic depth $T_d$.

From a theoretical perspective, the minimization of the computational complexity term would require to evaluate how far the actual computational complexity $C_c$ of a given application is from its lower bound, that is from the logical depth $L_d$ of the output that the application is intended to produce. However, as noted in Section 4, such solution is really hard to achieve, if not unsolvable at all (at least, given the current state of the art). First, it would require a general methodology for defining the computational complexity of a generic problem. Second, the actual computational complexity $C_c$ of the application should be properly determined. Third, a way to identify the shortest program that solves the problem should be determined (that is, define the logical depth $L_d$ of the problem). Fourth, a comparison between the values of $C_c$ and $L_d$ should be performed in order to evaluate how far the application is from the theoretical optimum. Since phases two and three cannot be completely automated, and would require to identify the minimum logical depth for each possible problem (which is a clearly not satisfiable requirement), we decided to focus on the definition of benchmarking methodologies to compare specific applications.

One of the first issues to be faced is to find suitable proxy metrics for computational complexity and thermodynamic depth of a given application.

Lloyd (2001) suggests a list of 42 different complexity metrics that could be used to characterize the complexity of a system from three different (yet complementary) perspectives: $a$) how hard is it to describe, $b$) how hard is it to create, and $c$) what is its degree of organization. Since our focus is to characterize the complexity of a software system, we plan to operationalize and apply such measures of complexity (or a subset of them) to software systems. Along with these new metrics, we are also
going to consider classic software quality metrics as validation terms of comparison, such as the McCabe's Cyclomatic Complexity (McCabe, 1976), the Halstead's Software Science (Halstead, 1977), and the design quality metrics for object oriented systems proposed by Chidamber and Kemerer (1994) and Brito e Abreu (1995). These metrics are not direct measures of computational complexity, but of software design quality and cohesion. Although there is not yet any empirical proof that these metrics are correlated with computational complexity, an high quality software is usually well structured and its operations follow a logical flow. Software quality metrics may be proxies, or indirect measures, of computational complexity. This hypothesis will be verified during the research.

With regard to the thermodynamic depth term, we acknowledge that different representations of the same computational problem (e.g., the file structure adopted to store data) can be more or less efficient, as well as the fact that different ways of representing single bits can have different commutation energy requirements. However, these issues will be included in our future works. Given the current absence of consolidated metrics for thermodynamic depths, we will perform our analyses on software systems and data sets of comparable dimensions, so to be in a situation of comparable thermodynamic depth.

After the definition of the theoretical framework and the operationalization of variables, we will develop a tool to measure these metrics by analyzing the code of an application.

We will then assess which subset of our metrics best represent computational complexity and thermodynamic depth and thus could be used to operationalize Expression (4). This will require to compare the data gathered by our tool with measures of actual energy consumption.

The measurement of energy consumption ($EC_{physical}$) related to the execution of specific applications will be performed through the use of ammeter clamps. Such methodology is commonly adopted (e.g., Isci and Martonosi, 2003) since it does not require particularly instrumented hardware, nor the definition of instruction-level energy consumption models for the target microprocessor (as done for example by Tiwari et al., 1994). Figure 4 shows the details of the power measurement setup that will be used to measure the actual power absorbed by a microprocessor.

We will conduct our experiments on a number of different hardware configurations, and for each configuration we will analyze the relationship between actual power consumption and the set of measured metrics. Metrics will be gathered for a sample of Open Source applications similar for domain, functionalities, and language. Different classes of applications will be considered and comparative analyses will be performed within each domain and for each specific hardware setting. Application classes will be selected according to relevance, usage (optimization is convenient only if usage is high), and also availability of a minimum number of Open Source projects. For example, ERP and DBMS systems are likely to be included in the analysis as they both respect all the criteria listed above. For each class specific scripts will be implemented to automatically generate benchmark workload and compare the consumption of the different applications. For example, for DBMS different kind of queries will be considered (e.g., CREATE/ INSERT/ DELETE on a test database with 20 fields and 1.000.000) and for ERP different activities flow will be created (e.g., create a new order, receive the goods in the warehouse, etc.).

Figure 4 – Power measurement setup.
We consider the realization of a software tool that can help IT managers in assessing the power consumption efficiency of software applications the first milestone of our research roadmap.

Finally, we will evaluate the impact of energy cost on the Total Cost of Ownership of an application and, more generally, of an information system.

6 CONCLUSIONS AND FUTURE WORK

In this paper we have proposed a research roadmap to identify a set of software complexity and quality metrics that may be used to assess the energy efficiency of a specific application. We plan to validate our theoretical framework by measuring the actual power consumption on a number of different hardware systems. Our research will result in a tool able to extract a set of energy-related metrics by analyzing the code of an application. Our tool will allow to compare the energy efficiency of two or more applications with the same functionalities, thus enabling a green-aware choice. Project managers, software developers, and software buyers will greatly benefit from our research as they will be able to assess the differences in power consumption among a set of software applications. An issue that will need investigation is how energy consuming will be the code analyzer tool, as it may turn out that the energy required to analyze and optimize an application outweigh the savings obtained. However, optimization costs occur one time only, whereas savings are repeated every time the application is executed. Moreover, most of the currently available code-based metrics can be easily and quickly computed by parsers and code analyzer tools.

After the first phase of our research, we foresee that we will extensively apply our methodology to Open Source applications so that we will gather a significant quantity of data to analyze. Our research will then focus on optimizing the energy efficiency of applications by identifying development best-practices.

References

F. Renzi, “Evolution scenarios for data centers technologies and architectures and impacts on energy consumptions”, IBM, Presentation held at the conference “Green ICT”, 22nd November 2007, Milan, Italy.
F. Bruschi, “From Green Grid to Efficient Enterprise”, Schneider Electric, Presentation held at the conference “Green ICT”, 22nd November 2007, Milan, Italy.
ENVIRONMENTAL RESPONSIBILITY AND GREEN IT: AN INSTITUTIONAL PERSPECTIVE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0477.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Green Computing/Green IT, Institutional theory, IT regulation, Regulation / Deregulation</td>
</tr>
</tbody>
</table>
ENVIRONMENTAL RESPONSIBILITY AND GREEN IT: AN INSTITUTIONAL PERSPECTIVE

Abstract

A recent Gartner Research report found that environmental concerns are increasingly exercising the minds of business and IT executives. This is reflected in the growing interest in the adoption of environmentally responsible approaches to the deployment, operation and use of IT. However, for the majority of firms, issues of cost reduction and energy efficiency appear to predominate. This paper argues that whether it is an interest in Green IT or in cost reduction, the concerns of business and IT managers are modulated by regulative, normative, and cultural-cognitive influences in the institutional environment. The study therefore applies institutional theory to develop a series of theoretical propositions which specify the effect that such influences have in shaping environmental responsibility in organisations. Important as such a theoretical contribution may be, there are, however, the pressing practical imperatives of formulating Green IT strategies, achieving energy efficiencies, and reducing carbon footprints—thus, the study also contributes to a practical understanding of the complex institutional influences at play in shaping such imperatives.

Keywords: Institutional Theory, Environment, Green IT
1 INTRODUCTION

The design, manufacture, deployment, operation, use, and disposal of information technology (IT) has significant implications for the environment (Aberdeen Group 2006). For example, IT (computers, networking, data and telecommunications technologies) consume significant amounts of energy when in use and contribute in no small way to the growth in greenhouse gas (GHG) emissions. Energy issues aside, IT artefacts also contain regulated and hazardous materials. Accordingly, the influence of environmentally-oriented regulatory and social pressures are especially evident in the IT sector (Murugesan 2007), particularly in manufacturing, but also in adoption, operation and use (Murugesan, 2008). IT is increasingly viewed as a significant and growing part of the environmental problem—something that few people take seriously, particularly IT executives (CNET Networks 2007).

This paper has as its focus Green IT. In the past, IT was designed, manufactured, and applied in business enterprises without regard to its environmental impact. In recent years, the term Green IT was coined by practitioners to differentiate between IT artefacts that had been designed with environmental sustainability in mind—that is, to contain the minimum amount of hazardous materials, to be energy efficient during the use period of their lifecycle, and to be disposed or recycled with the minimum effect on the environment and human health (Murugesan 2007). The IT industry has, of late, extended the scope of Green IT to include the manner in which IT is deployed to help reduce GHG emissions, to introduce energy efficiencies, or to reduce energy consumption (Adjugo 2008a,b; Gartner 2007, 2008). Academics such as Boudreau, Watson, and Chen (2008), argue that therefore IT-enabled information systems have a role to play in making business processes environmentally sustainable—they term such information systems Green IS. This study conceptualises Green IT as including all of the above, in order to avoid the possibility of definitional confusion and to maintain congruency with practitioners’ use of the term. This paper also introduces the concepts of corporate social responsibility (CSR) and corporate environmental responsibility. The link between Green IT and CSR has been commented upon (cf. Ryan 2008b); as has the relationship between CSR and corporate environmental responsibility (cf. Eisner 2004). For example, it is now common practice for many business enterprises to describe policies or measures aimed at environmental sustainability as part of their statements on corporate social responsibility—Green IT strategies are now included under this heading (cf. Ryan 2008a on Dell and Hewlett Packard).

In its recent Green IT Survey of Belgian executives, Adjugo (2008b) reports that 57% of respondents did not view Green IT as being part of their companies’ CSR (corporate social responsibility) initiatives. Nevertheless, Gartner reports that IT needs to become greener as energy is due to become a significant component of IT budgets, rising from 10% to 50%. Add to this the costs of planned carbon taxes and compliance with future environmental regulations. Current estimates suggest that IT accounts for over 2% of global CO₂ emissions (Gartner 2007)—roughly the same amount as generated by air travel. IBM (2008) argues that datacenters account for 2% of CO₂ emissions, implying that the true figure is much higher than this. Whichever figure is most accurate, the true cost to the environment of the design, manufacture, transport, and disposal of IT is much higher than estimated (Murugesan 2008).

Presently, there is little doubt of the impact that CO₂ and other greenhouse gases are having on the climate: consequently, all sectors should be taking responsibility to lower their GHG (Johnston 2008). However, a recent study by IBM argues that a majority of organizations do not possess a green IT strategy, let alone the capabilities to build one (cf. Gartner 2008, Adjugo 2008a). The reasons for this state of affairs are found in recent research by the UK’s Corporate IT Forum. This study on 150 large member organisations reported that while 81% of IT executives indicated that Green IT was higher up the corporate agenda than in 2007, it was just in eighth place behind security, legacy software, data quality and server consolidation on their list of priorities (Cooter 2008). Thus, issues of strategy aside, the costs associated with the transition to Green IT architectures are viewed as being problematic for business and IT executives. Both problems are related, however, as Cooter reports on a statement by
Ian Campbell at the European Green IT Summit in London to the effect that “managers often struggled to present a sound business case -- that was the biggest inhibitor” to the adoption of Green IT strategies. Campbell is quoted by Cooter as stating that perhaps the biggest problem faced by organisations is in measuring the impact of green IT policies. He indicated that there was an absence of standards—national and international—with which to evaluate and benchmark green initiatives. It is no doubt, with such issues in mind, that a recent study by Gartner Research (2008) argued that environmental concerns will increasingly impact on organizational decisions to purchase, deploy, operate and dispose of information technology (IT). There is, therefore, a need for business and IT managers to be knowledgeable of, and proficient in, all issues relating to the procurement and use of IT viz, the management of energy and power issues; data centre design and operation; IT server virtualization; regulatory compliance and eco-labelling of IT artifacts (in regard to the use of hazardous substances); green metrics and audits; teleworking and so on (Murugesan 2008).

Given the broad scope of the Green IT problem, this study will focus on the operation and use of IT to achieve green objectives. However, as Green IT is now a relatively mature topic for IT manufacturers, examples of their experiences in addressing regulative, normative, and cultural-cognitive institutional influences on their activities will be employed to illustrate the challenges that may confront end-using business enterprises. The question that arises from this introductory analysis and which guides the conduct of this study is: What institutional forces act to shape the ‘environmental concerns’ of business and IT managers in making the operation and use of IT green, and how can Green IT objectives be realised?

Green IT is a new and emerging issue for IS researchers. Webster and Watson (2002, p. xiv) recommend that “authors could tackle an emerging issue that would benefit from exposure to potential theoretical foundations” through the medium of a review paper. This paper follows Webster and Watson in addressing the phenomenon of the operation and use of Green IT. The first section of this paper therefore examines the forces shaping the ‘environmental concerns’ of business and IT managers through the lens of institutional theory.

2 APPLYING INSTITUTIONAL THEORY TO GREEN IT

Richard Scott (2004) argues that institutional theory is concerned with the influences that shape social and organisational structures, schemas, rules, norms, routines and, ultimately, the behaviour of social actors. According to Scott (2001, p. 33), institutions consist of “cognitive, normative, and regulative structures and activities that provide stability and meaning to social behaviour.” Thus, Scott (2001, 2004) argues that regulative, normative and cultural-cognitive institutional forces shape organizational processes and structures—such forces prescribe what is effective performance or efficient operation in organizations (Powell 1991). Focus is brought to this broad theoretical lens by DiMaggio and Powell’s (1983, p. 143) concept of an ‘organisational field’. An organisational field is constituted by a collection of “organizations that, in the aggregate, constitute a recognized area of institutional life”; inter alia, the field consists of “key suppliers, resource and product consumers, regulatory agencies, and other organisations that produce similar services or products.” While environmental ‘concerns’ shape Green IT responses across multiple organisational fields, due to the ubiquity of IT, the concept is nevertheless useful. Thus organisational fields are characterized by regulative and legislative (coercive) influences from government departments, state-sponsored agencies, the judiciary, and so on, in addition to normative and cultural-cognitive (mimetic — DiMaggio & Powell1983) influences from related organisations (professional and standards bodies, non-government organisations (NGOs), consulting organisations, professional bodies, etc.), and society at large (NGOs and community groups, for examples).
2.1 Regulative Influences on Corporate Environmental Responsibility

It is clear from Scott (2001) that the emphasis of regulative institutional influences is on coercion, indicators of which are rules and laws, which agents such as governments and regulatory agencies legitimize using legal mechanisms or sanctions to ensure Green or environmentally responsible behaviours or initiatives. Regulative institutional carriers, on the other hand, are social structures such as governance and power systems, which institute rules and laws, the organisational response to which is to institutionalize routines such as protocols and procedures (Scott 2001). Drawing on the introduction to this paper, we may now state our first general theoretical proposition:

Proposition 1: Governance and power systems, rules, and laws, will give structure to and shape organisational responses, procedures and protocols around the institution of Green IT initiatives.

Unlike organisational and societal end-users of IT, IT manufacturers are currently confronted with diverse regulations governing the design, production and performance of computers and other IT artefacts (Aberdeen Group 2006, Kellow 2006). Briefly, such regulations include the European Union’s (EU) Waste Electrical and Electronic Equipment Directive (WEEE), the Restriction of Hazardous Substances Directive (RoHS), the Registration, Evaluation and Authorisation of Chemicals (REACH) Regulation, and the Eco-Design for Energy Using Products (EuP) Directive (Hristev 2006). All of these place onerous compliance imperatives on IT manufacturers. Indeed, the rest of the world is catching up with the EU. In the US, for example, the Environmental Protection Agency (EPA) has numerous regulations covering environmental issues and hazardous substances across the whole range of manufacturing sectors. However, individual US states such as California are following the EU’s lead in introducing exacting WEEE and RoHS-like standards. Other jurisdictions are no less stringent as is the case with Japan, which also has highly demanding laws, while over the last two years, Korea, Australia, China and Canada have introduced legislation similar to the RoHS and WEEE directives.

The total amount of electricity consumed across all sectors in Europe is estimated to be growing at an annualised rate of 2% (Salhofer, Schiffleitner, and Stachura, 2008). Significantly, research by McKinsey (2007) indicates that energy consumption by computers and IT in general is projected to increase at an annual rate of approximately 3% to 2030—much higher than the overall trend. However, that estimate is conservative compared to that of Gartner (2007), who project that the energy portion of IT departments’ budgets is due to rise from 10% to 50% in a much shorter time period, with datacentres responsible for the lions share of this. Thus, corporate datacentres are singled out for special attention, as the EPA calculates that they presently account for an unacceptable 1.5% of US energy costs. In a different context, Alain Bandle VP and GM Dell Europe argues that the IT sector is responsible for 5% of total global consumption of electricity (850 terawatt hours) 520 million tones or 1.6% of global CO2 emissions (Bandle, 2008). These projected increases in consumption needs to be balanced with pressures to reduce consumption.

According to Peter Johnston (2008 p. 19), “[i]n the European Commission, we have identified ICT-enabled improvements in energy efficiency as one of the potentially most cost-effective ways in which Member States can meet their 2020 targets.” He adds that “Energy efficiency is therefore emerging as the cornerstone of energy and climate policies.” Take, for example, the EC Directive 2005 32/EC which focuses on energy efficiency in products: this sets a challenging target for energy savings of up to 9% in the period 2008-2016. Furthermore, the EU Green Book on energy efficiency estimates that energy savings of up to 30% for businesses is possible by 2020. Clearly, much will need to happen if residential and business electricity consumers are to meet EU targets. Failure to do so will, no doubt, result in further regulation in the area as energy generation in EU member states is responsible for over 24 % of the communities greenhouse gas (GHG) emissions. Thus, as EU member states fail to meet the greenhouse gas emissions targets, business organisations will be faced with regulations governing their energy efficiency. The present approach of depending on market forces and the price mechanism to control inefficient energy use by businesses will be replaced by legislation and energy audits. Thus, the option to go ‘green’ may not be left in the hands of business organisations for much
longer (cf. Johnston, 2008). The European Union continues to build a raft of environmental regulations and directives, the outcome of which will be increased carbon taxes and the imposition of further environmental compliance imperatives on GHG emissions (cf. Salnofer, Schiffleitner, & Stachura 2008). Likewise, Stephen Harper (2008) of Intel predicts that the US will adopt similar measures. Indeed, there is evidence that the US congress is already thinking along these lines (King 2007, cf. EPA 2007). Thus, business and IT managers which Cooter (2008) found to be recalcitrant in lowering the energy consumption (and associated GHG emissions) due to the cost of upgrading and replacing IT infrastructures, may have little choice in the matter in the not-too-distant future.

Take, for example, that the Eco-Design for Energy Using Products (EuP) Directive is presently being rolled out across electrical and electronics product groupings, including IT artefacts such as computers and monitors. Although still being formalised, it is being phased in and implemented across EU member states. The EuP Directive will require IT manufacturers (among others) to make voluntary declarations (presently) on the energy used in the design, packaging, delivery, and recycling of products across supply chains, in addition to the energy consumed during use. This new directive significantly extends Energy Star-like standards in EU member states such as Blue Angel in Germany Nordic Swan in Scandinavia. However, current thinking with the European Commission indicates that member states will have to regulate to have widespread efficient use of EuP, if they are to stand any chance of meeting GHG emission targets. The ever-increasing energy consumption of IT concerns the Environmental Protection Authority (EPA), as its recent report to the US Congress indicates (EPA 2007). With electricity cost and consumption rates to increase going forward, and with power shortages already evident in certain parts of the US, the EPA’s report makes some strong recommendations involving both carrots (i.e. incentives to conserve energy) and sticks (an energy efficiency tax for inefficient organisations)—all of which has serious implications for business enterprises, if adopted by the US government.

Proposition 2: Business and IT managers are more likely to be environmentally responsible (e.g. implement Green IT) if they are coerced into doing so by rigorous and comprehensive environmental regulations (e.g. aimed at reducing GHG emissions).

Thus, it is clear that in the absence of strong self-governance in relation to energy consumption, regulators in the EU, the US and elsewhere will become increasingly active in demanding compliance with energy efficiency and GHG emissions regulations and adherence to demanding targets (cf. Campbell 2007). It is certain that such measures will favour Green IT strategies that focus on structural power consummation avoidance with its emphasis on reducing the installed power base, as opposed to temporary avoidance strategies, which focus on the optimization of energy utilisation, without reducing the installed power base. Certainly, organisations that have made substantial efforts and costs to become greener will reap the first-mover benefits in terms of compliance, as has already been seen in the electrical and electronics sector (Aberdeen Group 2006).

2.2 Normative Influences on Environmental Responsibility

Scott (2001) argues that the emphasis of normative influences is on social obligation as a basis of environmental responsibility or Green behaviours: the basis of legitimacy for normative forces is moral governance. Scott also emphasises that the institutional carriers of such influences are the values and expectations that develop in cultures and organisational fields. Other carriers include authority systems in social structures and routines that reflect conformance and performance of duty.

In Joel Bakan’s The Corporation, Peter Drucker argues that executives who are concerned about the environment should be fired “...and fast”. Likewise Bakan reports that the economist and Nobel Laureate Milton Friedman holds similar views and observes that the only time that corporate social responsibility can be acceptable is when it is insincere. What Friedman means here is that ‘going green’ can only be a means to an end—the ends being to enhance brand image and increase market share in order to maximize shareholder wealth. Obviously Friedman’s views did not change since his 1970 article in the New York Times Magazine titled “The social responsibility of business is to
increase profits.” Indeed as Bakan (2005, p. 35) states “Corporations are created by law and imbued
with purpose by law. Law dictates what their directors and managers can do, what they cannot do, and
what they must do.” Indeed, this point weakens the argument for self-regulation and supports
Proposition 2.

In the four years since Bakan penned his argument, the business environment has changed. William
Ford, Jr., Chairman of the Ford Motor Company, argues, for example, that “corporations could be and
should be a major force for resolving environmental and social concerns in the twenty-first century”
(cited in Bakan, 2005, p. 31). Klaus Schwab, Executive Chair of the World Economic Forum, argues
that corporate social responsibility (CSR) focuses on the broader financial, social and environmental
effects of all that a company does (Schwab, 2008). Thus, companies who exercise CSR minimize any
negative effects of their commercial activities by taking responsibility and being accountable for
economic externalities in all spheres of its operations (Bakan 2005). Traditionally, companies attended
only to the bottom-line—their financial results, which usually did not account for externalities. Firms
practicing CSR, on the other hand, attend to a triple bottom-line, which includes being accountable for
what they are doing in terms of environmental and social responsibilities, in addition to their financial
obligations of maximizing profits (Schwab 2008). Schwab goes on to argue that CSR encompasses the
entire value chain to include suppliers and customers. The exercise of CSR by Ford Motor Co. reflects
this, as it requires all suppliers to be IS14001 compliant, thus using the supply chain as a vehicle to
drive environmental responsibility. In regard to Green IT, it is clear from Collett (2008) that CSR and
other environment-oriented initiatives introduced by industry standards bodies such as the Institute of
Electrical and Electronics Engineers, the Global e-Sustainability Initiative (which includes Microsoft,
Sun, Ericsson, Bell, BT, Dell, HP, and many others: cf. www.smart2020.com), the US-based
Electronic Product Environmental Assessment Tool (EPEAT) organisation, and so on, that self-
regulation via normative mechanisms is possible, this pleads to the following proposition:

**Proposition 3**: Organisations are more likely to be environmentally responsible if there are well-
defined systems of self-regulation with tangible recognition for rewarding often costly Green IT
initiatives.

Normative standards on computer environmental footprints are, as indicated, the Blue Angel in
Germany and Nordic Swan in Scandinavia. However, the most influential normative initiative is
undoubtedly that of the Environmental Protection Authority in the US which has a major focus on
energy-consumption-related GHG emissions. In 1992, the EPA’s instituted its Energy Star
programme—which concentrates on the energy use of products in use. Almost all major manufacturers
now produce Energy Star compliant products. Notably, only products in the first quartile (25%) of
energy efficiency standards set by the EPA and U.S. Department of Energy are awarded an Energy
Star rating. In 2006, the Green Electronics Council, which includes the EPA, U.S. Department of
Energy, and industry standards bodies introduced the Electronic Product Environmental Assessment
Tool (EPEAT). EPEAT is based on the Institute of Electrical and Electronics Engineers (IEEE) 1680
Standard, Section 4 of which governs environmental performance criteria for desktop PCs, notebooks,
and PC monitors. In terms of IEEE 1680 Standard Section 4, EPEAT is concerned with the (a) the
reduction/elimination of environmentally sensitive materials; (b) materials selection; (c) design for end
of life; (d) product longevity/life cycle extension; (e) energy conservation (Section 4.5 covers
adherence to the Energy Star standard); (f) end of life management (take-back and recycling); (g)
corporate performance (in terms of CSR); and (h) packaging (toxics, labelling). While the US
Government did not regulate on the IEEE standard or EPEAT, an Executive Order was issued by
President Bush in 2007 requiring all US federal agencies to purchase only EPEAT-certified artefacts
(Ryan 2008a) . Significantly, in July of 2007, a year after EPEAT’s launch, Energy Star 4 came into
force, which is a much more rigorous standard than its predecessor Energy Star 3. Currently, Hewlett
Packard (HP), Dell, and Toshiba manufacture computers to reach Gold standard certification (there are
also Silver and Bronze certificates). The success of the Energy Star programme is reflected in the fact
that in 2007 devices designed to Energy Star specifications resulted in savings of 40 million metric
tons of GHG emissions; this is equivalent to taking 27 million vehicles off the road annually (Hojlo
and Jacobson 2008). Yet further evidence of the effectiveness of standards in promoting environmental responsibility is the Leadership in Energy and Environmental Design (LEED) Standard, which is a Green Building Rating System developed by the U.S. Green Building Council. Deloitte is implementing those standards in the construction of its datacentre in Dallas (Desmond, 2008).

It is clear that practitioner publications, on- and off-line, such as Computerworld, NetworkWorld, CNET, the Cutter IT Journal have been raising awareness of late on environmental issues, particularly the issue of energy efficiency as it relates to Green IT. So too have bellwethers like Gartner Group, Forrester Research, the Aberdeen Group, trade and industry conferences such as Electronics Goes Green (EGG2008+), and so on. All this, in addition to normative prescriptions from the ISO and IEEE standards bodies, NGOs like the EPEAT, and so on, influences IT practitioners’ decisions. This leads to the following proposition:

**Proposition 4**: Business and IT executives are more likely to introduce environmentally responsible programmes, if Green IT is promoted by, and institutionalised through, industry standards, business publications, consultancy bodies, and other fora in which executives participate.

2.2.1 **Normative Influences of Stakeholders on Cost Reduction Vs. Green IT**

Returning to the point made by Milton Friedman, that if the social responsibility of business is to increase profits, then it is logical to assume that business will take steps to reduce costs if this increases profits. Thus, the findings of research by Forrester (2007) indicates that 55% of business and IT executive surveyed saw cost as the chief motive for instituting sustainable IT operations, not Green IT or concerns for environment-related energy efficiencies. It must be remembered, however, that it is only relatively recently that end-user organisations not in the IT sector, e.g., in the services sectors perceived the benefits of adopting ‘Green Strategies’. For example, Collett (2008) reports that Wachovia Corp. was ranked 12th in Computerworld’s survey of Top 12 Green IT Users in 2008. Likewise Desmond (2008) reports on Deloitte’s corporate Green IT initiative. Yet, there is not a widespread recognition by business and IT executives that costs can be reduced and profits increased through Green IT strategies. Significantly, Duffy (2008) reports that “80% of companies recently surveyed by Nemertes [Research] have no corporate green policies; only 13% knew datacenter energy costs; only 3% turn off their servers when not in use; and desktops are left on 50% of the time.” Clearly given the opportunity of cost reductions, IT are not serving stakeholders needs or exercising environmental responsibility. Yet as the following examples illustrate, the cost savings can be significant for relatively simple procedures.

Consultants from Adjugo (2008a) argue that “[i]mplementing quick wins like switching off computers after business hours can reduce the energy consumption up to 75% per year.” However, such straightforward changes in the management of workstations can have significant implications for an IT function’s methods of operation viz. conducting workstation and server software upgrades, patches and backups outside of normal business hours. One might be forgiven for thinking that such simple procedures would be within the ambit of business and IT managers; yet the opposite appears to be the case if stories carried in practitioner publications are to be believed. There are, nevertheless, several powerful examples of significant cost savings.

**Proposition 5**: Business and IT executives are more likely to introduce environmentally responsible programmes, if they are aware of the costs reductions associated with the implementation of Green IT initiatives and the resulting increase in profits.

2.3 **Cultural Cognitive Influences**

Cultural-cognitive influences that emphasize environmentally responsible behaviours are ‘taken for granted’. Beliefs and attitudes toward compliance are socially constructed and transferred mimetically, while being governed by a logic of orthodoxy. Institutional carriers include, according to Scott (2001),
socially constructed cultural categories and typifications and routines that are reflected in performance programs and scripts. Of interest to the present study are social structures that result in isomorphism of social and organisational practices.

Climate change and environmental concerns has led investors and investment managers to focus on green investments (Mincer 2007). Traditionally, investors have examined risk, liquidity and rate of return as investment criteria; now, however, an increasing number of investors look toward social and environmental factors (Kahlenborn 1999). Kahlenborn (ibid., p. 65) presents two contrasting definitions of ‘green investment’: in the first, it “can be understood as any form of financial investment whereby the investor pays attention to ecological goals as well as the traditional aims of investment. On the other hand, ‘green investment’ can be understood as an investment that successfully counteracts negative influences on the environment, or serves to produce goods or offer services that have positive effects on the environment.” Over the last ten years there has been a significant increase in individual investment and mutual funds focusing on what is in 2008 an established niche market (Kahlenborn 1999, Mincer 2007). Thus, in highly competitive business environments where ‘trust’ is an issue, and investor confidence seriously damaged, business enterprises that display CSR and act on environmental concerns may shape customer preferences and investment decisions in opting for ‘green strategies’ which could tap into an additional source of investment funding (cf. Campbell, 2007). It is clear that such funding is not limitless, so firms will also have to compete on their environmental credentials.

**Proposition 6:** Business and IT executives are more likely to introduce environmentally responsible programmes such as Green IT if doing so enhances their image and attracts Green Investors.

Thus, is the environmentally responsible behaviour of William Ford out of step with his corporate colleagues as Joel Bakan’s thesis would indicate? In an industry whose products are perhaps the most polluting and environmentally unfriendly when is use, presenting an environmentally positive image can only contribute to the bottom-line. So, given the nature of Ford’s business, are he and his company genuine in exercising CSR, or is this a form of ‘greenwash’? ‘Greenwashing’ products and brands has become commonplace as companies increasingly use a product’s ‘greenness’ as a basis for competition. Complaints to the UK’s Advertising Standards Authority (ASA) regarding ‘greenwashing’ in product advertising have increased dramatically in the recent past (Charles 2007): for example, complaints to the ASA averaged over 56 per month in the last two quarters of 2007. Likewise, a signal that all is not well in the IT sector comes from James Staten, a principal analyst at Forrester Research. He recently warned IT executives about the level of ‘greenwashing’ taking place in the industry, as manufacturers imply that whole product lines are green, when in fact only a proportion of products may be. The consequences of ‘greenwashing’ are significant, as stakeholders tend to lose confidence in a firm’s ability to be honest leading not only to potential damage to market share and the bottom line, but also to the withdrawal of valuable sources of investment (Zagenczyk 2004).

Greenpeace is one NGO that is especially significant in shaping public opinion regarding corporate environmental responsibility. They use their website to publish the results of studies and analyses of green and not-so-green products. Whether accurately or not, they also rank manufacturers based on their eco-friendly products and strategies. Take for example their Apple webpage ([www.greenpeace.org/apple/itox.html](http://www.greenpeace.org/apple/itox.html)) which states “Apple products – sleek looks, amazing design, and meticulous attention to detail. So what’s with the toxic chemicals inside, short life spans and allowing their products to be dumped in Asia?”. It must be noted that Apple launched its Greener Apple initiative in 2007, much to Greenpeace’s annoyance, as it considered Apple to be greenwashing. Companies like Nokia, on the other hand, are lauded for their efforts at environmental responsibility. This is one of the reasons why the Greenpeace presentation at the Electronics Goes Green 2008 conference in Berlin was, perhaps, the best attended with standing room only in the aisles. Clearly, NGOs like Greenpeace can shape public perceptions of corporations and products, hence they can exert cultural cognitive as well as normative influences on business activities. The press also shapes public perceptions and various publications are eager to report and rank firms based on their
approaches to corporate social and environmental responsibility. Take for example, McLean’s in Canada, whose Technology and Media section reported that Hewlett Packard’s efforts at social responsibility received an A+ ranking, while Dell and IBM achieved an A, NOKIA an A-, and Nortel and B+. Thus, the watchful eye of society at large is upon producers, with the media industry being all too ready to highlight issues of concern. Given the foregoing arguments, we now present the following proposition:

**Proposition 7:** Business and IT executives are more likely to introduce environmentally responsible Green IT programmes if their activities, processes and products are being monitored and reported upon by independent non-government organisations, the press, and society at large.

It is clear from Corporate Annual Reports (e.g. Dell, Intel and HP—Cf. Ryan 2008a,b) that corporate social and environmental responsibility is viewed as being increasingly vital to promote public perceptions and to add to the bottom line (cf. Campbell 2007, Adjugo 2008 a,b, Collett 2008). Thus, institutional theory would predict that mimetic behaviours would have business and IT managers imitate each other in terms of Green IT strategies. This is certainly occurring in large global corporations. Likewise, the increasing evidence of greenwashing has business ‘appearing’ to imitate. In the event that businesses do actually imitate each other and implement Green IT initiatives, then isomorphism that will benefit society as a whole will occur across firms. We may now present our next and final proposition.

**Proposition 8:** Business and IT executives will mimic successful Green IT strategies of others leading to industry-wide isomorphism in terms of environmental responsibility practices with regard to the operation and use of IT.

### 2.4 Directions for Future Research

This paper employs institutional theory to develop applied theoretical propositions; however, these require further elaboration in process-based studies, whether positivist, post-positivist or interpretivist. Such studies might employ the propositions as a theoretical lens in, for examples, exploratory or explanatory case-based research, using analytic generalisation or replication strategies (Yin 2003). The continued use of institutional theory is recommended, of course; however, following Currie (2009), we argue that multiple levels of analysis that encompass the institutional environment, organisational field, organisation and social actors would be of greatest utility. These could, for example, identify the institutional arrangements (i.e. specific Green IT strategies) organisations deploy in response to forces in the institutional environment, or the organisational field (DiMaggio and Powell 1983). In-depth case studies could also be used to inform practice as to why and how business enterprises are adopting and applying Green IT to reduce GHG/CO$_2$ emissions and introduce energy efficiencies. Such studies might also help identify how software solutions (or Green IS, if you will) can be designed and developed to make business and IT processes more environmentally sustainable across the business (cf. Boudreau et al. 2008).

Following Wheeler (2002), positivist or variance studies could build on process-based case studies to generate hypotheses and identify empirical indicators to subject the theoretical propositions/meso-level applied theory to test in order to attempt to falsify it. Here again, however, Currie’s (2009, p. 66) cautionary observation on the use of institutional theory that “Many variance and process studies examine the cause/effect relationships between IT-related constructs, such as adoption intention, assimilation and implementation, without considering the wider environmental and inter-organisational levels. Such an oversight is problematic, both theoretically and empirically, since the mainstay of institutional theory is an emphasis on multi-level and multi-stakeholder analysis.” Thus while institutional theory can contribute to a greater understanding of Green IT, Currie’s recommendations should be observed if such an understanding and theoretical contribution is to be achieved.
3 CONCLUSIONS

This paper examines a new and emerging research topic for IS researchers. Indeed, with IT literally up there with airline travel as a contributor of greenhouse gas emissions, the greening of IT should be uppermost in the minds of practitioners, as well as being high on the IS field’s research agenda. Yet despite the threats to the environment brought on by climate change, it appears from industry-based research that IT practitioners are leading the way only in their ignorance of the impact of energy inefficient practices relating to the deployment, operation, and use of information and communication technologies (ICT). Campbell (2007) argues that corporations behave in a socially responsible way when economies perform well, and are less than interested in exercising this responsibility, or are unable to do so, during economic downturns. It would be reasonable to assume that in the buoyant economic conditions that existed for the majority of firms over the past 10 years that business and IT executives would have exercised corporate environmental responsibility by making IT use more energy efficient and thereby reducing both related costs and greenhouse gas emissions. We conclude that they did not initiate energy conservation measures due to a general ignorance of the consequences and a failure to self-regulate. To be sure, there are many notable exceptions to this, but the facts speak for themselves—the majority did nothing, despite enormous increases in energy costs since 2004. Hence, such executives failed also in their primary objective of maximising shareholder wealth through increasing profits by lowering costs, both in the short term and over the long term.

This paper’s application of institutional theory to the phenomenon of Green IT indicates that a small but growing proportion of IT practitioners are responding to normative and cultural-cognitive influences and making their IT infrastructures environmentally responsible. It must be noted that many of these firms produce IT products and services (e.g. participants in Global e-Sustainability Initiative) and have for some time now been subject to regulative influences with regard to their production activities. Hence, they have been sensitive to environmental issues. Their Green IT initiatives have not, we believe, gone unnoticed by others in their organisational fields. Thus, they have shaped the broader institutional environment through cultural-cognitive influences that have engendered environmental responsibility in and across businesses through mimetic responses. At the time of writing, there is a global recession and, with corporate budgets being trimmed, Green IT initiatives are likely to suffer if they require initial investment to obtain energy efficiencies that will reap future cost reductions. Yet, in the face of ever increasing energy costs, there is hope that companies will attempt some of the ‘quick wins’ described in the literature (cf. Ryan 2008a). Unfortunately, the majority of such ‘quick wins’ will be in the front office, as improvements in energy efficiency in back-office ICT are not so easy to attain, as reducing power consumption requires structural measures such as reduction in installed power capacity—e.g. reducing the number of servers and using virtualisation to optimise server loads. Significantly, data centres account for much of the growth in IT energy consumption and related greenhouse gas emissions. (Hence, Intel’s argument that thin-client architectures are not as energy efficient as energy efficient power managed workstations and notebooks.) Thus, significant investments are required here to reduce both consumption and emissions in the operation of server farms and data centres. At the time of this submission (Dec 1st), government officials across the globe are in Poland to attend the UN Climate Change Conference, which aims to frame a replacement for the Kyoto Agreement. Likewise, the UK government is today talking about decreasing GHG emissions by 20% before 2020. Thus, it is more than likely that, economic downturn or not, governments and regulators will look to the long term and regulate inefficient energy use by business enterprises in order to address overall greenhouse gas emission targets, which are not presently being met. Indeed, many now argue that the targets set out in international agreements such as Kyoto are to be revised upwards, as with the UK—all of this has significant implications for energy efficiency in the deployment, operation, and use of IT infrastructures.
REFERENCES


CNET Networks (2007). IT departments slow to adopt green tech, URL: http://news.zdnet.co.uk/itmanagement/0,1000000308,39287053,00.htm


COLLABORATIVE SOURCING – THE MOTIVATION AND DESIGN OF DEMAND SIDE COMBINATIONS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0022.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Collaboration, Alignment, Offshoring / Outsourcing, interorganizational systems</td>
</tr>
</tbody>
</table>
COLLABORATIVE SOURCING – THE MOTIVATION AND DESIGN OF DEMAND SIDE COMBINATIONS

Borman, Mark, The University of Sydney, Room 422 Building H69, Sydney, NSW 2006, Australia, m.borman@econ.usyd.edu.au

Abstract

Collaborative sourcing is defined as a group of organisations combining on the demand side as part of the outsourcing process. Based upon a synthesis of the outsourcing and collaboration literatures, this paper seeks to examine both the benefits of collaborative sourcing and the options available. A series of case studies identifies that a number of alternative collaborative sourcing modes are possible with differing motivations – based on economies of scale but moderated by other influences – intensity – ranging from simply aggregating demand at the contracting stage to establishing a separate entity to actually provide the services – focus – from IT infrastructure through to business processes – and extent – the number of partners involved and activities covered.

Keywords: outsourcing, alignment, collaboration, interorganisational systems
1 INTRODUCTION

Since its genesis the outsourcing literature has developed considerably from viewing outsourcing as a singular concept (Loh and Venkatraman, 1992a) to recognising that many different types of outsourcing are possible (Marcolin and McLeLlan, 1998). Variations have been identified based upon the scope of the arrangement (Sambamurthy et al, 2001), its timeframe (Lacity and Willcocks, 1998), whether one or several suppliers are involved (Currie, 1998), the type of activity concerned – for example differentiating between information technology (IT) and business process outsourcing (Dibbern et al, 2004) – its geographical location (Bergkvist and Fredriksson, 2008) and the nature of the relationship between supplier and client (Fitzgerald and Willcocks (1994). Several classifications have been proposed. Millar (1994), for example, suggested four different types of outsourcing arrangement – differentiated by purpose, inclusion and relationship. Cullen et al (2005) classified outsourcing arrangements based on scope, supplier grouping, financial scale, pricing framework, contract duration, resource ownership and commercial relationship. Willcocks et al (2007) identified alternative outsourcing models – including traditional, netsourcing and customer-supplier joint ventures.

One area that the literature appears to have neglected however is the potential for collaborative sourcing – defined here as a group of customers combining their demand as part of the outsourcing process. Gallivan and Oh (1999) recognise a class of outsourcing where a group of organisations come together to obtain a common service from a supplier, a review of the literature suggests that little research has been conducted on the phenomenon. An exception is the work of Seddon (2001) which examined the IT outsourcing by clusters of agencies within the Australian Federal Government – however this concluded that the approach does not work. Yet the potential benefits of collaboration in other spheres are well recognised (Smith, Carroll and Ashford, 1995) and a number of collaborative sourcing ventures have been initiated. In the UK, for example, Lloyds TSB Bank, Barclays Bank and HSBC bank have formed a joint venture for cheque processing (Roberts, 2004). As such it does not seem to be appropriate for collaborative sourcing either to be ignored or summarily dismissed as unworkable.

The remainder of the paper comprises two sections. The first reviews literature relevant to collaborative sourcing to develop a framework to better understand the potential motivation and identify dimensions that define the resulting design. As with outsourcing, successful collaborative sourcing requires a clear motivation or objective (Dibbern et al, 2005). In addition, the plethora of outsourcing options that have emerged – see earlier – suggest that it is important to clearly specify the design of the resulting collaborative sourcing. The second section presents the results of an empirical study of nine collaborative sourcing arrangements in Australia conducted to assess the usefulness of that framework.

2 THE MOTIVATION AND DESIGN OF COLLABORATIVE SOURCING

Given the absence of an established literature regarding collaborative sourcing, existing research on outsourcing and collaboration formed the twin starting points for a literature review to develop a theoretical framework for understanding the motivation for, and design of, collaborative sourcing.

A wide variety of motivations have been posited for outsourcing; including a desire to focus on core competencies (Lacity and Willcocks, 2001), a response to the actions of other organisations (Loh and Venkatraman, 1992b) and the perception that external suppliers have superior capabilities (Poppo and Zenger, 1998). Perhaps the most frequently cited reason for outsourcing however is the reduction of operational costs (Lacity and Willcocks, 1998), often through economies of scale – the production and distribution efficiencies which come with
larger size (Chandler, 1990; Dibbern et al, 2005). For traditional outsourcing benefits can accrue in circumstances where in-house production does not achieve the minimum efficient scale (Venkatesan, 1992). Collaborative sourcing introduces a demand side dimension to this traditionally supply side phenomenon. As a group of organisations aggregate their demand a potential supplier should become better placed to realise economies of scale in meeting it.

While outsourcing research has examined the design of outsourcing arrangements it has often done so from a process rather than a structural outlook (see Dibbern et al, 2005) – in particular emphasising the relative role played by contracts and relationships – and from a customer–supplier(s) perspective (Willcocks et al, 2007). It has not examined in any depth the possibility of, and structures for, customers acting collaboratively (ie customer-customer). As such it is necessary to turn to research on collaboration. There is a considerable body of research examining collaboration from the perspectives of motivation, intensity and extent (Powell et al, 1996; de Rond and Bouchikhi, 2004). A wide variety of motivations for collaboration have been proposed including scale economies (Larsson et al (1990), competence development (Hamel et al, 1989), the spreading of risk (Guglar and Dunning, 1993), increased resource utilisation (Clemons and Row, 1992), access to complementary resources (Guglar and Dunning, 1993) and customer service enhancement (Lambert and Knemeyer, 2004). In terms of intensity, Moss-Kanter (1994) suggests that cooperative arrangements range along a continuum from weak and distant to strong and close, while Lambert and Knemeyer (2004) define three types of partnership based on the level of integration between partners. Kumar and van Dissel (1996) and Thompson (1967) similarly differentiate between collaborations based on the degree of interdependency. When considering the intensity of collaboration a number of authors (Pisano, 1991, Dyer and Singh, 1998) have suggested that consideration needs to be paid to the governance structure that is put in place. Yoshino and Rangan (1995) differentiated between collaborations based on whether they involved an exchange of equity or not.

While the bulk of collaboration research has focused on partnerships between two organisations (Doz, 1996; Adober, 2006) there have been some studies of multi-organisation arrangements (Kumar and van Dissell, 1996). Authors such as Oye (1986) and Coleman (1990) have argued that collaboration becomes harder as the number of partners increases – not least because of the possibility of free riding. Others however suggest the opposite arguing that risks, for example, are better able to be dispersed as the number of partners increases (Adobor, 2006). Harrigan (1988) suggested that the key influence on the appropriate number of partners is the scope of the activities concerned. Not only can there be variation with regard to the number of activities included (see for example Mariti and Smiley, 1983) but also with regard to its centrality or contribution to an organisation. Weill and Broadbent (1998) suggested that from an IT perspective a separation can be made between infrastructure and applications. It is also well recognised that beyond the IT itself are the essential business processes of an organisation (see for example Davenport, 1993).

**Figure 1: Collaborative sourcing: Motivation and design**
Synthesising the outsourcing and collaboration literatures suggests that empirical research would be useful to determine whether economies of scale – a common driver for both outsourcing and collaboration – is indeed the motivation behind collaborative sourcing and identify what form that collaboration takes – in terms of its intensity, focus and extent (see Figure 1).

3 METHODOLOGY

Given that little research has been conducted to understand the phenomenon of collaborative sourcing a qualitative approach was determined to be appropriate (Benbasat et al., 1987; Strauss and Corbin, 1990). The research was primarily outcome rather than process oriented – seeking to identify the motivation for, and design of, collaborative sourcing initiatives rather than the process of making those decisions (Patton, 2002). A multi-case design was selected with collaborative sourcing arrangements selected from both the private and public sectors in Australia\(^1\). A multi-case approach, it was felt, would provide for more robust analytical generalization with regard both to the appropriateness of the motivation and design constructs and – in the case of the latter – the specific choices to be made when constructing a collaborative sourcing arrangement (Yin, 1984). For each arrangement considered interviews were conducted with more than one participant to capture any variation in perspectives. As Table 1 illustrates the provider perspective was included where possible. Interviews were between one and two hours in duration and a semi-structured interview protocol was followed. While the underlying rationale was purposeful it was deliberately non-directive so as not to preclude the emergence of factors and influences not previously considered (Patton, 2002). As such it is in line with the methodology presented by Eisenhardt (1989).

<table>
<thead>
<tr>
<th>Collaborative sourcing arrangement</th>
<th>Sector</th>
<th>Interviewees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>Public</td>
<td>Group General Manager Corporate Services (Partner A) [GGMS-A-1]; Director Finance and Administration (Partner B) [DFA-B-1]</td>
</tr>
<tr>
<td>Case 2</td>
<td>Private</td>
<td>General Manager (Partner A) [GM-A-2], Finance Director (Partner A) [FD-A-2]</td>
</tr>
<tr>
<td>Case 3</td>
<td>Private</td>
<td>Managing Director (Partner A) [MD-A-3]; Corporate Services Director (Partner B) [CSD-B-3]</td>
</tr>
<tr>
<td>Case 4</td>
<td>Private</td>
<td>General Manager (Partner A, B, C) [GM-A-4, GM-B-4, GM-C-4]; General Manager (Provider) [GM-P-4]</td>
</tr>
<tr>
<td>Case 5</td>
<td>Private</td>
<td>Managing Director (Partner A) [MD-A-5], General Manager (Partner B) [GM-B-5]; Managing Director (Provider) [MD-P-5]</td>
</tr>
<tr>
<td>Case 6</td>
<td>Private</td>
<td>Director Finance (Partner A) [DF-A-6]; Director Business Partnerships (Partner B) [DBP-B-6]; Director Business Enterprises (Provider) [DBE-P-6]</td>
</tr>
<tr>
<td>Case 7</td>
<td>Private / Public</td>
<td>Director Finance (Partner A) [DF-A-7]; General Manager Corporate Services (Partner B) [GMCS-B-7]; General Manager Australia (provider) [GMA-P-7]</td>
</tr>
<tr>
<td>Case 8</td>
<td>Private</td>
<td>General Manager (Partner A) [GM-A-8]; Finance Manager (Partner B) [FM-B-8]; Managing Director (Provider) [MD-P-8]</td>
</tr>
<tr>
<td>Case 9</td>
<td>Public</td>
<td>Director General (Partner A) [DG-A-9]; Deputy Director General (Partner B) [DDG-B-9]; Managing Director (Provider) [MD-P-9]; Chief Information Officer (Provider) [CIO-P-9]</td>
</tr>
</tbody>
</table>

Table 1: Interviewees per case

\(^1\) Australia is recognised as a country where outsourcing is prevalent (see for example Seddon, 2001 and Willcocks et al)
With regard to analysis, data was first reviewed and coded in terms of its relationship either to the motivation for, or design of, the collaborative sourcing initiative – an approach in accord with the recommendations of Miles and Huberman (1994).

4 RESULTS

As highlighted by Figure 2 the empirical research found that collaborative sourcing was more complex than anticipated – in that there is neither a single motivation nor a single collaborative design. Each aspect – motivation, intensity, focus and extent – will be elaborated on below.

<table>
<thead>
<tr>
<th>Motivation</th>
<th>Intensity</th>
<th>Focus</th>
<th>Extent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economies of scale</td>
<td>Access</td>
<td>Coat-tail</td>
<td>IT infrastructure</td>
</tr>
<tr>
<td></td>
<td>Flexibility</td>
<td>Aggregate demand</td>
<td>IT application</td>
</tr>
<tr>
<td></td>
<td>Management resources</td>
<td>Joint management</td>
<td>Business process</td>
</tr>
<tr>
<td></td>
<td>Expertise</td>
<td>Joint venture</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Revenue</td>
<td>Joint operation</td>
<td></td>
</tr>
<tr>
<td>Politics / Public relations</td>
<td></td>
<td></td>
<td>Number of partners</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Number of activities</td>
</tr>
</tbody>
</table>

Figure 2: Collaborative sourcing – motivations and designs

4.1 Motivation

While economies of scale was cited as a motivation in all cases it was typically moderated by other factors.

- Economies of scale
  
  “The main goal is economy of scale” GGMS-A-1

  “simply by consolidating the economies of scale, we provided, in a lot of cases, services for three agencies for the price of two. So we achieved – you know, we just spoke to software vendors and other vendors and said, look, if you want to play with us into the future, it’ll cost you a 30 per cent reduction in cost” CIO-P-9

- Access

In a number of cases a goal was to access technology used by larger organisations that could not be afforded individually.

  “we simply can not afford some of the technology on our own” MD-A-3

  “all have similar needs – about 90% fit – and can pretty much assume that if it satisfies the larger ones it will be sufficient to meet our needs” GM-C-4
• Flexibility

Flexibility was not so much a pure goal but seen as an influence on how far some organisations would go down the collaboration path.

“who you link up with limits what you can do in the future .. makes you a bit dependent”
GM-A-2

“you can only standardise in certain things and only so far” DFA-B1

• Management resources

For a number of organisations a motivation was either to free up their own management resources to focus on other activities or to allow the development of specialist management skills – such as contract or supplier management – that could not be justified individually.

“maintain skills sets that can not afford individually” DF-A-7

“don’t have the competencies .. nor did we want those competencies since it was not a critical function” GM-A-2

“time you need to manage that relationship, how much time and how much effort and how much resource do you have to manage those relationships.” DBP-B-6

• Expertise

Collaboration was see as providing access to expertise – typically developed through the past experiences of one of the partners – that would be otherwise unavailable and that could make a significant contribution to improving activities.

“The IP [intellectual property] that drives the business comes out of [supplier] – the delivery capability, the management experience, the technical expertise and so on” DF-A-7

“we saw the need for a vehicle that allowed [organisations] to work with somebody else who’s got those expertise who’s looked at these things before, can bring in the processes and common systems and common sort of programs” DBP-B-6

• Revenue

Collaboration could help generate a new revenue stream such that services could be offered for sale to additional organisations beyond the boundaries of the collaboration.

“turn it into a revenue stream rather than just trying to manage your costs and minimise your costs .. ” GMA-P-7

“.. if you’re making a couple of cents a transaction and you get a big customer and they’re doing, you know, 20 - 30 million transactions a year, that’s where you make your money”
GM-A-8

• Politics / Public relations

Collaboration is seen as more acceptable than outsourcing.

“cost savings were demanded but outsourcing was off the table”DG-A-9

“jointly owned intermediary makes it more palatable – retain control and not direct with the private sector” DBP-B-6
4.2 Intensity

The research found that there was not a single design of collaborative sourcing but various alternatives which varied with regard to the intensity of the collaboration undertaken.

- Coat-tail

Where an organisation rides along on the coat-tail of another taking whatever option or solution that organisation chooses benefiting from the economies of scale but without any input or opportunity for modification. It is questionable whether such an option truly represents collaboration though in one of the cases examined there was a requirement that the coat-tailing organisation invest in the vehicle established for the purpose.

“assumption is if it suits us – the largest – it will suit them and they can come along and get the savings but we are not going to compromise our requirements” GM-A-2

“other organisations but we have not really taken them into account in the requirements” GGMS-A-1

- Aggregate demand

Organisations aggregate their demand and requirements at the tender stage but enter into separate contracts and operations with the selected supplier.

“The host agreements are all separate … So this is just purchasing power.. What we’re trying to do is to get as much of the cost benefit without selling your soul. We negotiate together, but at the end of the day we are separate entities. We are separate businesses, with the same supplier. That works for us.” MD-A-3

- Joint management

Organisations set up a jointly owned intermediary that negotiates on their behalf with suppliers and then manages the relationship with the selected supplier who provides a common service to all of them.

“Effectively what we’ve done is we’ve put a middle man in there to look after our interests.. they manage the relationship with [the external provider]. It suits [the external provider] to deal with one entity and one contract” GM-B-4

- Joint venture

Organisations enter into a joint venture with a party from outside the industry who is perceived to have the skill sets necessary to improve the activities of interest.

“They have the skill sets, the reference sites, the experience to be able to make it happen .. we have the contacts to provide the opportunity” DBP-B-6

“They have done this before successfully .. we haven’t .. do not want to see it fall over” GMCS-B-7

- Joint operation

A group of organisations establish their own service provider to provide a common service to all of them.

“It became evident in our normal cooperative effort that we needed to be combining our computer bureaus…” MD-A-5
It should be recognised that the options described are somewhat of a *simplification*. Within each category of collaboration there were variations with respect to the structure of the collaborative vehicle – in terms of the presence, and significance, of the equity invested – and the broader governance arrangements. With the exception of one of the coat-tail arrangements all of the collaboration arrangements involved the establishment of separate legal entities, companies, in which the partners purchase equity. However the level of that investment varied from the company being nominal – more-or-less in name only – to it being significant in its own right.

“no resources attached to it .. really just a vehicle” MD-A-3

With regard to governance, the structure put in place influenced the rights and responsibilities – potentially limiting input of some participants.

“So you’ve got to think deliberately about that, how the governance is designed, who should be involved and how” DF-A-7

In addition that structure was often quite complex containing multiple, nested agreements.

“When you create a structure like that, you need a lot of things to make things pump around here .. so you’ll see these services, sub-contracts and partnership agreements, second services contracts, service agreements, resourcing agreement.” DBP-B-6

### 4.3 Focus

The cases suggested that the activities collaborative sourcing has been engaged in for vary with regard to how tightly they are tied to the specific operations of the organisations involved – with collaboration identified at the IT infrastructure, IT application and business process levels.

- **IT infrastructure**

  At perhaps the most basic level is largely generic hardware, operating software and services

  “All we do is put the architecture and infrastructure in place” MD-P-5

- **IT application**

  Here standard or industry based applications are provided and supported.

  “it is [the application] opaque, it’s chugging away, and it has no bearing on the business.”
  
  MD-A-5

- **Business process**

  At the business process level the collaborative sourcing extends to the actual operation of an activity such as payroll processing, accounts payable etc. In many cases those activities are transactional in nature – high volume, standard activities requiring limited decision making.

  “its transactional, its process-driven and we’re good on process.” GMCS-B-7

  “it’s about standardisation and it’s about being able to add value to a process. So unless you can standardise it, you can’t get cost savings.” GM-A-2

However in some cases the activities concerned extended to those that were more ad hoc, unique and requiring of analytical and decision making capabilities. The inclusion of such “expert” services was seen as a means of enabling the complete removal of functions from businesses.

“I don’t think you could have the one without the other [transaction and expert services] because the whole point .. is to take those functions away from the agencies. If you’re going
to separate them out .. you’re not going to get the savings, so you really need to group them together” FD-A-2

While there is normally thought to be a hierarchy between infrastructure, applications and processes it was not found to be the case that a higher level of collaboration required that lower levels were included – it was possible to have collaborative sourcing for business processes but not applications or infrastructure.

“they have retained ownership of the underlying ERP [Enterprise Resource Planning] platform and we essentially pay for some of the access to the technology” DBE-P-6

4.4 Extent

The research suggested that collaborative sourcing arrangements varied with regard to their extent both in terms of the number of partners involved and the number of activities covered. However there did appear to be a trade-off between the two with collaborations with a larger number of partners covering a narrower range of activities and vice versa.

• Number of partners

The number of partners varied considerably – from three to over 50.

“it’s the size of the agencies and the number of them. More than three, initially, would have been messier than it was” MD-P-9

“Because we’re a private company, I think it’s capped by you can only have 50 shareholders. So 50 shareholders. So we’ve just changed them and we’re going to be a public unlisted company, so that we can bring those other people in as shareholders.” MD-A-5

As alluded to earlier, there could also be differing levels of partner involvement – at the most basic level between those actively involved in driving and running the venture and those serving simply as customers.

“varies with regard to how involved they want to be .. I am involved a couple of days a month on the board .. others take much more of a backseat” GM-B-5

“now focused not on more partners but on more customers” FM-B-8

• Number of activities

A difficulty in comparing the number of activities provided is that it is difficult to ensure that apples are being compared with apples – names of activities and what is included vary across the cases. However it is clear that at the highest level there are considerable variations in terms of the range of activities with some arrangements being focused, for example, on a single application or process area and others having a much larger remit.

“in terms of transaction management, it doesn't matter if you're doing the payroll or paying bills or moving assets, the concept is make sure you've got low decision making .. basically, it's a formula, it's getting accuracy the first time, and so what we did is actually combined all the transaction areas and just called it transactional services” FD-A-2

“We're very niche in terms of the services that we provide. very focused .. know what we are good at and want to focus on” GM-P-4
5 DISCUSSION

While clearly collaborative sourcing is more complex than originally envisaged (there is neither a single motivation nor design), the research suggests that the framework proposed remains useful with the dimensions capturing variations and facilitating a classification of collaborations.

While in all cases economies of scale is a motivation it is not always the only one and often appears to be moderated by other factors. In terms of the focus of the collaboration there is considerable variation. Furthermore there does not seem to be a simple hierarchical progression – from infrastructure through applications to business processes. With extent however there does appear to be something of a trade-off between the number of partners and activities – or the breadth and depth of the arrangement.

Combining the different elements of the framework it is possible to discern three distinct approaches to collaborative sourcing:

1. So long as I do not have to change [Cases: 1, 2, 3]
   While such organisations join with others to realise economies of scale the emphasis remains primarily fixed at the individual level and ensuring individual requirements are not compromised. “Collaborations” are typically narrow in terms of the number of partners but may cover a variety of activities and are focused on the provision of “as is” services.

2. Better in this together [Cases: 4, 5]
   Organisations recognise the benefits of sharing management resources as well as those of economies of scale. The focus of collaborations is typically broad in terms of the number of partners but tightly focused and related to infrastructure. The emphasis remains on the provision of services “as is”.

3. Grab a new opportunity [Cases: 6, 7, 8, 9]
   Organisations see an opportunity not only for economies of scale but also to develop new revenue streams through the provision of services to others. The collaboration may relate to infrastructure, applications or processes – but is typically very specific. There is often a trade-off between the number of partners and the number of activities. The emphasis is on improving the current way of doing things – and if the organisations do not have the expertise amongst themselves to do this they will bring in an external partner.

While the different groups have been described from the perspective of making their own choices, the research suggests that this might not always be the case with politics or public relations determining the path taken. There is also some evidence of collaborations in Group 2 developing over time to also provide services to customers – but in contrast to Group 3 on a “take it or leave it” basis similar to Group 1.

It should be noted though that the framework and classification is a simplification of reality. Many of the dimensions have additional complexity embedded within them. For example regarding business processes they are not all vanilla – with interviewees making the distinction between transaction and expert types. With partners, there would also seem to be differing levels of involvement – for example there are those that take an active role, those that are more passive and those that are simply customers.

Future research would be particularly useful in three directions. Firstly to examine in more detail the variations and complexities of the collaborative sourcing arrangements. Secondly to attempt to determine which arrangements are most likely to be successful – and in what
circumstances. Third to examine the arrangements over time – do they evolve as partners gain in experience and understanding? Do organizations get locked in to particular solutions and partners with alternatives becoming closed off over time?

References

Blogging to Express Self and Social Identities, Any One?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0079.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Blogging, Virtual identity, Social computing, Theory Building</td>
</tr>
</tbody>
</table>
Abstract

Blogging has fast emerged as an unparalleled social computing phenomenon not to be ignored. This paper begins by reviewing the motivation factors for blogging identified in the extant literature. These factors can be broadly classified as either activities-based or behavioral-based. We then propose that a root motivation factor based on the theoretical underpinnings of self identity and social identity can better explain blogging motivation compared to the existing factors. Specifically, we posit that individuals could be receptive toward a migration of their real world self and social identities to the online virtual blog environment and thereby cumulates in the formation of blogger virtual identities. Blogger virtual identities is conceptualized as a second order formative construct formulated through two first-order reflective subconstructs, namely the virtual self identities and virtual social identities. Drawing on the theory of planned behavior, we propose a virtual identities model of blogging intention to investigate how the blogger virtual identities can lead to blogging intention and actual blogging behaviour. A two-stage research methodology consisting of a qualitative content analysis and followed by a survey is also proposed. We conclude with a brief discussion on the potential theoretical and practical contributions of this paper.

Keywords: Blogging, Virtual identity, Social computing, Theory Building.
1 INTRODUCTION

Many would have agreed that one of the most exciting online activities in recent years has been blogging, i.e., writing on individual’s web page with reverse chronological sequence of dated entries (Herring, Scheidt, Bonus & Wright 2004; Sifry 2007). Research on blogging has become increasingly important and popular with significant amount of effort expended on identifying various motivation factors leading to a person embraces blogging and examining the social network of blogger (Gumbrecht 2004; Li 2005; Nardi, Schiano & Gumbrecht 2004a; Nardi, Schiano, Gumbrecht & Swarth 2004b; Schiano, Nardi, Gumbrecht & Swarth 2004; Trammell, Tarkowski, Hofmokl & Sapp 2006). More recently, other researchers have also investigated the motivation factors underlying microblogging, a relatively new derivation of blogging that allows bloggers to write short entries while on the move (see Java, Song, Finin & Tseng 2007). However, we believe that the extant literature does not adequately explain the root motivation factor of blogging (Nardi et al. 2004b).

This research in progress paper attempts to use the self and social identity notions as theoretical lens to address the knowledge gap. The rationale is that personal web pages have been known to be used by individuals to create their online identities as a mean of self presentation (Schau & Gilly 2003). Indeed, information provided on personal web pages pertaining to how the author would like to be regarded by others, i.e. identity claims, can portray the author’s personality accurately (Vazire & Gosling 2004). In other words, the need to project certain impression of oneself to others in the online world could plausibly explain why an individual could be motivated to blog (Schlenker 2005). To this extent, we contrive the theoretical notion of blogger virtual identities, which encompasses the online virtual self and social identities of an individual blogger, as the root motivation factor of blogging. Drawing on the Theory of Planned Behavior (Ajzen 1985; Ajzen 1991), we further propose a preliminary research model termed as the Virtual Identities Model of Blogging Intention to explain how blogger virtual identities can affect blogging intention and actual blogging behavior.

Understanding the root motivation factor of blogging from the identity perspective is important for several reasons. In particular, when individuals present themselves to others in the online world, it could either be a deception attempt or sincere portrayal of one’s self, regardless of whether it is done intentionally or unconsciously (Schlenker 2005). Our present study has the potential to demonstrate that bloggers are generally inclined towards a sincere portrayal of themselves to others. Furthermore, blogging is not only a popular online activity but has increasingly been incorporated as a pertinent feature of online social networking sites such as Friendster, Facebook and MySpace. The identity-based motivation factor can not only explain the use of blog but also be extended to examine the use of social networking sites beyond making friends to expressing one’s identity. In addition, the identity notion provides a theoretical basis for enhancing the functional design of blog tools to increase adoption and usage.

2 THEORETICAL BACKGROUND

2.1 Motivation Factors of Blogging

There are two interrelated fields of thoughts pertaining to the underlying motivation factors for a person to blog. The first field of thoughts posits that blogging arises as a form of social activity whereby bloggers create postings with their readers in mind and readers respond to the postings with

anticipation for more correspondents (Nardi et al. 2004a). This view is consistent with the notion of having a blog as a medium of communication between a blogger and the intended audiences (Gumbrecht 2004). In addition to the purpose of communication, five other object-oriented activities identified as blogging motivations are: 1) to maintain online personal journal; 2) to convey and convince others about one’s opinions; 3) to express emotions; 4) to develop one’s thinking process through blogging; and 5) to build an online community (Nardi et al. 2004a; Nardi et al. 2004b; Schiano et al. 2004). Similarly, from the microblogging perspective, people blog to talk about daily routine, converse with other bloggers, share information and report news (Java et al. 2007). Another field of thoughts anchors on the uses and gratifications (i.e., behavioral) perspective. This field of thoughts suggests that self-documentation of key life events, cultivating good writing habit and style, articulating one’s personality, acquiring information, passing time, socializing with other people, seeking personal entertainment and achieving professional advancement are the underlying motivation factors to blog (Li 2005; Trammell et al. 2006).

In our view, the two perspectives of blogging, i.e. activities and behavioral, could well complement each other to offer a more holistic perspective on a blogger’s motivation to blog. For instances, the desire to express one’s personality or life, i.e. self documentation, could account for the use of blog as personal journal and both opinions and emotions outlets. Additionally, the socialization perspective could shed insights as to why people use blogs as a communication tool and build community. However, the uses and gratifications framework primarily focuses on the motives for media use (Papacharissi 2002). It does not provide sufficient fundamental explanation for the causes and effects of categories of motivation that are chosen a priori for empirical investigation (Li 2005; Papacharissi 2002; Trammell et al. 2006). Thus, while we know that self expression and socialization have been repeatedly identified as important motivations for blogging (Li 2005; Trammell et al. 2006), the existing knowledge gap impede a fuller understanding of why bloggers would want to tell readers about themselves or develop an affinity for part of a community. Examining the two perspectives deeper, we suggest that both perspectives share a common underlying root of motivation: identity formation. We conceive that identity, in the forms of self identity and social identity, could synergize the existing two perspectives and serve as the theoretical lens to explain for why bloggers develop a particular behavior (e.g., expressing political viewpoints) or participate in a particular blogging activity (e.g., writing journal). We reason that the notion of virtual identity presents a concise and strong theoretical underpinning governing the examination of the motivation factors to use social computing applications, such as blog. These include social networking site and virtualization world that both emphasize the user’s online identity and the associated relationship with other people.

2.2 Self Identity and Social Identity

Identity of an individual could come in two forms: self identity and social identity (Hogg, Terry & White 1995; Lee, Lee & Lee 2001; Stets and Burke 2000; Thoits & Virshup 1997). The notion of self identity builds on the role identity theory (McCall & Simmons 1978) and the identity theory (Stryker 1980) to denote that one’s identity can be perceived as the expressive behavior of an individual’s role in society (Conner & Armitage 1998; Lee et al. 2001). For instance, an individual could have the identity of a doctor, a managing director, a housewife, an undergraduate student depending on his/her role played in the society. It is important to note that self identity refers to a collection of socially accepted behaviors constructed from the role in which an individual occupies in society. For instance, a doctor (role) who is engaging in a medical treatment of a patient (behavior) would be perceived to have the identity of a medical doctor. Given that there could exist a reciprocal link between an individual (i.e., self) and the society, an individual’s identity is likely to affect his/her social behavior (Hogg et al. 1995). Consequently, an individual is very likely to adjust his/her behavior in a manner that is in accordance with one’s existing/goal identity (Gecas & Burke 1995).

Social identity refers to the characteristics of the social category in which an individual perceives to belong to. According to the social identity theory, the acquired social identity prescribes one’s
attributes as a member of that group, i.e., what one should think and feel, and how one should behave. In other words, an individual regulates his/her behavior in a manner that is in compliance with the goal or image of a whole group or category of people with which the individual belongs to (Hogg & Abrams 1988). The evaluative nature of the social identity provides a strong motivation for groups and their members to adopt strategies for achieving or maintaining in-group/out-group comparisons that favors the in-group and individuals. Two sociocognitive processes may be used to achieve this objective, namely categorization and self-enhancement (see Hogg et al. 1995). The former is concerned with producing stereotypical and normative actions and perceptions that are unique to a particular group thus differentiating it from other categories. The later suggests that people tend to view members within the same group or category more favorably along the stereotypical dimensions compared to external individuals.

The two theoretical perspectives differ in one critical aspect: the self identity perspective describes individual’s role-related behaviors; comparatively, the social identity perspective focuses on group process and intergroup relations (Hogg et al. 1995). In terms of similarities, both perspectives of identities (i.e., self identity and social identity) demonstrate that the structure and function of the self is constructed socially. In other words, individual behavior is affected by the social structure and that this relationship is mediated by a multi-faceted and dynamic self (Hogg et al. 1995). Furthermore, both perspectives predict that the resulting identity mediates the relationship between society and individual social behavior (Hogg et al. 1995). Taking a holistic view of the two identity perspectives, people are organically linked to their groups through social identities and mechanically linked within groups through their role identities (Durkheim 1983). Consequently, it is necessary to consider both self identity and social identity in tandem so as to yield a holistic view and individual’s identity. Indeed, an individual’s membership in any social group or assumed role includes two important aspects, namely one’s identification with a category and the behaviors that he/she associates with the category; both of which are emphasized by depersonalization in the social identity theory and self-verification in the identity theory respectively (Stets & Burke 2000). Hence, individual could simultaneously assume a role and be affiliated to a group to influence one’s perception, affection and behavior (Stets & Burke 2000). This concept is best summarized by Thoits and Virshup’s (1997) analogy on group behavior. The authors observed that decision making committee could differentiate activities among members. Thus, it is possible that people may be simultaneously aware of themselves occupying a particular role identity, e.g. committee members occupying different positions, and as group members working for a collective identity, i.e. the committee as a whole.

Whether behavior driven by self identity or social identity is exhibited depends on both the fit of the identity to the situation, as posited by the social identity theory, as well as the structural embeddedness or commitment of the individual, as posited by the self identity theory (Stets & Burke 2000). An individual who identifies with a group need not always behave in accordance with the social identity (Van Knippenberg 2000). The same argument could also apply to self identity. Thus, understanding the antecedent of identity salience is critical to comprehend the behavior of individual. It has been empirically verified that the salience of an individual’s self identity is positively affected by the degree of commitment (vis-à-vis the number and strength of important social relationships associated with the particular identity) and evaluation (Hoelter 1983). If an individual is strongly committed to a role and evaluates his/her performance in that role positively, then he/she is likely to exhibit the behavior associated with that role. Additionally, role contingencies and identity prominence are also predictors of identity salience (Nuttbrook & Freudiger 1991). In the case of social identity salience, a “fit” between social categorization and subjectively relevant features of reality is likely to lead to social identity salience (see Haslam, Powell & Turner 2000). This “fit” consists of both “comparative fit” and “normative fit”. In addition, social identification itself also plays a role in social identity salience through the manifestation of perceiver readiness (Haslam et al. 2000).

It is also important to note that an individual may possess multiple self identities and multiple social identities since people may have different roles in society and be affiliated with different social groups (Hogg et al. 1995). However, it is likely that people may only exhibit behavior that is consistent with a
single identity at any one time depending on which identity is more relevant. This reasoning is consistent with the identity salience notion (Haslam et al. 2000; Hoelter 1983). With an understanding of the notion of identity, we apply the theoretical underpinning of identity to the blogging context.

2.3 Blogger Virtual Identities

Researchers have found that bloggers, in particular teenagers and women, are inclined towards expressing their identity on their blogs (Van Doorn, Van Zoonen & Wyatt 2007; Huffaker & Calvert 2005). Beyond personal information such as age and location (Huffaker & Calvert 2005), bloggers also express their gender identity through: 1) the use of emoticons and language styles (e.g., males are more direct and forceful whereas females are more indirect and intimate) (Huffaker & Calvert 2005); and 2) “representational domains” (e.g., males prefer occupation domain whereas females prefer domestic interests or hobbies) (Van Doorn et al. 2007, pp. 151). Clearly, it is possible for bloggers to develop some form of virtual self identities through blogs. New communication technologies have also been observed to exercise an impact on the formation and development of online communities and the resulting virtual social identity (Cerulo 1997). Specifically, it has been suggested that social identity can be developed in both virtual and real world (Bergami & Bagozzi 2000; Dholakia, Bagozzi & Pearo 2004). Consistent with the general rationale of the social identity theory, socialized bloggers will develop virtual social identities as the commitment to their blog communities, i.e., consisting of the bloggers and readers, increases gradually (Moon, Li & Sanders 2006).

Building on the virtual identities concept, we posit that individual could be receptive toward a migration of their real world identities to the online virtual environment and thereby cumulates in the formation of blogger virtual identities. In other words, individual possesses various role and group identities in their real life and are likely to extend these identities into the online world as virtual self and social identities manifested through their blogs. For instances, a college student might blog about his/her college life from the perspective of his/her role identity as a college student as well as a social identity acquired through his/her interaction with peers and teachers in the college environment. At the same time, the same student may wish to use the blog to express his/her opinion of current affairs or political issues as emanated from his/her self identity as an avid current affairs observer. Likewise, perhaps he/she routinely watches soccer matches on television with his/her friends and may have an urge to blog about the proceedings and outcomes of the matches. In the last example, the student brings the in-group behavior online to form a virtual social network based on soccer matches discussion. Collectively, we term such identities as virtual self and social identities in which the virtual identities resemble the real world offline identities.

Similar to most types of computer mediated communication medium, blog can provide bloggers with anonymity instead of using their real names (Huffaker & Calvert 2005). To this extent, bloggers may be inclined towards expressing covert self or social identities that they do not openly exhibit in their real life; but which they seek certain outlet such as blog to express these identities. We illustrate this view with examples of covert social and self identities. For instance, a computer programmer acquires a role identity of an expert programmer by virtue of his/her structural job position in society. Building on this role identity, he/she develops a covert virtual social identity with the predisposition to help other programmers in a blog community of programmers. As an illustration of a covert self identity, the same computer programmer could also be a computer hacker who frequents hacking blog community and discusses about hacking on his/her own blog. Regardless, it should be clear that the virtual social identity of expert programmer in a community and virtual self identity of computer hacker both arise from the real world offline identities possessed by the computer programmer.

In gist, we contrive blogger virtual identities as a second order formative emergent construct with each of the two separate virtual identities, namely virtual self identities and virtual social identities, as reflective first order factors. This qualifies our blogger virtual identities construct as a Type II second-order factor in Jarvis, Mackenzie & Podsakoff (2003) specifications. Our present construct exhibits several formative indicators properties suggested by Diamantopoulos and Winklhofer (2001). For
instances, omitting either of the two formative indicators would not completely represent the virtual identities of a blogger while the magnitude of the correlations among the two constituent identities cannot be precisely predicted given that an individual may possess none, one or both and the identities may or not may be related.

With the above definition of blogger virtual identities in perspective, we proceed to discuss its salience. Briefly, we anchor on the same salience notions of the offline world and propose that the salience of virtual identities is dependent on the same set of factors (Hoelter 1983; Nuttbrock & Freudiger 1991; Haslam et al. 2000) except that they apply to the likelihood of extending their real world identities online in the manifestation of blog. For instance, a positive commitment toward extending a role identity online and its positive evaluation should result in virtual self identity salience.

In summary, our present conceptualization of the blogger virtual identities construct is different from the notions of virtual identity proposed in the extant literature. For instance, Moon et al. (2006) anchored their virtual social identity construct upon the three aspects of social identity, namely cognitive, affective and evaluative. Other researchers such as Gross and Acquisti (2005), and Stutzman (2006) have examined the virtual identity of online social network communities’ members using demographic and personal information, including images. However, our present endeavor emphasizes the expression of the blogger real world offline identities to the online blog environment as reflected in the blog contents. To this extent, the blogger virtual identities construct is a more holistic one that both extends and complements the self-identity expressiveness advocated by Thorbjørnsen, Pedersen & Nysveen (2007) as well as identity claims proposed by Vazire and Gosling (2004).

3 RESEARCH MODEL AND HYPOTHESES

We postulate that the blogger virtual identities construct could account for the various “activities-based” and “behavioral-based” blogging motivations identified by prior researchers (Gumbrecht 2004; Herring et al. 2004; Java et al. 2007; Li 2005; Nardi et al. 2004a; Nardi et al. 2004b; Schiano et al., 2004; Trammell et al. 2006) through the explication of the self and social identities. For instance, a college student’s positive attitude toward explicating his/her self identities as a college student and current affairs observer should lead to blogging behaviors such as self-documentation and self expression respectively. These behaviors should then activate blogging activities such as writing online journal and expression self-opinion. Additionally, his/her social identity as part of an expert programmers group induces socialization behavior that leads to online community building activity. To this extent, blogger virtual identity presents a plausible theoretical predictor for blogging intention.

Figure 1 The research model – virtual identities model of blogging intention.

Our research model, termed as the Virtual Identities Model of Blogging Intention, is depicted in Figure 1. In the broadest sense, it is based on the Theory of Planned Behavior (TPB), which posits that
individual behavior is driven by behavioral intentions (Ajzen 1985; Ajzen 1991). Behavioral intentions themselves are affected by the individual’s attitude toward the behaviour. Thus, we posit that behavioral intention to blog is determined by the individual’s attitude toward their blogger virtual identities. In turn, behavioral intention should lead to actual blog usage thus explaining the various “activities-based” and “behavioral-based” motivations. Although blogging may not always be a planned and rational behavior, the identity salience notion (Hoelter 1983; Nuttbrock & Freudiger 1991) does suggest that an individual regularly exhibits behavior corresponding to a particular self or social identity that is relevant to them. We reason that over a period of time, it is plausible for the blogger to consistently exhibit the same relevant virtual identities in his/her blog.

Prior researchers have examined the effects of self identity and social identity on behavioral intention using TPB as the theoretical lens. Self identity was found to be a significant predictor of behavioral intention, i.e. when the behavioral role constitutes an important part of one’s self identity, an individual could be motivated to engage in the behavior (Sparks & Shepherd 1992; Terry, Hogg & White 1999). In addition, group norm is observed to have a significant effect on behavioral intention for participants with high group identification while a strong correlation exists between self identity and group norm among participants who identified strongly with the group (Terry et al. 1999). More closely relevant to our context, an individual with high self-identity expressiveness and social identity expressiveness could express a high intention to blog (Thorbjørnsen et al. 2007). In another study conducted by Dholakia et al. (2004) to examine the social influence model of consumers in virtual communities, the authors found significant direct relationship between group norms and we-intention to engage in joint behavioral action, such as a blogger posts blog entries and readers respond. This could suggest that virtual social identity might also serve as a mediator to explain the effect on blog usage on future intention of bloggers. Specifically, as observed by Moon et al. (2006), social interaction through blog and its enjoyment could lead to the development of strong virtual social identity while the resulting identity is expected to lead to a strong intention to expand the participants’ online social network and increase satisfaction with their life. The later finding provides partial support to our conjecture of a direct relationship between blogger virtual identities and intention to blog. Overall, these findings suggest the presence of a causal link between blogger virtual identities and intention to blog. Accordingly, we hypothesize a direct causality between blogger virtual identities and behavioral intention to blog:

**H1**: The blogger virtual identities construct is positively related to behavioral intention to blog.

(a) The virtual self identities subconstruct is positively related to behavioral intention to blog.

(b) The virtual social identities subconstruct is positively related to behavioral intention to blog.

According to the Theory of Reasoned Action (TRA) (Ajzen & Fishbein 1980; Fishbein & Ajzen 1975) and TPB, attitude serves as an important antecedent of behavioral intention and has an indirect effect on actual behavior mediated by intention. In a similar vein, we posit that blogger virtual identities can also function as an attitudinal factor exercising a separate and indirect effect on intention. Prior study has found that stronger social identity leads to stronger desires to participate in virtual community (Dholakia et al. at al. 2004). In other words, the relationship between social identity and desire was found to mediate a separate significant and indirect relationship between group norms and we-intentions. To the extent that desire has been defined as the motivations to decide in favor of acting as part of a virtual community (Dholakia al. et al. 2004) and that desire functions as a mediator between attitudes, anticipated emotions and subjective norms with intention (Bagozzi & Dholakia 2002), we believe desire and attitude are similar. Support for this line of reasoning can be found in their similar operationalization in prior studies. For instances, one can compare the operationalization of desire in Dholakia et al. (2004, pp. 252, Table 2) with the operationalization of attitude in Terry et al. (1999, pp. 232). Furthermore, self-identify expressiveness has been found to have an indirect effect on intention through attitude (Thorbjørnsen et al. 2007). Accordingly, we present the second set of hypotheses between blogger virtual identities and attitude toward blogging:

**H2**: The blogger virtual identities construct is positively related to attitude toward blogging.
The virtual self identities subconstruct is positively related to attitude toward blogging.

The virtual social identities subconstruct is positively related to attitude toward blogging.

The extent literature consists of numerous empirical studies that validated the relationship between attitude and behavioral intention (Sparks & Shepherd 1992; Terry et al. 1999). Accordingly, we hypothesize that:

**H3**: Attitude toward blogging is positively related to behavioral intention to blog.

According to TPB, whether behavioral intention leads to actual behavior is conditional upon individual’s control over the behavior (Ajzen 1985; Ajzen 1991). Along this line of reasoning, we posit that an individual’s behavioral intention to blog is also conditional upon his/her perceived ability and extent to blog at will. In particular, perceived behavioral control has been found to be a significant distinctive predictor of behavioral intention (Terry et al. 1999). Thus we further hypothesize that:

**H4**: Perceived behavioral control is positively related to behavioral intention to blog.

Both TRA (Fishbein & Ajzen 1975; Ajzen & Fishbein 1980) and TPB (Ajzen 1985; Ajzen 1991) posit that behavioral intention is a key predictor of usage behavior. Within the technology acceptance literature, this has also been a well researched area with strong empirical supports (see Venkatesh, Morris, Davis & Davis 2003 for a comprehensive review). In particular, an expanded version of the Technology Acceptance Model which incorporates two crucial social factors examined in our present study, namely self identity and social identity, has presumed the relationship between intention and behavior (Lee et al. 2001). Accordingly, we hypothesize that:

**H5**: Behavioral intention to blog is positively related to actual blogging behavior.

Finally, it is important to consider the effect of actual blog usage behavior on blogger virtual identities. Prior study has established that blog usage, i.e. social interaction though the blog and enjoyment of blog, could lead to the development of a strong virtual social identity (Moon et al. 2006). In particular, virtual social identity develops from the trust of and commitment to the blog community when there are intense social interactions within the blog community (Ridings, Gefen & Arinze 2002). Moreover, as the bloggers express themselves and are viewed by others as such over an extended period of time, their blogger virtual identities could further crystallize. This is consistent with the internalization of the bloggers’ self-presentation whereby the bloggers’ behaviors to create certain impression of themselves to others could ultimately influence how they viewed themselves (Schlenker, Britt & Pennington 1996). Accordingly, we present our final hypotheses to complete our proposed research model:

**H6**: Actual blogging behavior is positively related to the blogger virtual identities construct.

**a** Actual blogging behavior is positively related to the virtual self identities subconstruct.

**b** Actual blogging behavior is positively related to the virtual social identities subconstruct.

### 4 PROPOSED RESEARCH METHODOLOGY

We propose a two-stage methodology to investigate the blogger virtual identities phenomenon. In the exploratory first stage, we propose a qualitative content analysis similar to the technique used by Van Doorn et al. (2007). The qualitative data to be collected will be subjected to an interpretive analysis to understand how and why bloggers express their real world offline identities to their online virtual identities. The details of the process are outlined as follow. We begin by selecting 500 blogs at random using local blog directories, and then codify both the bloggers’ demographic and information about the blog to establish a profile for a typical blogger. Key demographic variables to consider include age, gender, occupation and ethnic group while key blog variables include age of the blog, update frequency, update recency and language (Herring et al. 2004). We then review these 500 blogs and select 6 typical blogs and 4 atypical blogs for the actual content analysis. The later will involve
examining the blogger’s profile and all postings, both textual and non-textual contents such as images, made in the last 6 months.

The results from the first stage study should enlighten us on the viability of the blogger virtual identities construct. If the findings are consistent with our conceptualization, we may then proceed to the second stage to empirically validate our proposed research model shown in Figure 1. The findings from the second stage study will reveal insights as to whether expressing blogger virtual identities does indeed leads to blogging intention and actual blogging behaviour. A survey research methodology will be used for the second stage study. The target population will consist of two groups: 1) existing bloggers recruited through local blog directories; and 2) internet users who have yet to start a blog to be recruited through local lifestyle-related internet forums. The objective of such a design is to investigate any systematic difference in behavioral intention to blog between existing bloggers and potential new bloggers. This also allows for better generalizability within each sample. To ensure that the non-bloggers are able to understand the terms used in the survey, participants from both samples will be given a simple factsheet on blog to read prior to answering the survey.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual Self Identities</td>
<td>1. I am willing to blog about characteristics that are an important part of who I am.</td>
</tr>
<tr>
<td>[Adapted from Callero 1985; Lee et al. 2001; Thorbjørnsen et al. 2007]</td>
<td>2. I am likely to blog about characteristics that represent who I want to be.</td>
</tr>
<tr>
<td></td>
<td>3. I find that blogging about personalities that truly reflect who am I is very meaningful.</td>
</tr>
<tr>
<td></td>
<td>4. I am happy to blog about life experiences related to my work responsibility.</td>
</tr>
<tr>
<td></td>
<td>5. Blogging about events related to my job position is something I will do enthusiastically.</td>
</tr>
<tr>
<td>Virtual Social Identities</td>
<td>1. I am willing to blog about characteristics that are an important part of my social groups.</td>
</tr>
<tr>
<td>[Adapted from Lee et al. 2001; Mael &amp; Ashforth 1992]</td>
<td>2. I am likely to blog about characteristics that are representative of people who I associate with.</td>
</tr>
<tr>
<td></td>
<td>3. I find it meaningful to blog about personalities that truly reflect my social groups.</td>
</tr>
<tr>
<td></td>
<td>4. I am happy to blog about experiences related to others in my social groups.</td>
</tr>
<tr>
<td></td>
<td>5. Blogging about events related to my social groups is something I will do enthusiastically.</td>
</tr>
<tr>
<td>Attitude Toward Blogging</td>
<td>1. Blogging to reflect my online character would be: Unpleasant – Pleasant</td>
</tr>
<tr>
<td>[Adapted from Ajzen &amp; Fishbein 1980; George 2004; Terry et al. 1999]</td>
<td>2. Blogging to reflect my virtual personality is a: Bad idea – Good idea</td>
</tr>
<tr>
<td></td>
<td>3. Creating a blog to display my online character would be: Unfavorable – Favorable</td>
</tr>
<tr>
<td></td>
<td>4. Creating a blog to display my virtual personality is a: Foolish idea – Wise idea</td>
</tr>
<tr>
<td></td>
<td>5. Starting a blog to exhibit my character on the internet is an idea I: Dislike – Like</td>
</tr>
<tr>
<td>Perceived Behavioral Control</td>
<td>1. If I wanted to, it would be easy for me to engage in blogging.</td>
</tr>
<tr>
<td>[Adapted from Ajzen &amp; Madden 1986; George 2004; Terry et al. 1999]</td>
<td>2. How much control do you have over whether you engage in blogging?</td>
</tr>
<tr>
<td></td>
<td>3. I am capable of blogging on the internet.</td>
</tr>
<tr>
<td></td>
<td>4. Blogging on the internet is entirely within my control.</td>
</tr>
<tr>
<td></td>
<td>5. I have the resources, knowledge and ability to blog on the internet.</td>
</tr>
<tr>
<td>Behavioral Intention to Blog</td>
<td>1. I intend to engage in blogging within the next 12 months: Extremely unlikely – Extremely likely.</td>
</tr>
<tr>
<td>[Adapted from Terry et al. 1999]</td>
<td>2. I intend to blog on the internet within the next 12 months: Extremely unlikely – Extremely likely.</td>
</tr>
<tr>
<td>(7-point Likert Like)</td>
<td>3. I am going to blog within the next 12 months: Extremely unlikely – Extremely likely.</td>
</tr>
<tr>
<td>Actual Blogging Behavior</td>
<td>1. I update my blog: Not At All (Stopped blogging), Once in a while (less frequent than monthly), Monthly, Weekly, Daily</td>
</tr>
<tr>
<td>[Self-developed]</td>
<td>2. On average, the number of posting(s) that I made on my blog in a week is: 0, 1-3, 4-6, 7-9, 10 or more.</td>
</tr>
<tr>
<td>(5-point scale)</td>
<td>3. On average, the number of lines of text per posting that I made on my blog is approximately: 1-5, 6-10, 11-15, 16-20, More than 20.</td>
</tr>
</tbody>
</table>

Table 1 Instrument scales used to measure constructs.

Except for the blogger virtual identities and actual blogging behavior constructs, the remaining three constructs in our research model are operationalized based on validated instrument scales from the extant literature. The instrument scales for all the constructs are listed in Table 1. Unless otherwise stated, the scale type is 7-point Likert (1 – Strongly Disagree to 7 – Strongly Agree). The instrument scales will be validated through a formal procedure as described in Moore and Benbasat (1991). Emphasis will be placed on the instrument testing phase which includes pilot and field tests.
Specifically, the pilot tests will be used to assess the construct validity of blogger virtual identities using a second-order factor analysis in LISREL.

5 CONCLUSION

A substantial amount of research efforts are required to implement the proposed two-stage methodology. However, there are several potential contributions that can be realized in addition to the noteworthy achievement of a better understanding of the blogging phenomenon. Theoretically, the blogger virtual identities construct may be extended to examine other social computing applications, tools or phenomenon that emphasize the user’s online identity and the associated relationship with other people. For instance, online social networks such as Friendster, Facebook, MySpace and virtual world such as Second Life2. Practically, this paper provides a theoretical basis for refining the current design layouts and elements of blog to explicate or facilitate the development of both self and social identity in a blog environment. Specifically, design principles or elements that improve explicating of an individual’s virtual identities should lead to greater use of blog and potentially other social computing applications.

In conclusion, this paper summarizes the activities-based and behavioral-based blogging motivation factors found in the extant literature. We further highlight the inadequacies of these factors in providing a good understanding of why people would want to blog. Accordingly, we put forth the notion of blogger virtual identities that is based on the theoretical underpinnings of self and social identity. We reason that the propensity for an individual to extend both his/her real world self and social identities to the online blog environment cumulating in the formation of blogger virtual identities can explain why a person would want to blog. More specifically, we put forth the Virtual Identities Model of Blogging Intention, which is based on TPB, for investigating the effect of blogger virtual identities on blogging intention and actual blogging behavior.

References


UNDERSTANDING VIRTUAL WORLD USAGE: A MULTIPURPOSE MODEL AND EMPIRICAL TESTING

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0201.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Virtual world, Adoption, Consumer behavior / choice / demand / empowerment / reviews / consumerism, Technology Acceptance Model (TAM)</td>
</tr>
</tbody>
</table>
UNDERSTANDING VIRTUAL WORLD USAGE: A MULTIPURPOSE MODEL AND EMPIRICAL TESTING

Tibert Verhagen
Knowledge, Information and Networks research group, Faculty of Economics and Business Administration, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Tel. +31.20.4446050, Fax: +31.20.4446005, E-mail: tverhagen@feweb.vu.nl

Frans Feldberg
Knowledge, Information and Networks research group, Faculty of Economics and Business Administration, VU University Amsterdam.

Bart van den Hooff
Knowledge, Information and Networks research group, Faculty of Economics and Business Administration, VU University Amsterdam.

Selmar Meents
Knowledge, Information and Networks research group, Faculty of Economics and Business Administration, VU University Amsterdam.

Abstract

This study reports an attempt to enhance our understanding of the reasons behind virtual world usage. By providing a mixture of utilitarian and hedonic value, virtual worlds represent an emerging class of multipurpose information systems (MPIS). Previous research seems to fall short in explaining MPIS adoption, especially since key extrinsic and intrinsic motivators are not considered. Drawing upon IT adoption research, motivation theory and the consumer behavior literature, this research extends available works and provides insight into the influence and roles of extrinsic and intrinsic motivation. Hypotheses are postulated and tested using a sample of 1627 users of the virtual world Second Life. The results confirm the majority of the hypotheses and support the comprehensive model. The findings indicate instantaneous effects of extrinsic and intrinsic motivation, and highlight reinforcing effects of intrinsic motivation. Implications for research and practice are discussed.

Keywords: virtual worlds, multipurpose information system, extrinsic motivation, intrinsic motivation, adoption behavior, escapism, visual attractiveness, entertainment value.
INTRODUCTION

With increasing usage, the value of virtual worlds has been growing at an impressive rate. Parallel to their growth and diversified usage, virtual world systems have gradually advanced into more complex systems. Initially developed to serve gaming communities, virtual worlds were primarily designed in line with the intrinsic motives of those intending to use the system, that is, to create pleasurable experiences and provide self-fulfilling value (Ryan and Deci, 2000b). As such, these virtual worlds could be classified as hedonic information systems (cf. Van der Heijden, 2004). More recently, virtual worlds have started to fulfill a more instrumental role by enabling their users to perform rather utilitarian tasks such as getting insurances, following courses, and building and leasing stores. Accordingly, virtual worlds increasingly adapt to those willing to use the system for extrinsic motives, which relates to the achievement of goals/benefits external to the system-user interaction (Van der Heijden, 2004). From this perspective, virtual worlds contain characteristics that are distinctive for utilitarian information systems (see Van der Heijden, 2004). Given the mixture of hedonic and utilitarian functions, it becomes clear that the traditional distinction between these functions no longer applies to information systems like virtual worlds and thus demands revision. Mirroring the view of Hong and Tam (2006) on IT beyond work settings, we view virtual worlds as multipurpose information systems (MPIS). MPIS are defined here as information systems that provide an integrated suit of hedonic and utilitarian functions to their users.

In this paper we embrace the MPIS perspective, and introduce and test a comprehensive model of users’ motivations to adopt virtual worlds. Given the multipurpose characteristics of virtual worlds, using an MPIS perspective encompassing both intrinsic and extrinsic motives seems crucial. The vast majority of research on IT system adoption, however, has focused on either hedonic (e.g., Van der Heijden, 2004) or utilitarian systems (e.g., Venkatesh, 2000). Consequently, these works fall short since they address only about half of a system’s nature and therefore fail to provide an overall picture of the intrinsic and extrinsic motivators behind virtual world use. Furthermore, since the interrelationships between both types of motivators are relatively unexplored, possible synergies between hedonic and utilitarian system characteristics that might guide system developers remain untainted. This research intends to further our understanding of virtual worlds by introducing an MPIS model that captures both intrinsic and extrinsic believes and as such integrates the hedonic and utilitarian system paradigms.

The paper is structured as follows. We first discuss the literature on the nature and value of information systems, and address the relationships between information system value and consumers’ IT adoption motivations. We conceptualize the key constructs in our model and formulate hypotheses. The research model is then tested in a large-scale survey within the virtual world Second Life. We will discuss the results of this empirical study, and the conclusions that can be drawn from them. Finally, we will describe the theoretical, methodological and practical implications of our findings.

THEORETICAL BACKGROUND

2.1 Information system value

Research on the adoption of IT has classified information systems according to the value they provide to users. Reflecting the user’s overall experience of interacting with the system based on both the system’s usefulness and provision of enjoyment and/or fun (adapted from Holbrook, 1986; Babin, Darden and Griffin, 1994), system value is assumed to be an important determinant of adoption behavior. A common value-based classification originated in the consumer behavior literature (e.g., Babin et al., 1994; Hirschman and Holbrook, 1982, Holbrook and Hirschman, 1982a), and is used to differentiate between utilitarian and hedonic information systems (cf. Van der Heijden, 2004).
Utilitarian information systems refer to systems that are mainly used to achieve goals that are external to system usage (Van der Heijden, 2004). As such, they provide instrumental utility to the user. Hedonic information systems on the other hand, are systems that are used for their own sake and provide feelings of enjoyment, pleasure, excitement and escapism (cf. Babin et al., 1994). The usage of hedonic systems is mainly experiential in nature, which implies that users strive for instant hedonic pleasure while the consequences of usage appear in the fun and pleasure of the experience itself (Holbrook and Hirschman, 1982a; 1982b).

While the distinction between utilitarian and hedonic information systems is widely accepted, and research has demonstrated the predictive power of this value dichotomy in explaining the relative weights of IT adoption determinants (e.g., Van der Heijden, 2004, Wakefield and Whitten, 2006), recent technological developments seem to challenge its applicability. In particular, we refer to the emerging class of systems that provide both utilitarian and hedonic value (e.g., smart phones; mobile data services; virtual worlds). These systems, alluded to in our introduction as MPIS, incorporate instrumental and experiential functions in one environment (Hong and Tam, 2006). A main challenge for developers of MIPS is how to apply an integrated suite of instrumental and experiential functions, in order to enable productive and prolonged use. Such integration demands a thorough understanding of the drivers of the adoption process and stresses the need for an integration and extension of theories concerning IT adoption and use.

2.2 Motivation theory

A significant body of research has applied motivation theory (Deci, 1975; Deci and Ryan, 1985) to predict the acceptance of information systems. Referring to “enduring predispositions that arouse and direct behavior toward certain goals (Engel, Blackwell and Miniard, 1995, p. 49), motives are seen as key determinants of IT adoption. Two types of motivation exist: extrinsic motivation and intrinsic motivation.

Extrinsic motivation is defined as “doing something because it leads to a separable outcome” (Ryan and Deci, 2000b, p. 55). Extrinsically motivated behavior is driven by the instrumental value of the performed activity (Ryan and Deci, 2000b), which is derived from the outcome of the activity instead of the activity itself (Davis, Bagozzi and Warshaw, 1992). Given the instrumental value, extrinsic motives are theorized as dominant predictors of utilitarian system usage (Van der Heijden, 2004; Wakefield and Whitten, 2006). Most theories on IT adoption (see Venkatesh, Morris, Davis and Davis, 2003 for an overview) are productivity-oriented and use extrinsic motives such as ‘perceived usefulness’ and ‘perceived ease-of-use’ as key determinants of utilitarian system use.

Intrinsic motivation refers to “doing an activity for the inherent satisfaction of the activity itself” (Ryan and Deci, 2000a, p. 71). Intrinsically motivated behavior is self-determined, volitional (Deci and Ryan, 1985) and involves people engaging in activities that they find interesting, new, and optimally challenging (Deci and Ryan, 2000, p. 235). Intrinsic motivation derives its value from the appreciation of the activity itself rather than its instrumental outcome (Davis et al., 1992; Mathwick, Malhotra and Ridgon, 2001). Intrinsic motivation is closely associated with the general need for feelings of competence and autonomy, and integrally relates to emotions such as enjoyment and excitement (Deci, 1975; Deci and Ryan, 1985). The vast majority of IT adoption studies has centered on the intrinsic belief ‘perceived enjoyment’, which has been related successfully to hedonic system usage (see Davis et al., 1992; Van der Heijden, 2004; Wakefield and Whitten, 2006).

3 RESEARCH FRAMEWORK AND HYPOTHESES

This study applies an integrated approach to put forward a model of virtual world adoption. The logic behind the model structure is predicated on three theoretical considerations. First, using the attitude towards use as key independent variable, we include the general technology beliefs perceived
usefulness and perceived ease of use, as proposed in TAM. The empirical robustness of this structure supports the assumption that both constructs explain a substantial part of the attitude variance, and thus our decision. Second, we challenge the parsimony and claim to sufficiency of TAM and expand our model by including the extrinsic belief escapism and intrinsic beliefs visual attractiveness, intrinsic enjoyment and entertainment value. In line with research suggesting direct effects of ‘external variables’ on behavior (e.g., Hong and Tam, 2006; Van der Heijden, 2004) we posit direct effects of these beliefs on attitude. Third, based on classical drive theory (see Hull, 1943), we view intrinsic motivation as a reinforcer of extrinsic motivation. In the remainder of this section we elaborate on the model structure, and detail the research constructs and their hypothesized relationships.

3.1 General technology beliefs

Since its conception in 1989, the Technology Acceptance Model (TAM) has spawned a whole research tradition into the factors affecting adoption and use of various IT applications. This model is an adaptation of the theory of reasoned action (TRA) proposed by Fishbein and Ajzen (1975) which aims to explain and predict the behaviors of people in a specific situation. The core of TAM is built on two specific beliefs explaining user acceptance of information technology (Davis, 1989), namely perceived usefulness and perceived ease of use. Both these beliefs are extrinsic motivators, focusing on the instrumental value of system use (Venkatesh, 2000).

A fundamental assumption in TAM (and one that has received much empirical support) is that both these beliefs are positively related to attitude towards system use. This leads to our first two hypotheses:

**H1.** Perceived usefulness positively influences the attitude towards using virtual worlds.

**H2.** Perceived ease of use positively influences the attitude towards using virtual worlds.

Previous TAM-based research found that both beliefs are also mutually related, in the sense that a system’s perceived usefulness increases as its ease of use gets higher (Legris, Ingham and Collerette, 2003; Moon and Kim, 2001). Therefore, we hypothesize:

**H3.** Perceived ease of use positively influences perceived usefulness.

Escapism

As conceptualized by Hirschman (1983), escapism is an extrinsic motivation. It “offers the individual an avenue to a more desirable state of being than the one presently experienced” (p. 64), which means it has a utilitarian function as an anxiety reduction mechanism. At its core, escapism refers to users’ motivation to “leave the reality in which they live in a cognitive and emotional way” (Henning and Vorderer, 2001, p. 101). Although in consumer behavior research such motivations are often directly related to hedonic values (Babin et al., 1994; Bridges and Florsheim, 2007), and escapism can also be conceptualized as related to “playfulness” and “pretending” (Mathwick et al., 2001), we feel it is primarily an extrinsic motivation. The instrumental value that is derived from technology use (“getting away from it all”) is an important driver for behavior here. This is in line with Hirschman and Holbrook’s (1982) concept of “imaginative constructs of reality”, where acts are not based on what users know to be real, but rather on what they desire reality to be. Projective fantasies or absorbing experiences (Hirschman and Holbrook, 1982) can be strong drivers for the use of environments such as virtual worlds. The social play inherent in virtual worlds can be expected to positively contribute to this value, and thus be positively appreciated by individuals with escapist motivations, leading to the following hypothesis:

**H4.** Escapism positively influences the attitude towards using virtual worlds.

In line with the above, it can be assumed that this motivation will not only directly exert a positive influence on attitude towards use, but will also be positively related to perceived usefulness, where
anxiety reduction is the benefit that can be derived from system use (Hong and Tam, 2006). Therefore, we hypothesize:

**H5.** Escapism positively influences perceived usefulness of virtual worlds.

### 3.2 Elaborating intrinsic motivations: enjoyment, entertainment and visual attractiveness

#### Intrinsic enjoyment

The TAM has received considerable empirical support over the years (Adams, Nelson and Todd, 1992; Davis and Venkatesh, 1996; Legris et al., 2003), and has seen numerous extensions in terms of new variables and relationships being added to the model. Because social play is an important element supported by virtual worlds, a relevant extension of TAM in the context of the study reported here is the inclusion of perceived enjoyment. Davis et al. (1992) first introduced perceived enjoyment as “the extent to which the activity of using the computer is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated” (p. 1113).

The variable of enjoyment is especially relevant in relation to information systems with a hedonic function. Previous research in Web-based environments has found empirical support for the influence of enjoyment on attitude towards using, and intention to use such environments (Moon and Kim, 2001; Van der Heijden, 2003). In terms of the hedonic function of using virtual worlds, enjoyment can be expected to be an important factor:

**H6.** Intrinsic enjoyment positively influences the attitude towards using virtual worlds.

In line with the relationship between perceived ease of use and perceived usefulness, perceived ease of use can also be associated with enjoyment. Venkatesh (2000) argues that users’ perceptions of ease of use are adjusted based on actual experiences with a system. More specifically, Venkatesh argues, the more a user enjoys the interaction with a system, the more they tend to “…’underestimate’ the difficulty of the means or process of using a system because they quite simply enjoy the process and do not perceive it as being effortful…” (p. 348). This line of reasoning is also found in Venkatesh (1999) and Wakefield and Whitten (2006) and leads to the following hypothesis which shows a reinforcing effect of intrinsic on extrinsic motivators:

**H7.** Intrinsic enjoyment positively influences perceived ease of use of virtual worlds.

#### Entertainment value

Entertainment value can be defined as the perceived degree to which the use of an information system is a pleasant and likeable experience (Ducoffe, 1996) and lifts the user’s spirit (Mathwick et al., 2001). Although both intrinsic enjoyment and entertainment value are fun-related notions, they are conceptually distinct. An individual may experience an information system as entertaining, regardless of the reason why he uses it. In contrast, when this individual intrinsically enjoys using the information system, the pleasure is not just an added benefit of using the system to pursue some other goal, but rather is actively sought after and a purpose in itself (cf. Brown and Venkatesh, 2005; Mathwick et al., 2001).

As was argued in the above, based on the literature it can be expected that perceived fun, in the sense of intrinsic enjoyment, positively impacts the attitude towards using the virtual world. Given that intrinsic enjoyment and entertainment value concepts share the element of fun, the same arguments apply to entertainment value. Indeed, the literature provides a substantial body of evidence suggesting a positive effect of entertainment value on digital environment related attitudes such as the attitude towards using an information system (Dabholkar and Bagozzi, 2002), towards a website (Richard, 2005), and towards online advertising (Xu, Liao and Li, 2008). Accordingly, it can be anticipated that:
H8. Entertainment value positively influences the attitude towards using virtual worlds.

Arguably, when the use of an information system is perceived to be fun this is likely to increase the degree to which this fun is seen as an end in itself. Therefore, it is expected that:


A system that entertains the user provides this individual with a service and shows that it is capable of achieving something that the user values, regardless of why the system is used. Therefore, it stands to reason that the more the system is seen as entertaining, the more this gives the user the indication that the system will also perform adequately when it is put to use to increase the individual’s performance. This is supported by the research of D’Ambra and Rice (2001) who found a positive influence of entertainment value of websites on perceived personal performance. Consequently, we again hypothesize a reinforcing effect of an intrinsic on an extrinsic motivator:

H10. Entertainment value positively influences perceived usefulness of virtual worlds.

Visual attractiveness

The impact of the visual appeal of information systems on user perceptions and behavior has been investigated in the context of the World-Wide Web, Human Computer Interaction (HCI) and mobile commerce. For instance, visual attractiveness was found to be an important determinant of user satisfaction related to web site usage (Lindgaard and Dedek, 2003). Research performed in the field of mobile commerce revealed that design aesthetics appear to influence users’ loyalty towards mobile applications (Cyr, Head and Ivanov, 2006). This line of reasoning can be extended to virtual worlds, as virtual world technologies fully support the development of visually attractive environments aiming at creating game-like user experiences. In line with previous research (e.g., Van der Heijden, 2003) we assume that the visual attractiveness of virtual worlds will be associated with a positive attitude towards using such environments:

H11: Visual attractiveness positively influences attitude towards using virtual worlds

Visual attractiveness has also been found to be related to perceived ease of use. For example, Tractinsky, Katz and Ikar (2000) found a tight relationship between users’ perceptions of interface aesthetics and perceptions of the usability of an information system. Characteristics of the user-system interaction, such as visual appeal and usability, were found to play a role in driving perceived ease of use (Venkatesh, 2000). Van der Heijden (2003, p. 544) introduced the notion of “perceived visual attractiveness” defined as “the degree to which a person believes that the website is aesthetically pleasing to the eye” and found perceived visual attractiveness to positively influence both perceived usefulness and perceived ease-of-use. All in all, these findings lead to the following hypothesis, which again establishes a relationship between an intrinsic and an extrinsic motivator:

H12: Visual attractiveness positively influences perceived ease of use of virtual worlds.

4 METHOD

4.1 Procedure

We conducted a survey to collect empirical data. The sample consisted of Dutch registered users of the virtual world Second Life. An e-mail with an invitation to participate freely in the study was sent to a mailing list of 50,000 registered users. The invitation included a link that led to a web-based survey. As incentive, respondents were asked to fill in their e-mail address to engage in the raffle of ten book tokens of 20 euro. The online questionnaire addressed perceptions of perceived usefulness, perceived ease of use, escapism, intrinsic enjoyment, entertainment value, visual attractiveness, and attitude towards use. Moreover, socio-demographics were included.
4.2 Measures

Multi-item scales were used to measure the research constructs (see appendix). For each scale, items were collected based upon previous literature. We took care to solely select items that added to the conceptualization of the constructs as applied in our research. All items were part of reliable and validated measurement instruments. We tailored the items to the context of our study (i.e., made them target specific). All constructs were measured using 7-point Likert scales (strongly disagree-strongly agree).

4.3 Sample

The data was collected from the 29th of June to the 10th of July 2007. Eventually, 1627 respondents filled in the online questionnaire. The majority of the respondents was male (n=925; 56.9%), between 30 and 50 years old (n=862; 53%). A slight majority (52%, n=846) indicated to use Second Life once per week or more, while a minority (39.9%, n=649) reported to buy at Second Life. To investigate whether nonresponse bias posed a threat to the internal validity of the study, we compared the sample demographics with those of the population of Dutch Second Life users (cf. Armstrong, J.S., and Overton, 1977; Pavlou et al., 2007). Overall, no large differences in these demographics were found. This suggests that nonresponse bias was not a major concern in this study.

5 RESULTS

We first applied EFA to test whether the measurement items only tapped into one underlying dimension. Then, Partial Least Squares (PLS) modeling was utilized to further assess the measurement model and to estimate the structural model\(^2\). PLS was selected since our model is rather complex and our research intends to develop rather than confirm theory (see Fornell and Bookstein, 1982). Moreover, PLS has some specific advantages including minimal demands on data distribution (Chin, 1995).

5.1 Test of dimensionality

EFA was run using the principal components model with the oblique rotation technique (Direct Oblimin). The data passed the thresholds for sampling adequacy (KMO MSA 0.957, Bartlett’s test of sphericity 56554.6, p <.001). Four items were removed since they demonstrated high cross-loadings. The final factor solution was then derived. Together, the 7 factors explained 84.6% of the variance in the data. Unidimensionality of the measures was confirmed since each item loaded highest on its intended factor. Moreover, except for one intrinsic enjoyment item, all items loaded very strongly on their underlying factors. As such, preliminary evidence for convergent validity and discriminant validity was provided.

---

\(^1\) Data about the demographics of the Dutch Second Life user population was provided to us by the company responsible for Secondlife.nl.

\(^2\) We initially planned to use Structural Equation Modeling (SEM). First analyses resulted in poor fit. In such situations, PLS can be seen as feasible alternative for SEM as it places fewer demands on normality of data distributions and residual distributions (Fornell and Bookstein, 1982).
5.2 Measurement model

PLS was used to assess the reliability and validity of the measures. We utilized the software package Smart PLS (Ringle, Wende and Will, 2005) of the University of Hamburg to compute factor loadings, Cronbach’s alpha, composite reliability and Average Variance Extracted (AVE). In addition, SPSS (version 13, Mac) was used to compute minimum item-to-total correlations.

The results strongly confirmed the reliability of the measures. The alphas and composite reliability scores exceeded the value of 0.90, and all AVE’s surpassed the 0.50 guideline (see Ping, 2004). Convergent validity was assessed by factor loadings (PLS), alphas, AVE’s and minimum item-to-total correlations. All scores exceeded accepted rules of thumb (factor loadings: 0.70, see Ko, Kirsch and King, 2005; alpha: 0.80, see Ping, 2004; AVE: 0.50, see Wasko and Faraj, 2005; minimum item-to-total correlations, 0.40, see Jayanti and Burns, 1998). As such, convergent validity was established. Finally, we assessed for discriminant validity in two steps. First, we used the PLS output to study the within-construct item factor loadings and compared these loadings to across-construct item loadings (cf. Wasko and Faraj, 2005). Since all within-construct item loadings were high, and lower than the cross-loadings, discriminant validity can be assumed. Second, we studied the individual AVE’s and compared the scores with the squared correlations among the constructs (cf. Chin, 1998a).

![Table 1: Discriminant validity: AVE’s versus cross-construct squared correlations (note: the boldscores (diagonal) are the AVE’s of the individual constructs. Of the diagonal are the squared correlations between the constructs).](image)

<table>
<thead>
<tr>
<th>Construct</th>
<th>Perceived usefulness</th>
<th>Perceived ease of use</th>
<th>Escapism</th>
<th>Intrinsic enjoyment</th>
<th>Entertainment value</th>
<th>Visual attractiveness</th>
<th>Attitude towards use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived usefulness</td>
<td>0.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived ease of use</td>
<td>0.16</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Escapism</td>
<td>0.38</td>
<td>0.11</td>
<td>0.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrinsic enjoyment</td>
<td>0.21</td>
<td>0.17</td>
<td>0.25</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entertainment value</td>
<td>0.41</td>
<td>0.23</td>
<td>0.44</td>
<td>0.53</td>
<td>0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual attractiveness</td>
<td>0.23</td>
<td>0.16</td>
<td>0.27</td>
<td>0.34</td>
<td>0.51</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>Attitude towards use</td>
<td>0.32</td>
<td>0.13</td>
<td>0.22</td>
<td>0.27</td>
<td>0.37</td>
<td>0.26</td>
<td>0.82</td>
</tr>
</tbody>
</table>

All AVE’s exceeded the values of the squared correlations among the constructs in the corresponding rows and columns. As such, discriminant validity was demonstrated.

5.3 Structural model

We then estimated the path coefficients (β) and $R^2$ values of the structural model using the bootstrapping technique (1627 re-samples; cf. Hesterberg et al., 2003). To evaluate the significance and effect sizes of the path coefficients, we conducted two-tail t-tests with a significance level of 0.01 (see Blalock, 1979). The procedure of Chin (1998a) was applied to compute effect sizes.

<table>
<thead>
<tr>
<th>Hyp. Path</th>
<th>β</th>
<th>t-value</th>
<th>Sign.</th>
<th>Effect</th>
</tr>
</thead>
</table>

Proceedings ECIS 2009
Table 2: PLS path coefficients, t-tests and effect sizes (n=1627).

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Path Coefficient</th>
<th>t-test</th>
<th>p-value</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Perceived usefulness → Attitude towards use</td>
<td>0.294</td>
<td>10.528</td>
<td>&lt;.001</td>
<td>0.09</td>
</tr>
<tr>
<td>2 Perceived ease of use → Attitude towards use</td>
<td>0.031</td>
<td>1.360</td>
<td>n.s.</td>
<td>0.00</td>
</tr>
<tr>
<td>3 Perceived ease of use → Perceived usefulness</td>
<td>0.114</td>
<td>5.298</td>
<td>&lt;.001</td>
<td>0.02</td>
</tr>
<tr>
<td>4 Escapism → Attitude towards use</td>
<td>-0.004</td>
<td>0.152</td>
<td>n.s.</td>
<td>0.00</td>
</tr>
<tr>
<td>5 Escapism → Perceived usefulness</td>
<td>0.345</td>
<td>14.151</td>
<td>&lt;.001</td>
<td>0.13</td>
</tr>
<tr>
<td>6 Intrinsic enjoyment → Attitude towards use</td>
<td>0.153</td>
<td>4.986</td>
<td>&lt;.001</td>
<td>0.03</td>
</tr>
<tr>
<td>7 Intrinsic enjoyment → Perceived ease of use</td>
<td>0.262</td>
<td>9.255</td>
<td>&lt;.001</td>
<td>0.06</td>
</tr>
<tr>
<td>8 Entertainment value → Attitude towards use</td>
<td>0.207</td>
<td>4.970</td>
<td>&lt;.001</td>
<td>0.03</td>
</tr>
<tr>
<td>9 Entertainment value → Intrinsic enjoyment</td>
<td>0.730</td>
<td>48.047</td>
<td>&lt;.001</td>
<td>1.14</td>
</tr>
<tr>
<td>10 Entertainment value → Perceived usefulness</td>
<td>0.356</td>
<td>14.794</td>
<td>&lt;.001</td>
<td>0.12</td>
</tr>
<tr>
<td>11 Visual attractiveness → Attitude towards use</td>
<td>0.125</td>
<td>4.173</td>
<td>&lt;.001</td>
<td>0.02</td>
</tr>
<tr>
<td>12 Visual attractiveness → Perceived ease of use</td>
<td>0.246</td>
<td>8.438</td>
<td>&lt;.001</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Figure 1: Results structural model. Note: all path coefficients are significant at p <.001.

Overall, the results strongly confirm the predictive power of the model. The amount of variance explained is high, implying a good fit to the data. The results indicate that ten hypotheses (H1, H3, H5,
H6, H7, H8, H9, H10, H11, H12) are supported and two (H2, H4) are rejected. Six significant paths are of substantial strength (>0.19; cf. Chin, 1998b), while significant but smaller effects are reported for four paths.

6 DISCUSSION

6.1 Key Findings

Our research yields a number of key findings. First, the study reveals direct effects of the extrinsic belief perceived usefulness and the intrinsic beliefs intrinsic enjoyment, entertainment value and visual attractiveness on the attitude towards virtual world usage. Together, these concepts explained 45% of the attitude variance, which is quite impressive. Perceived usefulness and entertainment were the strongest predictors. The influences of intrinsic enjoyment and visual attractiveness were less strong. Overall, these findings confirm the proposed expansion of IT adoption drivers beyond general technology beliefs, and support the inclusion of intrinsic enjoyment, entertainment value and visual attractiveness. Second, our work provides strong support for the reinforcing role of intrinsic motivation on extrinsic motivation. We demonstrate that higher levels of enjoyment, entertainment, and visual attractiveness contribute to the perceived usefulness and perceived ease of use of virtual world systems. All proposed effects were significant and rather strong in nature. Adding to our first key finding, this finding suggests that intrinsic beliefs function as first- and second-order determinants of the attitude. Third, the results show that perceived usefulness mediates the effects of the extrinsic motives escapism and perceived ease of use on attitude. Contrasting our expectations, perceived ease of use and escapism did not directly relate to attitude. An explanation for the non-significance of perceived ease of use comes from literature claiming that ease of use solely affects behavior when a) users are inexperienced (Venkatesh, 2000) and/or b) the system is highly unusable (Van der Heijden and Verhagen, 2004). Since the vast majority of our respondents used Second Life on a regular base, these conditions are unlikely to apply to our research. Regarding the non-significance of escapism, we notice that, in line with the literature, its conceptualization reflects the instrumental value of a virtual world to achieve a particular goal (i.e., to ‘escape’). As such, this task-specific behavioral belief reflects a perceived system benefit and is most likely to be used directly to form assessments about the general behavioral belief usefulness (Venkatesh and Davis, 2000; see also Ajzen and Fishbein, 1980), and not to judge about an overall disposition towards system use (i.e., the attitude).

6.2 Theoretical implications

This research is one of the first to understand the adoption of virtual worlds. We introduced and tested a model that integrates and extends existing models on utilitarian and hedonic system use. Given the fit of our model, our results seem to suggest that existing utilitarian and hedonic models fall short in explaining virtual world use. Our findings underline that user adoption of virtual worlds depends on two belief configurations, and support that extrinsic and intrinsic motives function simultaneously in adoption decision-making (cf. Engel et al., 1995). Future theoretical efforts could incorporate this evidence and test its applicability across different contexts.

To advance our understanding of the behavioral dynamics behind virtual world use, this study suggests two patterns concerning the consequences of intrinsic motivation. First, we challenged the claim to sufficiency of previous models such as TRA and TAM (cf. Venkatesh, 1999), and demonstrated that intrinsic motivation directly affects the attitude. Second, we showed that intrinsic motivation reinforces the effect of extrinsic motivation on attitude. Although this proposed order effect is theoretically plausible, and aligns with the fact that we included intrinsic motives to an existing base of extrinsic technology beliefs (and not vice versa; see also Venkatesh, 2000), this finding contrasts with previous IT adoption works showing that extrinsic motivation precedes intrinsic motivation (e.g.,
Davis et al., 1992; Van der Heijden, 2004). Apparently, the relationships between extrinsic and intrinsic motivation are more complex than initially thought and more research is needed.

6.3 Implications for practice

Our findings have several practical implications. First, our findings highlight the roles of intrinsic and extrinsic motivation as direct determinants of virtual world adoption. To directly affect adoption behavior, developers should enhance the usefulness, entertainment, enjoyment, and visual attractiveness of a given virtual world. As such, both productive and prolonged use can be triggered in a direct manner. Second, our research draws attention to the potential of integrating instrumental and enjoyable, new, challenging functions and features. The reinforcing influence of intrinsic motivation on extrinsic motivation guides such integration. For instance, employing features like attractive designs, funny navigation (e.g., flying; using cars, boats, planes), and entertaining video and sound can enhance usefulness and ease of use. Third, to further prioritize development efforts, our findings underline the vital function of perceived usefulness and entertainment value. Both constructs have strong direct effects on the attitude. Moreover, usefulness mediates the influence of ease of use and escapism on the attitude, and entertainment value has strong second-order effects via usefulness and intrinsic enjoyment. Being key influencers of virtual world adoption, a focus on both constructs and their determinants is likely to enhance the effectiveness of system development endeavors in the broader context of IS development in general.

6.4 Limitation and additional recommendation

The potential limitations of this study concern sample bias. While no evidence of nonresponse bias was found, self-selection effects caused by the use of a self-administered questionnaire cannot be completely ruled out. In addition, the sample of Dutch Second Life users implies external validation in Western culture. Scholars in the fields of information system research (e.g., Al-Gahtani, Hubona and Wang, 2007; Straub, Loch and and Hill, 2001), however, have shown that cultural differences are likely to affect IT adoption. For example, in comparing collectivistic versus individualistic cultures Davis, Wang and Lindridge (2007) demonstrate that online consumers in collectivistic societies suppress the exploration and expression of emotions. In such settings, perceptions of pleasure can be assumed to have a relatively weak impact on online system use (Davis et al., 2007). Extrapolating this research finding to virtual world adoption, it is believable that the influence of intrinsic motivation is weaker in non-Western cultures. Future research may test this assumption by cross-validating our research across different cultures.

7 REFERENCES


APPENDIX

*** = removed after EFA

Perceived usefulness (general) (Hong, Thong and Tam, 2006; Porter and Donthu, 2006; Van der Heijden, 2003). Mean (SD) = 3.17(1.466)
1. I find <name virtual world> useful in my life.
2. I find that <name virtual world> adds value to my life.
3. Overall, <name virtual world> is useful.
4. Using <name virtual world> helps me accomplish things more quickly.
5. Using <name virtual world> makes my life easier.
6. Using <name virtual world> can make one productive.***
7. <Name virtual world> can make things easier. ***

Perceived ease of use (Hong et al., 2006; Porter and Donthu). Mean (SD) = 3.91 (1.628)
1. Learning how to use <name virtual world> is easy.
2. <name virtual world> is clear and understandable to use.
3. It is easy to become skilful at using <name virtual world>.
4. Overall, <name virtual world> is easy to use.

Escapism (Kim and Kim, 2005; Mathwick et al., 2001). Mean (SD) = 3.96 (1.617)
1. Using <name virtual world> makes me feel like I am in another world.
2. Using <name virtual world> “gets me away from it all”.
3. I get so involved when I use <name virtual world> that I forget everything else.
4. Using <name virtual world> truly feels like “an escape”.

Intrinsic enjoyment (Mathwick et al., 2001; Van der Heijden, 2003). Mean (SD) = 5.43 (1.263)
1. Irrespective of whether <name virtual world> gives me what I want, I enjoy using it for its own sake.***
2. I use <name virtual world> for the pure enjoyment of it.
3. I use <name virtual world> for pleasure.
4. Using <name virtual world> is an agreeable way of passing time.

Entertainment value (Mathwick et al., 2001; Richard, 2005). Mean (SD) = 4.87 (1.426)
1. I think <name virtual world> is very entertaining.
2. The enthusiasm of <name virtual world> is catching, it picks me up.
3. Using <name virtual world> entertains me.
4. I think <name virtual world> is exciting. ***
5. I think <name virtual world> is imaginative.
**Visual attractiveness** (Mathwick et al., 2001; Van der Heijden, 2003). Mean (SD) = 4.92 (1.336)
1. The way things are displayed in <name virtual world> is attractive.
2. <name virtual world> is aesthetically appealing.
3. I like the way <name virtual world> looks.
4. Overall, I find that <name virtual world> looks attractive.

**Attitude** (Moon and Kim, 2001). Mean (SD) = 4.86 (1.183)
1. Using <name virtual world> is a good/bad idea.
2. Using <name virtual world> is a wise/foolish idea.
3. Using <name virtual world> is a pleasant/unpleasant idea.
4. Using <name virtual world> is a positive/negative idea.
USER ACCEPTANCE OF SECOND LIFE: AN EXTENDED TAM INCLUDING HEDONIC CONSUMPTION BEHAVIOURS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0269.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Technology Acceptance Model (TAM), Virtual world, Consumer behavior / choice / demand / empowerment / reviews / consumerism, Education</td>
</tr>
</tbody>
</table>
USER ACCEPTANCE OF SECOND LIFE: AN EXTENDED TAM INCLUDING HEDONIC CONSUMPTION BEHAVIOURS

Saeed, Nauman, Faculty of ICT, Swinburne University of Technology, PO Box 218, Hawthorn 3122, Victoria, Australia, nsaeed@swin.edu.au

Yang, Yun, Faculty of ICT, Swinburne University of Technology, PO Box 218, Hawthorn 3122, Victoria, Australia, yyang@swin.edu.au

Sinnappan, Suku, Faculty of Higher Education, Swinburne University of Technology, PO Box 218, Lilydale 3140, Victoria, Australia, ssinnappan@swin.edu.au

Abstract

Second Life is a 3-D multi-user virtual environment which has gained widespread popularity amongst academic community in the recent years. However, due to its infancy very little is known about the factors driving users’ intention to use Second Life especially in the educational context. This paper presents findings from an ongoing study about the impacts of using multi-user virtual environments in higher education. In information systems research, several models and frameworks have attempted to predict the acceptance of new technology. However some recent studies suggest that the traditional technology acceptance approaches may not work well with today’s entertainment-oriented technologies such as multi-user virtual environments. They also recommend exploring those facets of human behaviour that are likely to capture the hedonic consumption of such technologies. In this paper, we propose an extended technology acceptance model (TAM) including hedonic consumption behaviours in order to explain the usage and acceptance of Second Life in the educational context. The proposed model is empirically evaluated using survey data collected from 122 users about their perceptions of Second Life. Findings suggest that hedonic consumption behaviours are strong predictors of Second Life usage as compared to traditional motivational constructs of usefulness and ease-of-use.

Keywords: TAM, Second Life, education, hedonic consumption behaviours.
1 INTRODUCTION

Second Life is an emerging 3-D multi-user virtual environment (MUVE) which is increasingly gaining wide spread acceptance from educators and students. A large number of prominent educational institutions have established their virtual campuses in Second Life (a complete list of participating institutions is available at: http://simteach.com/wiki/index.php?title=Second_Life_Education_Wiki#Institutions_and_Organizations_in_SL/). Despite the fact that emerging Web technologies like Second Life are seen as a next-generation platform for Web users, their acceptance remains a big challenge in order to become mainstream (Shin & Kim, 2008). Also, due to its infancy, very little academic research is available on the usage and acceptance of Second Life. Therefore it is important to understand the significant predictors of Second Life usage in order to utilise it as an effective teaching and learning tool.

In information systems research, several theoretical models or frameworks attempted to explain the acceptance of new technology. Among these, technology acceptance model (TAM) is the most researched one. TAM is originally developed by Davis and his colleagues to explain or predict individuals’ acceptance of computer based systems and underlying influencing factors (Davis, 1989). TAM has been validated for a number of productivity-oriented technologies like word processors, spread sheets, Web-based learning environments and multimedia application. However, some recent studies suggest that traditional technology acceptance approaches such as TAM may not work well for today’s entertainment-oriented technologies like multi-player online games or MUVEs (Heijden, 2004; Holsapple & Wu, 2007). They also recommend to consider those facets of human behaviour that are likely to capture the hedonic consumption of such technologies (Hsu & Lu, 2004). Since Second Life also inherits a large entertainment element, it is important to capture its hedonic consumption in order to explain its usage and acceptance. In this paper, we present an extended TAM including hedonic consumption behaviours in order to explain the user acceptance of Second Life. An online survey was developed using constructs from previously published research and participants were employed through various resources within Second Life as well as externally. A total of 122 participants including students, teachers, researchers and academic managers responded to our survey. The PLS (Partial Least Squares) approach was used for data validation and hypotheses / model testing. All of the hypothesised relationships formulated at the start were confirmed by the data. Findings suggest that hedonic consumption behaviours are strong predictors of Second Life usage than traditional motivational behaviours. Our model explains 51.4% of the users’ intentions to use Second Life, which is a significant outcome when compared with similar studies of technology acceptance.

The rest of the paper is organised as follows. Section 2 presents an overview of Second Life and its applications in academia, technology acceptance model and the theory of hedonic consumption behaviours. The research hypotheses and the model are presented in Section 3. We present the research methodology in Section 4. Section 5 presents results of data validation and hypotheses / model testing. We discuss the study findings, implications and limitations in Section 6. Section 7 concludes the study and points out future work.

2 BACKGROUND

2.1 Second Life and education

Second Life is a 3-D multi-user virtual environment (MUVE) launched by Linden Labs in 2003. Second Life is a world solely created by its inhabitants, called ‘residents’. Residents have the opportunities to create their digital proxies called ‘avatars’ and design their clothing, hair colour, dresses and even appearances (Coffman & Klinger, 2007). Avatars can walk, run, or even fly in the virtual environment. They can converse with other avatars using text, images, gestures or even voice.
Residents can move or ‘teleport’ from one location to another. Second Life provides enormous opportunities to imitate real world situations in a virtual environment, to name a few: reincarnation of ancient architecture and civilisations (Harrison, 2009); advertising and selling of real life commodities (Lui, Piccoli, & Ives, 2007); experiencing complex medical procedures discounting dangerous outcomes (Thompson & Hagstrom, 2008); library services (Bell, Peters, & Pope, 2007); conducting classes and labs (Holmberg & Huvila, 2008); and many more.

3-D MUVEs like Second Life offer a variety of potential benefits for educational use including: collaboration and communication, engagement, conducting activities in a risk-free environment, alternative space for instruction and tasks, and visualisation of difficult content (Eschenbrenner, Nah, & Keng, 2009). (Richter, Anderson-Inman, & Frisbee, 2007) have identified at least five different types of learner engagement that are possible in Second Life: experiential, diagnostic, demonstrative, role-play and constructivist. Second Life has the potential to be a useful educational tool for teaching and learning by using a constructivist approach (Coffman & Klinger, 2007), which is the theory of knowledge acquisition obtained through interactions and building upon own knowledge and which produces the highest type of learning according to Bloom (Cheal, 2007). Following this approach, students can discover and create meaningful content and interactions (Stevens, 2006). Teachers in higher education have found Second Life a convenient place to conduct online classes, conferences, presentations, and meetings with students (Richter, et al., 2007). However, in order to explore the teaching and learning potential of Second Life, it is important to understand the factors that affect users’ intentions toward Second Life usage. In this paper, we aim to propose and evaluate a technology acceptance model to explain user acceptance of Second Life within educational domain.

2.2 Technology acceptance model (TAM)

TAM posits that user perceptions of usefulness and ease-of-use determine attitudes towards using a system or technology. An individual’s attitude is hypothesised to influence the behavioural intention to use a technology, which in turn leads to the actual use. In the follow-up model, TAM2 (Venkatesh & Davis, 2000), the attitude component was dropped and perceived technology characteristics directly influenced the individual’s intention to use the technology. Social influences (also referred as subjective norms) were also included in the follow-up model. Both TAM and TAM2 have established themselves as being robust and parsimonious for predicting user adoption of a variety of new technologies (Raaij & Schepers, 2008) and have been validated for a variety of productivity-oriented technologies including word processors, e-mail, spreadsheets, Web-based learning systems, and multimedia learning systems (Halawi & McCarthy, 2007; Lederer, Maupin, Sena, & Zhuang, 2000; Saade, Nebebe, & Tan, 2007). However, some previous studies suggest that traditional technology acceptance approaches like TAM may not work well with today’s entertainment-oriented technologies such as multi-player online games or MUVEs (Heijden, 2004; Hsu & Lu, 2004; Koufaris, 2002). Because these technologies enable users to fantasise, role-play and be entertained, it is important to measure their hedonic consumption along with the traditional user acceptance behaviours, in order to better understand their usage patterns. In this paper, we present an extended TAM2 including hedonic consumption behaviours to predict the user intentions towards Second Life usage.

2.3 Theory of hedonic consumption behaviours

With its roots in marketing research, hedonic consumption designates those facets of user behaviour that relate to the multi-sensory, fantasy and emotional aspects of one’s experience with products (Hirschman & Holbrook, 1982). Hedonic perspective seeks not to replace traditional theories of consumption but rather to extend and enhance their applicability. The theory asserts that emotional and imaginative responses are the main drivers of hedonic consumption, which can be explained through the constructs of emotional involvement, enjoyment and role projection (Lacher & Mizerski, 1994). Several studies have reported significant impact of these constructs in explaining the consumption of entertainment-oriented technologies or systems: including online games, virtual
learning environments, online retail shopping, music, and gambling (Childers, Carr, Peck, & Carson, 2001; Lacher & Mizerski, 1994; Lee, Cheung, & Chen, 2005; Mun & Hwang, 2003; Shin & Kim, 2008; Titz, Andrus, & Miller, 2002). The traditional economic view of products as objects would seem inappropriate for products whose usage are based upon satisfying emotional desires rather than fulfilling utilitarian functions (Kim & Forsythe, 2007). Thus for the systems that are hedonic in nature, hedonic factors could be the dominant predictors of attitude towards their usage. Similarly, for MUVEs (such as Second Life) that are largely hedonic in nature, we can expect hedonic behaviours to be a strong predictor of attitudes towards using these technologies. Therefore, in this paper, we aim to examine the impact of hedonic consumptions behaviours on user intentions toward Second Life usage.

3 RESEARCH HYPOTHESES AND MODEL

3.1 TAM2 hypotheses

Perceived usefulness and perceived ease-of-use are the basic TAM constructs. In general, perceived usefulness reflects an individual's subjective estimation of the job performance enhancement that is likely to result from the use of a new technology, whereas perceived ease-of-use refers to the degree to which he or she expects the use of the technology to be free of effort (Davis, Bagozzi, & Warshaw, 1989). Both of these constructs constitute a significant influence on an individual's intention to use a technology or system (Ma & Liu, 2004). Perceived ease-of-use is also reported a significant determinant of perceived usefulness (Davis, 1989). We follow this trend and hypothesise the following:

H1. The perceived ease-of-use (PEU) of Second Life will have a positive impact on perceived usefulness (PU) of Second Life.

H2. The perceived ease-of-use (PEU) of Second Life will have a positive impact on behavioural intention (BI) to use Second Life.

H3. The perceived usefulness (PU) of Second Life will have a positive impact on behavioural intention (BI) to use Second Life.

Subjective norms refer to the degree to which a person believes that those who are important to him or her think that he or she should perform the behaviour in question (Fishbein & Ajzen, 1975). Previous studies have shown that both peers’ and superiors’ influences can affect a person’s decision to accept a new technology (Mathieson, 1991; Taylor & Todd, 1995). Subjective norms can also influence technology acceptance via perceived usefulness, which is referred as ‘internationalisation’ mechanism (Venkatesh & Davis, 2000). On the basis of this, we hypothesise the following:

H4. The subjective norms (SN) will have a positive impact on perceived usefulness (PU) of Second Life.

H5. The subjective norms (SN) will have a positive impact on behavioural intention (BI) to use Second Life.

3.2 Influence of hedonic consumption behaviours

The hedonic consumption theory focuses on positive behavioural experience: emotional and imaginative responses, which are key human factors and are likely to capture the entertainment nature of the technology of interest (Holsapple & Wu, 2007). These responses can be explained through emotional involvement, enjoyment and role projection constructs. Emotional involvement is defined as the degree to which an individual is emotionally engaged in a behaviour; enjoyment is defined as the degree to which performing an activity is perceived as providing pleasure and joy in its own right, aside from performance consequences (Venkatesh, 2000); and role projection involves the mental
activities whereby individuals project themselves into particular roles or characters (Holsapple & Wu, 2007). Some recent studies have shown the significant effect of these constructs on user acceptance of entertainment-oriented technologies (Depradine, 2007; Shin & Kim, 2008). We thus hypothesise the following:

H6: The perceived emotional involvement (PEI) will have a positive impact on the behavioural intention (BI) to use Second Life.

H7: The perceived enjoyment (PEN) will have a positive impact on the behavioural intention (BI) to use Second Life.

H8: The perceived role projection (PRP) will have a positive impact on the behavioural intention (BI) to use Second Life.

3.3 Extended TAM2

On the basis of above hypotheses, we present an extended TAM2 including hedonic consumption behaviours (in terms of emotional involvement, enjoyment and role projection) as significant predictors of users intentions to use Second Life, as shown in Figure 1.

![Figure 1. An extended TAM2 including hedonic consumption behaviours](image)

4 METHODOLOGY

4.1 Participants and data collection

Since the aim of our study was to capture the user acceptance of Second Life in the educational context, our target subjects were academic-centric participants who either have attended, designed, conducted classes or involved in some form of educational activities in Second Life. Data for this research were collected through in-world (within Second Life) and external resources. First of all, the authors joined several educational and research groups within Second Life to get an insight of the
ongoing educational activities and to make social contacts with the community. We visited a large number of educational islands and personally invited the residents to take part in our online survey. We also sent invitations to various educational groups within Second Life. The external resources used for data collection were the two popular mailing lists; Second Life Educators mailing list (SLED) and Second Life Research Listserv (Slrl). These lists are populated by large number of geographically distributed and active academic-centric individuals and are considered a constant ground for data collection. Most of the data used in this study come from external resources. Some recent studies have adopted similar approaches in exploring this new technology (Second Life) within the education context (Alvarez, 2006; Boostrom, 2008; Richter, et al., 2007). The total number of valid responses we obtained from all above resources was 122, which sufficed to perform PLS analysis.

4.2 Measures

Multiple items were adopted from previously published scales for the constructs used in our model. All items were measured on a seven point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). The scales for perceived ease-of-use (PEU), perceived usefulness (PU) and subjective norms (SN) were adopted from (Davis, et al., 1989) and (Igbaria, 1990). The scales for behavioural intention (BI) were adopted from (Venkatesh, Morris, Davis, & Davis, 2003). The scales to measure the perceived emotional involvement (PEI), perceived enjoyment (PEN) and perceived role projection (PRP) were adopted from (Hirschman, 1983) and (Hirschman & Holbrook, 1982). The complete list of measurement items is presented in the Appendix. Demographic items are not included due to space limit.

5 RESULTS

The research model was tested using the PLS approach. PLS is considered as a powerful tool in analysing structural models involving multiple constructs and multiple indicators. Previous research shows that PLS approach is more suitable for prediction as compared to other approaches like LISREL and EQs, because it assumes that all the measured variance in the study will be explained (Chin, 1998; Saade & Bahli, 2005). The PLS approach has been used in several other studies of technology acceptance such as (Raaij & Schepers, 2008), (Saade & Bahli, 2005) and (Mun & Hwang, 2003), thus deemed suitable for our study. After ensuring the reliability and validity of the scales, we tested the hypothesised relationships between hedonic consumption behaviours and TAM2 constructs using bootstrap re-sampling method (Cotterman & Senn, 1992). Unlike other structural equation modelling approaches such as LISREL, the primary objective of the PLS approach is the maximisation of variance explained, not the minimisation of the difference between the observed and the reproduced covariance matrices (Hulland, 1999). Thus the quality of the PLS approach can be determined by examining the R^2 values of the dependant constructs.

5.1 Demographics

Our sample included 79 females and 43 males. The mean age of the participants was 42, ranging from 17-65 years. Teachers (27%) and students (25%) constituted the larger groups while researchers the smallest (13%) group. The majority of participants appeared well educated as 67.2% of them held postgraduate qualifications. The survey results also showed that 82.8% of participants had Internet experience of more than 9 years, 94.3% of them used the Internet several times a day and the primary access location of the Internet for 56.6% of participants was the home. In addition, the majority of participants (81.9%) had at least 6 months experience of using Second Life. About half of them admitted of accessing Second Life at least once a day and the primary access location for 72.1% of participants was the home. These results suggest that our participants were mature, well educated and geographically distributed. They had extensive experience of using Web-based applications with
adequate experience of using Second Life in the educational contexts. Thus they fit well into our target participants’ profile.

### 5.2 Data validation and reliability

Table 1 presents the summary of all construct reliability measures including factor loadings, t-values, composite reliability (CR), and average variance extracted (AVE). The factor loadings provide evidence for convergent validity as all our constructs load greater than the threshold of 0.50 as suggested by (Peterson, 2000). The t-values derived from our analysis also provide evidence for convergent validity since all values exceed the threshold of 1.96 as suggested by (Gefen & Straub, 2005). Internal consistency appears significant for all of our constructs since the composite reliability values exceed the minimum of 0.70 as suggested by (Nunnally & Bernstein, 1994).

![Table 1. Construct reliability measures.](image)

Discriminant validity was met using the Fornell and Larcker test (Fornell & Larcker, 1981). The procedure involves computing the square root of the AVE of each construct, which should exceed the correlation shared between the construct and other constructs in the model. Table 2 shows that square roots (in bold) of all AVEs (on the diagonal) are greater than the cross-correlations of all other constructs. Thus all our constructs demonstrated a good degree of validity and reliability.

![Table 2. Discriminant validity of constructs.](image)
5.3 Hypotheses and model testing

Figure 2 summarises the results of hypotheses testing. We found a direct positive impact of perceived ease-of-use (PEU) on perceived usefulness (PU) and behavioural intention (BI), thus supporting hypotheses H1 and H2 respectively. The perceived usefulness (PU) also had a direct positive impact on behavioural intention (BI), supporting hypothesis H3. Subjective norms (SN) exhibited direct positive impact on perceived usefulness (PU) and behavioural intention (BI), thus supporting hypotheses H4 and H5 respectively. The three relationships (H6, H7, and H8) between the hedonic constructs (PEI, PEN, PRP) and the behavioural intention (BI) also appeared significant, with perceived enjoyment (PEN) as the strongest predictor of users’ intentions to use Second Life. Thus all hypothesised relationships formulated at the start were confirmed by the data. Our model explained 51.4% of the variance through user’s intentions to use Second Life, which is a significant outcome when compared with similar studies of technology acceptance such as: (Mun & Hwang, 2003; Raaij & Schepers, 2008; Shin & Kim, 2008).

Figure 2.PL5 results.
*Path coefficient significant at the 0.5 level; ** at the 0.01 level; *** at the 0.001 level

6 DISCUSSION

The study broadens our understanding of technology acceptance by including emotional and imaginative responses of hedonic consumption behaviours as key determinants of Second Life usage. All eight hypotheses postulated at the start were confirmed by the data. Several insightful results could be summarised form our research model as follows: First, perceived enjoyment was the most important predictor of users’ intention to use Second Life. This is consistent with the findings of (Lee, et al., 2005), which have reported the direct and significant impact of perceived enjoyment on user intentions to use Internet-based learning medium (ILM). Similarly, (Childers, et al., 2001) have reported the strong impact of enjoyment on online retail shopping behaviour.
Second, the next strongest and direct impact on use intentions was of perceived emotional involvement followed by the perceived role projection. Thus, the hedonic consumption behaviours as a whole appeared as the strongest predictor of user intentions to use Second Life. This is consistent with the findings of (Kim & Forsythe, 2007), which have reported the stronger impact of hedonic motivations than functional motivations toward using product virtualisation technologies.

Third, TAM2 constructs of perceived ease-of-use, perceived usefulness and subjective norms also had significant impact on user intentions to use Second Life but with lesser significance than that of hedonic consumption behaviours. This is consistent with the findings of (Shin & Kim, 2008), which suggests that in case of Web 2.0 technologies users may need a clear motivations of usefulness and ease-of-use than those of traditional Web technologies. This is also in line with some recent studies which have predicted that the traditional technology acceptance approaches such as TAM may not work well with today’s multi-user and entertainment-oriented technologies like Second Life (Heijden, 2004; Holsapple & Wu, 2007; Hsu & Lu, 2004).

The implications of our results for educators and researchers are multi-fold: The educators of Second Life should include elements of enjoyment, emotional involvement and role projection while designing academic activities as the findings suggest that the more the users enjoy, get emotionally involved and be able to project themselves in the virtual environment, the more likely they will use Second Life, thus enhance their learning. Similar trends are illustrated in practice by the more popular islands in Second Life such as the Genome Island (http://connect.educause.edu/Library/EDUCAUSE+Review/GenomeIsland/47234?time=1226972674) and others (see http://sleducation.wikispaces.com/educationaluses). Here, students are presented with opportunities to role play, co-create and get involved in real time educational process. This presents students a much more rich and rewarding learning experience. For researchers, the study offers a platform to further explore hedonic consumption of today’s social and entertainment-oriented technologies like multiplayer online games, social networking Websites, Web authoring tools, and other emerging Web technologies. The study also suggests that hedonic behaviours should be given due consideration while measuring the user acceptance of entertainment-oriented technologies.

Like any other user study, our study also has limitations. First of all, our sample may be biased as our respondents were more likely to be engaged with the Second Life environment than non-respondents. Therefore, respondents could have been captured to the ‘hedonic consumption behaviours’. Secondly, the study results cannot be generalised to other MUVEs because our sample represents users of Second Life only. Another limitation of our study is the use of self-reported usage data which are often measured by log files. However, it was beyond our control to obtain such data as majority of the participants were employed through external resources and we had no means of maintaining their usage log files.

7 CONCLUSION AND FUTURE WORK

Our study provided a useful insight into the usage and acceptance of Second Life in the educational context. Instead of focusing on the traditional motivational and performance-based determinants of technology acceptance, the study focused on the hedonic consumption of technology because the targeted technology (Second Life) possessed a large entertainment element. Empirical findings provided support for our proposed model. Although hedonic consumption behaviours explained a significant amount of users’ intentions toward Second Life usage, more research would be carried out in our future studies to explore other key usage predictors such as user involvement, presence, flow, critical mass, fantasy, escapism, arousal and social pressures. We would also consider conducting similar studies with other popular MUVEs such as Haboo, There or Active Worlds to better understand user acceptance of MUVEs in education and to utilise them as effective teaching and learning tools.
References


Proceedings ECIS 2009


**Appendix – Measurement Items**

<table>
<thead>
<tr>
<th>PEU1</th>
<th>Learning to use Second Life is easy for me.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEU2</td>
<td>I find it not difficult to get Second Life to do what I want it to do.</td>
</tr>
<tr>
<td>PEU3</td>
<td>I find Second Life to be flexible to interact with.</td>
</tr>
<tr>
<td>PEU4</td>
<td>It is easy for me to become skilful at using Second Life.</td>
</tr>
<tr>
<td>PU1</td>
<td>Using Second Life enables me to accomplish my tasks more quickly.</td>
</tr>
<tr>
<td>PU2</td>
<td>Using Second Life improves my class or work performance.</td>
</tr>
<tr>
<td>PU3</td>
<td>Using Second Life increases my productivity.</td>
</tr>
<tr>
<td>PU4</td>
<td>Using Second Life makes it easier for me to understand lecture.</td>
</tr>
<tr>
<td>PU5</td>
<td>Using Second Life makes it easier for me to communicate with lecturer/friends.</td>
</tr>
<tr>
<td>PU6</td>
<td>Overall, I find Second Life useful in my study/work.</td>
</tr>
<tr>
<td>SN1</td>
<td>People who influence my behaviour think that I should use Second Life.</td>
</tr>
<tr>
<td>SN2</td>
<td>People who are important to me would think that I should use Second Life.</td>
</tr>
<tr>
<td>SN3</td>
<td>People whose opinion I value would prefer me to use Second Life rather than other 3D MUVEs.</td>
</tr>
<tr>
<td>SN4</td>
<td>I think that those people who are important to me would want me to use Second Life rather than other 3D MUVEs.</td>
</tr>
<tr>
<td>PEI1</td>
<td>When I am using Second Life, I feel &quot;carried off&quot; by the 3D virtual environment.</td>
</tr>
<tr>
<td>PEI2</td>
<td>When I am using Second Life, I feel as if I am part of the 3D virtual environment.</td>
</tr>
<tr>
<td>PEI3</td>
<td>When I am using Second Life, I feel deeply about the 3D virtual environment.</td>
</tr>
<tr>
<td>PEN1</td>
<td>I have fun using Second Life.</td>
</tr>
<tr>
<td>PEN2</td>
<td>Using Second Life provides me with a lot of enjoyment.</td>
</tr>
<tr>
<td>PEN3</td>
<td>I enjoy using Second Life.</td>
</tr>
<tr>
<td>PRP1</td>
<td>Using Second Life enables me to project myself into a particular role.</td>
</tr>
<tr>
<td>PRP2</td>
<td>Using Second Life enables me to project myself into a particular character.</td>
</tr>
<tr>
<td>PRP3</td>
<td>Using Second Life enables me to project myself into a particular task.</td>
</tr>
<tr>
<td>BI1</td>
<td>Assuming I had access to Second Life, I intend to use it.</td>
</tr>
<tr>
<td>BI2</td>
<td>Given that I had access to Second Life, I predict that I would use it.</td>
</tr>
<tr>
<td>BI3</td>
<td>I will use Second Life frequently in the future.</td>
</tr>
</tbody>
</table>
# GENDER AND WEBSITE DESIGN ACROSS CULTURES

| Journal: | 17th European Conference on Information Systems |
| Manuscript ID: | ECIS2009-0081.R1 |
| Submission Type: | Research Paper |
| Keyword: | Cross-cultural issues, Cross-national study, Human computer interaction (HCI), Interface design |
GENDER AND WEBSITE DESIGN ACROSS CULTURES

Abstract

Women and men process information differently and this phenomenon is confirmed in the context of online shopping where women are known to have diverse preferences from men related to website design and website trust. Yet despite these known differences, little research has examined gender and website design. Further, although online shoppers hail from all corners of the globe, no research has examined gender differences concerning website design in a multicultural sample. To fill this gap, the current investigation examined differences between men and women with a focus on website trust, website satisfaction, and website design in eight countries. Results demonstrate significant differences between men and women on all variables tested. Further analyses of four of the countries in the sample (China, Canada, the United States, and Germany) are compared related to socio-cultural values for masculinity-femininity. As expected, countries more alike on this dimension exhibited more similarities in terms of gender comparisons. All hypotheses in the study are supported. Suggestions are made regarding avenues for future research.

Keywords: Website Design, Website Trust, Website Satisfaction, Gender, Culture.
1 INTRODUCTION

The number of women shopping online is increasing, although relatively little is known about gender differences when buying on the Web. Prior research that does examine gender related attitudes and activities on the Internet suggests women are less interested in the Internet than men, spend less time online than men, and are less likely to purchase online (Allen 2001, Garbarino & Strahilevitz 2004, Rodgers & Harris 2003).

The design of the website is expected to impact user preferences, which in turn may produce differing reactions between men and women (Chen and Dhillon 2002). Although few studies have examined gender and website design - there are some exceptions. For instance previous research has examined gender differences related to online social presence and enjoyment (Cyr & Hassanein & Head & Ivanov 2007), website design and satisfaction (Cyr & Bonanni 2005), email usage (Gefen & Straub 1997), social norms (Venkatesh & Morris 2000), online risk (Garbarino & Strahilevitz 2004), or trust (Awad & Ragowsky 2008, Cyr & Bonanni 2005).

Building on previous research, the goal of the current investigation is to examine differences between men and women concerning their reactions to an online shopping website. The research variables considered include: website design elements (such as information design, information content, visual design, navigation design) website trust, and website satisfaction. To expand the generalizability of the findings, data was collected in eight countries (Canada, the United States, India, Germany, Japan, Mexico, Chile, and China) as well as separately for English and French Canadians within Canada. In addition, comparisons are made between four countries based on levels of masculinity-femininity (as elaborated by Hofstede 1984). Of interest is whether there is any similarity between user perceptions based on gender in countries which are similar on psychological-cultural values such as masculinity-femininity.

This paper begins with an introduction to gender and information processing, as well as cultural values related to gender with a focus on masculinity-femininity. This is followed by a review of the literature related to the various dependent variables, and with respect to prior research that examines the variables in the context of gender. The method, results, and discussion follow, with both theoretical and practical implications outlined.

2 INFORMATION PROCESSING AND GENDER

Men and women are known to process information differently, and this spills into various areas of information technology communications (Gefen & Geri & Paravastu 2007, Gefen & Ridings 2005) including online shopping. Dittmar et al. (2004, p. 440) write: “Men are more functional in their buying attitudes…whereas women stress social-experiential and identity-related concerns, and in particular, emotional involvement”. Rodgers and Harris (2003) noted that inadequately perceived emotional benefits may be an underlying reason why women are less involved in e-commerce activity.

Social explanations have been advanced with relevance for why men and women vary in their perceptions and expectations regarding online shopping. Men tend to vigorously pursue self focused goals having great personal consequences, while women are guided by communal concerns emphasizing interpersonal affiliation and harmonious relationships (Petrevu 2001). To better understand differences in information processing between men and women, two theories – selectivity interpretation and item-specific versus relational processing – are advanced. The selectivity hypothesis

---

1 It should be noted that the purpose of this paper is not to compare the research variables across countries but rather to focus on gender differences within each country or group with respect to the research variables. Separate research has been conducted by this author on website design across cultures. Please refer to Cyr (2008a).
(first proposed by Meyers-Levy 1989) asserts that men are selective processors who rely on highly available and salient cues. Women are comprehensive processors who are apt to assimilate all available information before arriving at a conclusion. Translated to an online shopping experience, men tend to pursue a minimizing approach whereby they make a selection as quickly as possible. Women, on the other hand, will spend considerably more time gathering information about products and comparing the merits of each prior to making the purchase decision.

Alternately, in item-specific information processing, attributes that are unique or distinctive to a message are important and there is a tendency to focus on a few salient attributes (more male oriented). When relational processing occurs (more female oriented), similarities, shared themes, or interrelationships among disparate pieces of information are sought (Einstein & Hunt 1980). In the context of online shopping these theories find relevance in that women require more social interaction than men in the shopping experience - and which may not be possible on the Web. Hence women perceive less emotional benefit from online shopping, and have less favorable perceptions of the experience than men (Cyr & Bonanni 2005, Simon 2001).

3 SOCIO-CULTURAL VALUES OF MASCULINITY-FEMININITY

In addition to different perceptions in online shopping related to gender, culture is known to have an effect as well (Cyr & Bonnani & Bowes & Ilsever 2005, Cyr 2008, Gefen & Heart 2006, Srite & Karahanna 2006). Differences in online communication strategies for target markets occur between Japan, Spain and the United States (Okayazaki & Rivas 2002). Over the years, researchers have often used Hofstede’s (1984) classifications to study social psychological phenomena including website design and experience (Gefen & Heart 2006, Jarvenpaa & Tractinsky & Saarinen & Vitale 1999). Simon 2001).

In this research, Hofstede’s (1984) classification for masculinity-femininity will be considered for its relatedness to psychological-social phenomena with respect to gender. Masculine values emphasize work goals such as material success and having challenging work. Feminine values are focused on quality of life, nurturing, and modesty (Hofstede & Associates 1998). As defined by Hofstede (1984) works goals include a focus on recognition, challenge, advancement, earning, and achievement defined by earnings. This set of goals, termed ego goals or ego enhancing goals, is thought to be associated with masculine cultures. Alternately, quality of life work goals emphasize a supportive and friendly work environment, cooperation, job security and achievement determined with respect to work relationships and human contacts. According to Hofstede, these social or relationship enhancing goals are associated with feminine cultures. Specific to the countries in this investigation, two countries that are ranked medium in masculinity (Canada - 52; China - 50) are compared; as are two countries that are ranked high in masculinity (the United States - 62; Germany - 62). As outlined in the introduction, it is of interest to see if in similar countries based on socio-cultural values of masculinity-femininity if there are any gender parallels for these countries.

4 WEBSITE DESIGN AND GENDER

Effective website design engages and attracts online consumers (Agarwal & Venkatesh 2002). Modes of information presented on the Internet, and the quality of graphics have a significant impact on user experience (Chau & Au & Tam 2000, Cyr & Head & Larios & Pan 2009). Others have conducted research into website design related to website trust, satisfaction and loyalty and have considered separate elements of information design, visual design, and navigation design (Cyr et al. 2005, Cyr 2008). In addition, Cyr (2008) included the additional design element of information content. Building on this earlier work, in this study the following are investigated:

Information Content - information that is complete, sufficient, and effective;
Information Design - information that is logically presented and organized;
Navigation Design - the navigational scheme or format used to help or hinder users as they access different sections of a website;

Visual Design - design elements such as balance, emotional appeal, aesthetics, and uniformity of a website’s overall graphical look; this may include colors, photographs, shapes, or font type.

Few studies have explicitly explored website design differences between men and women. Although in one investigation differences in website perceptions were examined for men and women concerning information richness, communication effectiveness, and communication interface (Simon 2001). Results from this study revealed that women had a lower perception of websites than men. In another study, Cyr and Bonanni (2005) found differences on selected items only between men and women for information design, navigation design and visual design. In total, five of 11 design items were statistically different using t-test comparisons. Based on these results as well as interview data, men felt the information on the website was better organized and presented, and they were more satisfied with navigation. With respect to visual design, women were more attracted by the colors on the website while men liked the more interactive and “flashy” design elements.

Further, based on four of the eight countries in the sample, it will be interesting to explore if there are similarities between how men and women perceive website design elements in a context that includes socio-cultural values of masculinity-femininity. More specifically, it is expected that countries that are both medium in masculinity such as Canada and China will have similar results for gender comparisons. Alternately, other countries such as the United States and Germany that are both high in masculinity will likewise be similar, and in turn different from Canada and China with respect to gender comparisons for website design.

Based on the preceding, the following hypotheses are posited:

**H1: There will be differences between men and women in their perception of information content.**

**H2: There will be differences between men and women in their perception of information design.**

**H3: There will be differences between men and women in their perception of navigation design.**

**H4: There will be differences between men and women in their perception of visual design.**

**H5: Canada and China (both medium masculinity) will be similar in gender results compared to the United States and Germany (both high masculinity), and alternately the United States and Germany will be more similar compared to Canada and China for information content, information design, navigation design, and visual design.**

5 TRUST, SATISFACTION AND GENDER

According to Jarvenpaa et al. (1999) online trust refers to consumer confidence in the website and a “willingness to rely on the seller and take actions in circumstances where such action makes the consumer vulnerable to the seller” (p. 4). Building on this definition, in the current research online trust refers to general trust of the website, and that the user can trust the transaction process as well as information presented. Consumer trust in the website is fundamental to loyalty including online purchase intentions (Flavián & Guinalíu & Gurrea 2006, Gefen 2000, McKnight & Kacmar & Choudhury 2004, Pavlou & Gefen 2005) and willingness by consumers to buy from an online vendor (Flavián et al. 2006, Laurn & Lin 2003, Pavlou 2003, Wang & Benbasat 2008). Overall, women are less likely to trust a website than men (Cyr & Bonanni 2005, Rodgers & Harris 2003).

It appears that perceptions of online trust are influenced by gender, although few studies examine this phenomenon. In this investigation, website trust was examined in terms of whether users feel they can trust the website viewed. This leads to the next hypothesis:

**H6: There will be differences between men and women for website trust.**
An effectively designed website may engage and attract online consumers resulting in satisfaction with an online vendor (Agarwal & Venkatesh 2002). Palmer (2002) validated design metrics for websites and found information content, site organization, and navigation are important to website success, including intent to return to the site. In other research, website design and the “ambience associated with the site itself and how it functions” is related to online satisfaction (Szymanski & Hise, p. 313). In the current investigation, website satisfaction refers to overall contentment with the online experience (Anderson & Srinivasan 2003, Flavián et al. 2006), and may include access to information, a positive navigation experience, and perception of a well designed website (Balasubramanian & Konana & Menon 2003). While few studies have examined online satisfaction and gender, there is evidence that women tend to be less satisfied with the online experience than men (Dittmar et al. 2004, Simon 2001). Further, men and women differ in their ‘web acceptance and usage processes’ (p. 19) with men more driven by instrumental factors with stronger perceptions of usefulness of the Web than women (Sanchez-Franco 2006).

This results in the following hypothesis:

**H7:** There will be differences between men and women for website satisfaction.

Further, to our best knowledge, there is no previous research in which perceptions of website trust and website satisfaction are tested with respect to cultural differences.

**H8:** Canada and China (both medium masculinity) will be similar in gender results compared to the United States and Germany (both high masculinity), and alternately the United States and Germany will be more similar compared to Canada and China for website trust and website satisfaction.

### 6 METHOD

#### 6.1 Participants

A total of 1156 participants located in English Canada (232: Female (F) = 124), French Canada (80: F=39), the United States (197: F=116), India (106; F=90), Germany (122; F=61), Japan (78; F=31), Mexico (71: F=41), Chile (48; F=37), and China (222; F=109) completed an experimental task and online survey. To ensure participants are “of the culture” it was determined each had lived in the country the majority of their lives and spoke the native language as their primary language. Participants were recruited from a wide range of sources including universities, institutes, and companies. Average age across countries is similar with an overall average of 27.4 years. Participants are experienced online shoppers with a mean of 3.8 years of online shopping experience. They are well educated with 273 having graduate education, 487 undergraduate education, 50 technical training, and 263 with high school completion. To determine if significant differences existed across cultures based on demographics, ANOVA tests were run for age, education, and Internet and online shopping experience. Overall, no differences occurred between cultures that would influence the constructs tested in this research.

#### 6.2 Task and Website Design

This research targets user impressions of B2C Web pages. For the research treatment participants responded to the local version of the SonyStyle website represented in their native language. The SonyStyle website was chosen after an extensive search for a well localized vendor website in which the design of the website was adapted to be appropriate to the culture of each user as determined by a design expert who rated each country website. To recruit participants, the researcher sent an email to international colleagues with a link where instructions for the research and an online survey were found. These colleagues further distributed the email to students and members of their respective organizations. Participants were requested to initially view the home page of the local website, followed by navigation of the website to choose a cell phone they would hypothetically purchase. This methodology is consistent with Cyr et al. (2005) and Cyr (2008a, 2008b). Once participants concluded
this task each completed an online questionnaire. Background information to the study, and all other written content including the questionnaire were translated and back-translated into each required language.

### 6.3 Instrument Validity and Reliability

Content validity ensures construct items are representative and drawn from a universal pool (Cronbach 1971). With the exception of Information Content, the constructs used in the investigation are based on previous research and have been previously validated from a variety of sources and exhibit content validity. More specifically: Information Design (Cyr 2008a, Egger 2001), Navigation Design, Visual Design (Cyr 2007, Cyr 2008a), Website Trust, Website Satisfaction (Cyr et al. 2005, 2007, 2008a). Items for Information Content were derived from existing literature including Garrett’s (2003) website classifications, and this construct was successfully validated in this investigation. All items were assessed on a 7-point Likert scale from strongly disagree to strongly agree. The questionnaire was pre-tested with 62 undergraduate students. Categories were evaluated for item validity and reliability and several items were revised for better fit. Table 1 shows the results of the principle component analysis with Varimax rotation.

<table>
<thead>
<tr>
<th>Principal component</th>
<th>InfoCont</th>
<th>InfoDes</th>
<th>NavDes</th>
<th>VisDes</th>
<th>Trust</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td>InfoCont1</td>
<td>.671</td>
<td>.206</td>
<td>.093</td>
<td>.000</td>
<td>.148</td>
<td>.136</td>
</tr>
<tr>
<td>InfoCont2</td>
<td>.814</td>
<td>.088</td>
<td>.184</td>
<td>.105</td>
<td>.141</td>
<td>.105</td>
</tr>
<tr>
<td>InfoCont3</td>
<td>.645</td>
<td>.303</td>
<td>.168</td>
<td>.110</td>
<td>.193</td>
<td>.114</td>
</tr>
<tr>
<td>InfoCont4</td>
<td>.791</td>
<td>.112</td>
<td>.129</td>
<td>.157</td>
<td>.018</td>
<td>.107</td>
</tr>
<tr>
<td>InfoDes1</td>
<td>.204</td>
<td>.748</td>
<td>.252</td>
<td>.148</td>
<td>.192</td>
<td>-.020</td>
</tr>
<tr>
<td>InfoDes2</td>
<td>.225</td>
<td>.668</td>
<td>.367</td>
<td>.220</td>
<td>.134</td>
<td>.139</td>
</tr>
<tr>
<td>InfoDes3</td>
<td>.327</td>
<td>.617</td>
<td>.278</td>
<td>.107</td>
<td>.100</td>
<td>.227</td>
</tr>
<tr>
<td>NavDes1</td>
<td>.156</td>
<td>.253</td>
<td>.796</td>
<td>.137</td>
<td>.124</td>
<td>.149</td>
</tr>
<tr>
<td>NavDes2</td>
<td>.179</td>
<td>.197</td>
<td>.816</td>
<td>.120</td>
<td>.196</td>
<td>.098</td>
</tr>
<tr>
<td>NavDes3</td>
<td>.214</td>
<td>.229</td>
<td>.696</td>
<td>.163</td>
<td>.200</td>
<td>.118</td>
</tr>
<tr>
<td>VisDes3</td>
<td>.215</td>
<td>.203</td>
<td>.270</td>
<td>.709</td>
<td>.215</td>
<td>.033</td>
</tr>
<tr>
<td>VisDes4</td>
<td>.114</td>
<td>.147</td>
<td>.130</td>
<td>.818</td>
<td>.109</td>
<td>.180</td>
</tr>
<tr>
<td>Trust1</td>
<td>.171</td>
<td>.107</td>
<td>.211</td>
<td>.195</td>
<td>.664</td>
<td>.063</td>
</tr>
<tr>
<td>Trust2</td>
<td>.143</td>
<td>.211</td>
<td>.170</td>
<td>.084</td>
<td>.781</td>
<td>.152</td>
</tr>
<tr>
<td>Trust3</td>
<td>.135</td>
<td>.095</td>
<td>.191</td>
<td>.111</td>
<td>.714</td>
<td>.182</td>
</tr>
<tr>
<td>Sat 2</td>
<td>.262</td>
<td>.194</td>
<td>.189</td>
<td>.173</td>
<td>.270</td>
<td>.706</td>
</tr>
<tr>
<td>Sat 3</td>
<td>.322</td>
<td>.083</td>
<td>.276</td>
<td>.141</td>
<td>.156</td>
<td>.665</td>
</tr>
<tr>
<td>α-value</td>
<td>.840</td>
<td>.817</td>
<td>.868</td>
<td>.714</td>
<td>.832</td>
<td>.800</td>
</tr>
<tr>
<td>AVE</td>
<td>.734</td>
<td>.679</td>
<td>.771</td>
<td>.765</td>
<td>.721</td>
<td>.689</td>
</tr>
</tbody>
</table>

NB: InfoCont = Information Content; InfoDes = Information Design; NavDes = Navigation Design; VisDes = Visual Design; Sat = Satisfaction

**Table 1. Factor Analysis**

---

2 The questionnaire is not included but may be obtained from the author.
Construct reliability was assessed using Cronbach’s $\alpha$-value. In Table 1 $\alpha$-values ranged from 0.714 (for Visual Design) to 0.868 (for Navigation Design). The Cronbach $\alpha$ of a scale should be greater than 0.5 for items used together and ideally higher than 0.7 (Rivard & Huff 1988). Therefore all constructs possess construct reliability. The average variance extracted (AVE) for a construct should exceed 0.5 (Fornell & Larcker 1981). In Table 1 this criterion is satisfied for all constructs, and the constructs used in this study exhibited satisfactory convergent validity (demonstrated by the principle component factor analysis) and satisfactory discriminant validity (shown from inter-construct correlation analysis).

7 RESULTS

A goal of the investigation is to determine if gender differences exist for an international sample of participants. Comparisons were conducted using Tukey HSD testing (N=1156). Results for gender differences for the mixed country sample appear in Table 2. For all dependent variables significant differences between men and women are noted. This serves to confirm Hypotheses 1, 2, 3, 4, 6 and 7.

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>F-Value</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Content</td>
<td>36.166</td>
<td>.000***</td>
</tr>
<tr>
<td>Information Design</td>
<td>62.839</td>
<td>.000***</td>
</tr>
<tr>
<td>Navigation Design</td>
<td>51.185</td>
<td>.000***</td>
</tr>
<tr>
<td>Visual Design</td>
<td>79.843</td>
<td>.000***</td>
</tr>
<tr>
<td>Trust</td>
<td>69.360</td>
<td>.000***</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>28.283</td>
<td>.000***</td>
</tr>
</tbody>
</table>

*** p <.001

Table 2. Gender Differences for Mixed Country Sample (ANOVA)

Means and standard deviations appear in Table 3. In all cases, men score higher than women.

<table>
<thead>
<tr>
<th>Construct</th>
<th>InfoCont</th>
<th>InfoDes</th>
<th>NavDes</th>
<th>VisDes</th>
<th>Trust</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
</tr>
<tr>
<td>Male</td>
<td>5.199 (1.132)</td>
<td>5.494 (1.023)</td>
<td>5.612 (1.066)</td>
<td>5.718 (1.150)</td>
<td>5.529 (0.996)</td>
<td>4.904 (1.282)</td>
</tr>
<tr>
<td>Female</td>
<td>4.761 (1.192)</td>
<td>4.927 (1.232)</td>
<td>5.070 (1.317)</td>
<td>4.999 (1.385)</td>
<td>4.955 (1.181)</td>
<td>4.478 (1.288)</td>
</tr>
<tr>
<td>Total</td>
<td>4.945 (1.186)</td>
<td>5.165 (1.182)</td>
<td>5.298 (1.247)</td>
<td>5.301 (1.339)</td>
<td>5.196 (1.142)</td>
<td>4.657 (1.302)</td>
</tr>
</tbody>
</table>

Table 3. Gender Differences for Mixed Country Sample [Sample N = 1156]

Comparisons for gender differences for China, English Canada, the United States and Germany appear in Table 4. As predicted, China and English Canada which both are medium for masculinity are more similar for the various dependent variables compared to the United States and Germany, which are both high for masculinity. This confirms Hypotheses 5 and 8.

Means and standard deviations appear in Table 5.
<table>
<thead>
<tr>
<th>Construct</th>
<th>China</th>
<th>English Canada</th>
<th>United States</th>
<th>Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F-Value</td>
<td>Sig.</td>
<td>F-Value</td>
<td>Sig.</td>
</tr>
<tr>
<td>Information Content</td>
<td>5.384</td>
<td>.021*</td>
<td>.699</td>
<td>.404</td>
</tr>
<tr>
<td>Information Design</td>
<td>1.497</td>
<td>.223</td>
<td>3.147</td>
<td>.077</td>
</tr>
<tr>
<td>Visual Design</td>
<td>3.300</td>
<td>.071</td>
<td>.003</td>
<td>.956</td>
</tr>
<tr>
<td>Trust</td>
<td>1.174</td>
<td>.280</td>
<td>.706</td>
<td>.402</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>.148</td>
<td>.701</td>
<td>5.056</td>
<td>.025*</td>
</tr>
</tbody>
</table>

* p < .05   ** p <.01   ***p <.001

Table 4. Gender Differences for China, English Canada, the United States and Germany

<table>
<thead>
<tr>
<th>Construct</th>
<th>InfoCont</th>
<th>InfoDes</th>
<th>NavDes</th>
<th>VisDes</th>
<th>Trust</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
</tr>
<tr>
<td>China [Sample N = 211]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>4.792 (0.872)</td>
<td>4.589 (1.003)</td>
<td>4.699 (1.155)</td>
<td>4.840 (1.247)</td>
<td>4.496 (1.044)</td>
<td>4.394 (1.252)</td>
</tr>
<tr>
<td>Female</td>
<td>4.468 (1.132)</td>
<td>4.399 (1.245)</td>
<td>4.565 (1.345)</td>
<td>4.524 (1.284)</td>
<td>4.325 (1.237)</td>
<td>4.329 (1.185)</td>
</tr>
<tr>
<td>Total</td>
<td>4.626 (1.024)</td>
<td>4.492 (1.135)</td>
<td>4.631 (1.255)</td>
<td>4.678 (1.273)</td>
<td>4.408 (1.148)</td>
<td>4.361 (1.216)</td>
</tr>
<tr>
<td>English Canada [Sample N = 231]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>4.975 (1.078)</td>
<td>5.190 (1.096)</td>
<td>5.370 (1.091)</td>
<td>5.606 (1.231)</td>
<td>5.429 (1.048)</td>
<td>4.555 (1.269)</td>
</tr>
<tr>
<td>Female</td>
<td>5.107 (1.293)</td>
<td>5.454 (1.158)</td>
<td>5.547 (1.182)</td>
<td>5.615 (1.314)</td>
<td>5.550 (1.133)</td>
<td>4.959 (1.443)</td>
</tr>
<tr>
<td>Total</td>
<td>5.045 (1.196)</td>
<td>5.329 (1.135)</td>
<td>5.464 (1.141)</td>
<td>5.611 (1.273)</td>
<td>5.493 (1.093)</td>
<td>4.769 (1.376)</td>
</tr>
<tr>
<td>United States of America [Sample N = 196]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5.432 (0.963)</td>
<td>5.725 (0.794)</td>
<td>5.821 (0.921)</td>
<td>5.950 (0.977)</td>
<td>5.525 (0.971)</td>
<td>5.088 (1.196)</td>
</tr>
<tr>
<td>Female</td>
<td>4.951 (1.284)</td>
<td>5.319 (1.221)</td>
<td>5.282 (1.334)</td>
<td>5.440 (1.398)</td>
<td>5.348 (1.156)</td>
<td>4.643 (1.425)</td>
</tr>
<tr>
<td>Total</td>
<td>5.147 (1.185)</td>
<td>5.485 (1.084)</td>
<td>5.502 (1.209)</td>
<td>5.648 (1.266)</td>
<td>5.421 (1.085)</td>
<td>4.824 (1.351)</td>
</tr>
<tr>
<td>Germany [Sample N = 120]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>4.263 (1.203)</td>
<td>4.767 (1.211)</td>
<td>5.017 (1.523)</td>
<td>4.892 (1.360)</td>
<td>4.878 (1.097)</td>
<td>4.159 (1.383)</td>
</tr>
<tr>
<td>Female</td>
<td>5.084 (1.168)</td>
<td>5.489 (1.190)</td>
<td>5.639 (1.085)</td>
<td>5.442 (1.303)</td>
<td>5.373 (1.257)</td>
<td>4.842 (1.145)</td>
</tr>
<tr>
<td>Total</td>
<td>4.673 (1.251)</td>
<td>5.128 (1.249)</td>
<td>5.328 (1.353)</td>
<td>5.167 (1.355)</td>
<td>5.125 (1.200)</td>
<td>4.500 (1.310)</td>
</tr>
</tbody>
</table>

Table 5. Gender Differences for China, English Canada, the United States, and Germany

It is interesting that in the medium masculinity countries where socio-psychological values would be more similar, there are few gender differences related to website design. For China, the only significant gender difference is for information content. Men score higher than women. For English Canada, the only significant difference is for satisfaction and rather surprising, women report greater satisfaction than men. The U.S. and Germany both have numerous significant differences by gender. For the U.S. 5 of 6 variables are significantly different, and in each case men score higher than women. For Germany, all variables are significantly different, and unexpectedly men score lower than women in all cases.
8 DISCUSSION

8.1 Theoretical and Practical Contributions

This study, based on a large sample of participants from eight country locations, serves to strongly reinforce the reality that men and women have a different experience of the same website. In the mixed country sample significant gender differences were found for all dependent variables. Further, the direction of the findings was as expected from previous research for website design (Cyr & Bonanni 2005, Simon 2001), trust (Cyr & Bonanni 2005), and satisfaction (Dittmar et al. 2004, Simon 2001). In all cases women score lower than men. Looking at each variable separately, mean scores are reasonably high for both men and women and range for men from 4.9 (out of 7) for satisfaction to 5.7 for visual design. Alternately, mean scores for women range from 4.5 for satisfaction to 5.1 for navigation design. While scores for women are lower, there are parallel patterns of preference for both men and women. This data provides insights into website characteristics that are most important for the user, and as such offers designers important information as to how to please online shoppers.

An interesting discovery from this research is that based on socio-cultural values associated with masculinity-femininity patterns emerge related to gender preferences between countries. This suggests that countries that are more similar on the masculinity-femininity dimension also score similarly in terms of gender differences for the dependent variables as tested in this research. It would appear that there are fewer differences between men and women in countries where masculinity is moderate (as in Canada and China) versus higher (as in the U.S. and Germany). Perhaps this can be explained by more similar socio-cultural values in those countries - regardless of gender. For Web designers this might feasibly suggest that for countries with similar cultural values, that the requirement for website localization (or adaptation) is less than for companies where cultural values are more discrepant.

8.2 Limitations of the Research

Data was collected in eight countries, including separate samples for English Canada and French Canada with a total of 1156 participants. A large and diverse sample population is a positive feature of this investigation. Participants are from a variety of sources including companies, universities, and institutes which offers generalizability of the findings. Concerning limitations to the research, a single vendor website (SonyStyle) was used. The possibility of response biasing could occur as respondents are aware of the company name and reputation, and may have previously established impressions of the company. In addition, a single task was used of searching for a cell phone for hypothetical purchase on a product-based website. No actual purchase was required. While this procedure is consistent with other e-commerce research, this may limit transferability of the findings to real e-commerce situations. In the future, similar research could be expanded to include a greater variety of tasks on both service and product websites, a larger sample of websites, or websites without specific branding.

8.3 Directions for Future Research

The findings for Germany are puzzling, with eight of nine variables significantly different between men and women and higher in each case for women. This finding merits additional investigation, perhaps using qualitative methodologies, to discover why exactly women are responding differently from their female counterparts in other countries. To date, there is no evidence to suggest why this result occurred.

Since the investigation of website preferences and the online shopping experiences between men and women is relatively new, there is much scope for future investigations in this field. Further research is suggested to more deeply probe online trust and satisfaction. In addition, differences in design preferences for men and women indicate this area offers multiple opportunities to expand understanding as to what is gender relevant when browsing or shopping online. Although website localization is usually considered in a cultural context, researchers and web designers might also want to consider the notion of gender localization – given the results of this study. With large numbers of
both men and women shopping online, vendors who create gender appropriate websites may reap
rewards from customers who are more satisfied and loyal.

References
procedure for measurement of usability. Information Management Research, 13(2), 168–121.
http://www.emarketer.com/analysis/eCommerce_b2c/20010228_b2c.html
shopping: An empirical evaluation of a broadband interactive shopping service. Journal of
Organizational Computing Electronic Commerce, 10(1), 1-22.
Cyr, D. (2008a). Modeling web site design across cultures: Relationships to trust, satisfaction, and e-
Proceedings for the Seventh Pre-ICIS HCI Research in MIS Workshop (HCI/MIS’07), Paris.
Electronic Business, 3(6), 565–582.
loyalty in e-service environments. Interacting with Computers, 19(1), 43-56.
conventional buying motivations. Sex Roles, 50(5/6), 423–444.
Einstein, G.O. and Hunt, R.R. (1980) Levels of processing and organization: Additive effects of
individual item and relational processing. Journal of Experimental Psychology: Human Learning
and Memory, 6(5), 588-598.
Egger, F. N. (2001) Affective design of e-commerce user interfaces: How to maximize perceived
trustworthiness,” in Proceedings of the International Conference on Affective Human Factors
Fornell, C. and Larcker, D. (1981). Evaluating structural equation models with unobserved variables
and the effects of receiving a site recommendation. Journal of Business Research, 57, 768-775.
Riders Publications, Indianapolis, IND.
Management Science, 28(6), 725–737.


CROSS-CONTEXTUAL USE OF INTEGRATED INFORMATION SYSTEMS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0528.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Enactment theory, Global information systems, interpretivist research, Standards</td>
</tr>
</tbody>
</table>
CROSS-CONTEXTUAL USE OF INTEGRATED INFORMATION SYSTEMS

Gasparas, Jarulaitis, Norwegian University of Science and Technology, Sem Sælands vei 7-9, 7491 Trondheim, Norway, Gasparas@idi.ntnu.no
Eric, Monteiro, Norwegian University of Science and Technology, Sem Sælands vei 7-9, 7491 Trondheim, Norway, Eric.Monteiro@idi.ntnu.no

Abstract

Large-scale organizations are increasingly promoting more collaborative and collective work practices across organizational boarders. A predominant way to achieve better collaboration in large-scale heterogeneous contexts is to establish an integrated and standardized technological infrastructure. Ethnographically inspired studies, on the other hand, have challenged such perspective and illustrated that generic technology does not fit in local contexts and needs to be worked-around. Similarly, this paper empirically exemplifies local workarounds and illustrates ongoing and persistently imperfect integration of a collaborative infrastructure in a global oil and gas company. More importantly, however, our analysis focuses on how integrated technology is used across contexts. We illustrate how local workarounds, as a result of tight technological integration, shape use patterns across contexts. Integrated systems establish interdependencies across contexts, thus, the use implies cross-contextual rather than local enactment. Since the trajectory of enactment is influenced by cross-contextual constrains, our study is addressing the existing overemphasis on studying/analysing the use of technology in isolated local contexts. Practically, our study suggests considering workarounds as an intrinsic part of every day work, which should be calculated as additional costs of making the generic technology to work in practice.

Keywords: workarounds, integrated systems, situated action, standardisation
1 INTRODUCTION

Abandoning earlier and overly structuralist accounts, there has been a steady increase in information systems research exploring contextual aspects of how technology is developed and used (Avgerou and Ciborra 2004). Ethnographically inspired studies have demonstrated beyond any reasonable doubt the situated nature of how information systems are appropriated (Orlikowski 1996; Robey and Sahay 1996; Walsham 2001). Theoretically, there has been an ‘agentic turn’, which has “led increasingly to the theoretical positions that privilege human agency over social structures and technological futures” (Boudreau and Robey 2005, p.3), for instance by advocating how technology is always ‘enacted’ (Orlikowski 2000). The locus of attention is local work practices and how technology is enacted in a situated context, where context is limited to individual actor’s engagement and “recurrent interaction with the technology at hand” (Orlikowski 2000, p. 47). Since, enactments more than often deviate from intended system design, a practical concern, then, relates to whether workarounds need to be eliminated (Azad and King 2008) or considered as an intrinsic part of every day work (Rolland and Monteiro 2002).

Integrated, collaborative systems (e.g. enterprise resource planning (ERP) systems, coordination technology (Lotus Notes, MS SharePoint), customer relationship management (CRM) systems) are attractive to business and public sector for their promise to promote more collaborative and collective work practices. Working more collectively across geographical, professional and organizational boundaries entails that one previously local, independent context of use gets linked with (i.e. becomes dependent on) other contexts. As opposed to largely local, independent contexts of enacted technology, use of integrated systems implies the interdependent enactment across the contexts now linked as a result of the integration.

The main purpose of this paper is to analyse the form and implications of cross-contextual enactment of integrated systems. We explore questions such as: how does one local workaround affect other contexts of use; how does local appropriation of technology ‘travel’ to other contexts mediated by the integration, possibly creating unintended consequences there?

The empirical basis of our paper is an ongoing, longitudinal (2007-2008) case study of a global oil and gas company (OGC, a pseudonym to maintain anonymity) where we also earlier (1997-1998) studied integrated, collaborative systems (Monteiro and Hepsø 2002). Operating across significant geographical, professional, business and organizational boundaries, OGC is struggling to move towards more collaborative modes of working. Integrated systems are a strategically recognised vehicle to address this challenge. Our study reports from an ongoing effort to deploy an integrated system based on Microsoft SharePoint (MSP) technologies1. We trace out local enactment (e.g. workarounds) of MSP, but more importantly demonstrate the cross-contextual nature of this enactment i.e. how workarounds in one context affect local appropriation of MSP in another context.

The structure of the remainder of this paper is organized as follows. In the next section we conceptualize the use of integrated information systems. Then, we outline our research approach and introduce historical context and intentions of changing collaborative infrastructure in OGC. Thereafter, we illustrate and discuss how local workarounds, as a result of the tight integration in MSP, shape use patterns across contexts. Finally, we provide analytical implications for studying the use of integrated IS and offer practical implications for managing generic infrastructures.

1 http://www.microsoft.com/SharePoint/default.mspx
CONCEPTUALISING THE USE OF INTEGRATED SYSTEMS

2.1 Using technology in a situated context

Mirroring a more general interest in the social sciences for practice theory (Gherardi 2000; Savigny, Knorr-Cetina et al. 2001), information systems research has for some years studied its principal ‘practice’ viz. the practices that go into the use of information systems. Analyses of how users perceive, appropriate and subsequently use information systems demonstrate the highly contextual or situated nature of the use (or practice of use, if you want) of technology.

An early and influential contribution was Gasser’s (1986) study of users’ strategies of fitting, augmenting and working around the intentions inscribed into the functionality of the system. Gasser (ibid.) empirically illustrated what people actually do when confronted with rigid and unreliable computing procedures. The author vividly illustrated that users in fact do not use information systems as they are designed, but invent various ad-hoc strategies to fit the technology for a particular task. The major conceptualization from this study was the notion of workaround, which refers to “using computing in ways for which it was not designed or avoiding its use and relying on an alternative means of accomplishing work” (ibid., p.216).

While Gasser (1986) studied the rigid inventory control system, other more flexible types of systems were studied as well. For instance Orlikowski (1996) investigated how quite small customer support department used a new system to provide a better service for customers. Orlikowski (ibid.) illustrated how users contingently appropriate technology over the time. The central characteristic of appropriation is continuous change with unpredictable trajectory (‘improvisation’) rather than stability.

Central to these studies is to understand how technology is used in a situated context. While the notion of context is certainly vague (Chalmers 2004), context, in this case, is limited to individual actor’s engagement and “recurrent interaction with the technology at hand” (Orlikowski 2000, p. 47).

2.2 Integrated systems

Several scholars have been interested in how integrated systems are used (‘enacted’). For instance, Boudreau et al. (2005) in their recent study showed how, despite inherent rigidity of an ERP system, users are working-around the system in unintended ways. They argue that users first avoid the system (due to inertia), later learn by improvising (rather than in formal training) and finally reinvent the system in not-planned ways. Thus, the authors emphasize the human agency perspective over the technological logic and argue that “technology’s consequences for organizations are enacted in use rather than embedded in technical features” (ibid. p.14). Other researchers have similarly emphasized the impossibility of large-scale systems to be universal across contexts due to local relevance or cultural fit (Joshi, Barrett et al. 2007). Some researchers have suggested that workaround is an intrinsic part of every day work rather than negative or unwanted effect (Rolland and Monteiro 2002).

In general, it was suggested that technologies would always drift from initial plans due to the improvisational capability of a human actor. In turn, the same technology can produce contrasting effects in similar organizational contexts (Robey and Sahay 1996) and these should be addressed employing the logic of opposition (Robey and Boudreau 1999).

Overall, studies on IS use tend to overemphasize local practices and do not “adequately address the longer-term co-evolution of artefacts and their social settings of use” (Pollock, Williams et al. 2007, p.257). Certainly, the relationship between contexts and spanning-effects are discussed by several scholars (Hanseth, Ciborra et al. 2001; Ellingsen and Monteiro 2006), however, there are few if any studies which conceptualize how local workaround does influence other contexts of use. For instance Boudreau et al. (2005) do identify the relationship between local appropriations: “An error occurring
at one level of the system would have a ripple effect at other levels” (ibid., p.13), however they do not elaborate on the issue, nor do the authors elaborate what is the role of local enactments in larger contexts: “one cannot categorically argue that unintended actions are good, any more than one can argue that they are bad. Like any other aspect of organizational behaviour, evaluations of effectiveness are relative, not absolute” (ibid., p.16).

To sum it up, the study of integrated information systems has been framed to date largely along the lines of practice theory in the sense that local strategies for appropriation and use have been highlighted. Our study supplements this with a more systematic attention to the interdependence of cross-contextual appropriation mediated by integrated systems.

3 METHOD

We report from an ongoing longitudinal research project started in January 2007. Our research approach can be conceptualized as an interpretive case study (Walsham 1993) as we “attempt to understand phenomena through the meanings that people assign to them” (Klein and Myers 1999, p.69).

Data collection activities started at the beginning of 2007 with the primary aim to explore the change associated with the implementation of MS SharePoint technologies. We have employed 3 modes of data gathering: informal and formal interviews, observation and document studies.

We have conducted 25 in-depth interviews, on average lasting 1-1.5 hours. First interviews were open ended and aimed to identify IT strategic visions and implementation activities related to MS SharePoint. During later interviews, we analysed specific infrastructural components, work practices or individual engagements with technology. The technological complexity and intentions behind the new infrastructure were discussed with developers, administrators and managers of the collaborative infrastructure. The use of collaborative infrastructure was explored with actors from several organizational units. Interviewed users represent such disciplines as technology managers, human resources, senior researchers and various engineers involved in oil and gas production activities.

Participatory observations and informal discussions were mainly carried out in one of the OGC research centres, where both authors were granted access since the beginning of data collection. Since January 2008, one of the authors has been granted an office space, an access badge and access to OGC IT network. Since then, the researcher has been spending 2-3 working days a week in the research centre. Significant amount of time spent on-site forms the understanding of how work is carried out in practice and what problems and frustrations users experience on a daily basis. Additionally, being on-site gives an opportunity to have informal but informative chats around a coffee machine or during lunch breaks.

The third major empirical data source is internal OGC documents. We have extensively studied strategic documents related to planning and implementation activities of MSP. Additionally, we analysed technical descriptions, formal presentations and training materials on various MSP infrastructural components. A number of policy documents, which define how particular technology should be used or how specific work has to be carried out, were studied in detail. Finally, OGC intranet portal provided extensive contextual information on diverse OGC activities.

Data analysis is ongoing and iterative. Considering changing researchers involvement and overlapping but not the same research focus, the analysis of empirical data has many trajectories. This difference gives us a unique opportunity to analyse implementation process from slightly different perspectives. It is quite often that after interviews, if conducted together, we have a discussion and analyse what new aspects we have uncovered or what needs more attention in the subsequent data collection steps. In our faculty, there are several actors (not only the authors of this paper) exploring MSP implementation activities in OGC. We meet and discuss quite often either around a coffee machine or having more formal discussion sessions. Considering that the authors of the paper are involved
researchers, significant part of data analysis and validation is actually occurring with the help of OGC actors. During informal or formal meetings, we frequently present our findings to various OGC actors. In turn, we are challenged, supported or directed to the issues that need more attention. For instance, several record’s information managers supported our early findings on the metadata use in research and development activities, but we received extensive comments and suggestions to collect more empirical data in operative environments. Adjustments to some generalizations were made and empirical data collection directions were embraced.

In general, empirical data is classified in broad themes reflecting specific organizational project, practice or technical component. Such classification is neither all encompassing nor exhaustive; it is rather overlapping and continually changing. Theory has an important role in the analysis process. It provides an analytical lens to sort out and reclassify empirical data. For instance, in relation to this paper, the concept of workaround implied to determine when is a workaround (in relation to formal policy) and classify empirical data according to why, where and how workarounds are practiced. The concept of ‘generification’ implied to analyse how local workarounds ‘travel’ across contexts and what effects they produce.

4 CASE: COLLABORATION AND INTEGRATION

4.1 Context and history

Established only in the 1970s, the global oil and gas company (OGC, a pseudonym) has grown from a small, regional operator in Northern Europe to a significant energy company, currently employing some 30,000 people with activities in about 40 countries across 4 continents. OGC has grown largely organically, but with selected, important national and international acquisitions. Facing limited growth potential in its region of origin, OGC is actively pursuing a strategy to grow globally. To boost its financial capacity and flexibility, in the 1990s OGC diversified and expanded its shareholder ownership including getting listed at the New York Stock Exchange.

Alongside its growth in size, geography and business areas, OGC has been engaged in a number of corporate-wide initiatives to improve communication and collaboration. These initiatives have relied heavily on information systems. The first comprehensive effort to establish a corporate, collaborative information systems infrastructure was in the early 1990s (Monteiro and Hepsø 2002), at a time of oil industry recession, falling oil prices and dollar rates. Centralization, standardization and market orientation of IT services was the direct outcome of several projects whose primary aim was to solve the problems of fragmented and incompatible IT solutions. The outcome of standardization activities led to the establishment of the Lotus Notes-based collaborative infrastructure.

The Lotus Notes based infrastructure has proven successful inasmuch as it has been widely used for a range of different purposes. A key vehicle for facilitating collaboration within projects in OGC has been Lotus Notes Arena (Arena for short) databases for collective storing and dissemination of documents. The challenge, however, with the Lotus Notes based infrastructure has been to promote communication across the project-defined boundaries of the Arena databases. The Arena databases had no central indexing functionality, meaning that it was impossible to retrieve a document by searching if one did not know which database to search. With Arena databases thriving apparently ‘out of control’ – there were some 5000 databases by the latest estimates – locating relevant information stored outside your immediate project scope was non-trivial. Each user had in addition access to both personal (G disc) and departmental storage (F disc) areas. In short, information was scattered and duplicated over many local storage arenas.
4.2 New collaborative strategy – higher efficiency with tighter integration

To overcome the problems with Lotus Notes and establish more effective ways of collaboration, coordination and experience transfer, OGC formulated a new strategy in 2001. According to this strategy, OGC already had a set of general collaboration tools, but “these tools are poorly integrated”, and “there is a particular need for better and more integrated coordination tools, better search functionality and improved possibilities for sharing information with external partners” (internal strategy documents). The change in the collaborative infrastructure was defined as a necessity and catalyst in order to achieve goals formulated in the strategy. The decision was made in 2003 and the rollout of a new infrastructure based on the Microsoft SharePoint (MSP) started. MSP was selected exactly for its potential to overcome the fragmentation resulting from project-specific Arena databases. Recent accounting regulation in the aftermath of Enron added pressure to ensure more systematic and consistent documentation of business decisions to inform the stock market and the public.

MSP is a core element in the new OGC collaborative infrastructure. The central element of MSP is so-called team site (TS), the virtual arena for collaboration. TS provides functionality to check-in and check-out documents, post announcements, share links and create discussion boards. Another important element of TS is a so-called workspace. A workspace is a web site connected to a TS (sometimes called baby-team site), used for production and sharing of a specific document or collection of documents. While MSP is mainly used for documents management, the technology is integrated with a corporate-wide search engine, an archive system and MS Exchange system.

While the technology itself (MSP) is customizable for specific contexts, the OGC decided to make the solution as generic as possible so that it would fit all contexts (internally it is referred as one-size-fits-all strategy). The strategic choice to rollout ‘out-of-the-box’ solution with minimum customization was highly influenced by the previous implementation experiences. In particular, the straightforward MSP implementation process was planned in contrast to recent experiences with an opposite (extensive customization) strategy when implementing a several hundred million dollars worth corporate-wide ERP solution. These experiences were translated in the standardization of both the functionality and the interface of every TS. The only element that differentiates team sites is metadata. The metadata standard provides a common and standardized classification scheme on how the information should be classified. Thus, the metadata can be seen as the main element in the collaborative infrastructure, which should fit a generic TS to a specific local situation. The metadata standard represents quite complex classification scheme with 13 different ‘elements’ and corresponding ‘sub-elements’. In total there are more than 120 sub-elements in the metadata standard. Taking into account all sub-elements, the standard describes “identity, authenticity, content, structure context and essential management requirements of information objects” (OGC internal).

5 ANALYSIS: CONTEXTUAL AND CROSS-CONTEXTUAL USE

What studies of contextual use of information systems have convincingly established is the importance, indeed, necessity, of users’ active appropriation of technology to local circumstances and concerns. In other words, local workarounds (or appropriation, tweaking, improvisation, drift etc.) are not anomalies or design shortcomings but constitutive elements of working technologies (Rolland and Monteiro 2002).

Zooming in on two extended illustrations from OGC’s implementation of MSP (the use of workspaces and classification of documents), we first reiterate this point. More importantly, however, we go on to identify how these local workarounds – as a result of the tight integration in MSP – shape use patterns in other contexts of use. Table 1 summarizes this. One way, then, to formulate the gist of our analysis is to say that the local strategies of appropriation, the prerequisite of working information systems, are
simultaneously non-local side-effects significantly influencing patterns of use in other contexts (i.e. cross-contextual).

<table>
<thead>
<tr>
<th></th>
<th>Local context: appropriation</th>
<th>Cross-contextual: side-effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>The use of workspaces</td>
<td>Reproducing folder structure from entrenched practices</td>
<td>Invisibility of information due to the incorrect information and access rights management</td>
</tr>
<tr>
<td>Classification of documents</td>
<td>Overriding default values to fit local context</td>
<td>Undermining the search engine and distributing additional work</td>
</tr>
<tr>
<td></td>
<td>Incorrect, but convenient classification of sensitive information</td>
<td>Availability of sensitive information and suspension of the search engine for 5 months</td>
</tr>
</tbody>
</table>

Table 1. A summary of local (contextual) and non-local (cross-contextual) enactments of integrated information for two examples within the MSP infrastructure.

5.1 Workspaces: local appropriation as replicating existing practices

Prior to the implementation of MS SharePoint an important decision not to have folder structure in team sites was made. This decision was made due to the technical inability to create a complex folder structure in team sites (limitations of the URL length). On the other hand, users in OGC had quite long experience with folder structure in Lotus Notes infrastructure and indeed as Boudreau et al. (2005) explain did have difficulties to ‘forget’ previous practices. In the initial stages of MS SharePoint implementation users tended to avoid the new system and used file servers to share information instead. As one manager explained, the amount of documents in file servers exploded when team sites were introduced. However, after some time users got acknowledged with the system and found out that it is possible to replicate previously existed folder structure with the help of workspaces:

“From the beginning it was very clearly communicated that we are not allowed to use workspaces to replicate folder structure. That was the intention… however people have been using computers here for the last 15 years, so they actually continue to make folders with workspaces despite the fact that they are told no to do so… All our team sites have a pile of workspaces [the user navigates to one of the team sites and shows a list of approximately 30 workspaces]” (User working with operational support for offshore activities)

Overall, replicating folder structure with workspaces is quite popular workaround in OGC. In particular, the workaround is practiced in information-rich contexts. The ‘popularity’ of this workaround illustrates that users do not adhere to OGC policies but are actively engaging and experimenting with technology. The existence of such workaround is not surprising; it can be explained with the concept of ‘installed base’, used by Boudreau and Robey (2005). Such treatment of workaround requires shifting the focus from identification to explanation of the effects a workaround is producing. Precisely because the MSP infrastructure is integrated, local appropriations, outlined above, are not only local; they have cross-contextual implications.

The initial intention of using workspaces was related to the possibility for team site users to create ‘areas’ (i.e. workspaces) with custom access rights. In that sense, a workspace was considered as a temporary ‘arena’ in order to limit or expand the original access rights in a team site. However, in practice, users sometimes create folder-like workspace structures and use the functionality of limiting the access. Importantly, in its current configuration, corporate-wide search engine is only returning those documents that a user has access to. Thus, documents stored in workspaces with limited access rights will be visible and retrievable only to specified users and ‘invisible’ to others.

Cross-contextual effects are especially experienced by users working across contexts i.e. whose ability to find information depends not only on their skills, but also on how others manage information.
locally. For instance, a well engineer responsible for conducting well interventions across different fields explains:

“It is quite often that we do not have access to necessary information. When planning a well intervention we have to know a lot of technical information about a particular well and history of the well in general [this includes information about previous challenges/problems, and conducted interventions with corresponding experience reports produced after each intervention]. Sometimes you do not find information just because you do not have access… so you have to call various people and ask… it is very time consuming and I know some people do not bother spending all their time on that… however, not having important information means more uncertainty during operation, and this can increase the risk and cost of operation.” (Well engineer; emphasis added)

Oil and gas exploration, production and export activities span across many disciplines not only in the OGC, but a significant part of activities is carried out by various external contractors. For instance, while a plan to drill a new well is primarily produced by several internal disciplines, an external contractor can perform drilling. This means that for a certain period of time an external contractor needs access to information related to the new drilling activities and probably to some historical reports. It adds complexity to access management, and workarounds made some time ago tend to pop-up here:

“It is quite often that I get a call asking for help to find information or to give access. So I have to use a lot of my time on this… I would like them [contractors] to be more independent … to avoid this [access problems], for instance, after a meeting with contractors I am sending two emails, one to internal OGC employees with a link to a document and another one to external ones with attachment.” (Drilling engineer; emphasis added)

Many users are aware of cross-contextual effects of incorrect access rights management by explaining that they do not know whether a specific information is existing or not: “the worst thing is that if you don’t find information it does not necessarily mean that information is non-existing” (Engineer). Such effects lead to uncertainty and distrust the capabilities of the search engine. In general, these examples show that local workarounds change use patterns across contexts. In that sense, working in an ‘integrated environment’ means shifting the focus from how the system fits locally to how the system fits across contexts.

5.2 Classification of documents: the power of the default value

The decision to have flat document structure did not fuel too much enthusiasm for end-users and, as argued in the previous section, invoked workarounds. By removing folder structure the implementation team understood that some alternative way to classify information should be provided. In turn, it was decided to develop a common predefined classification scheme, which would form the basis for information structure and help both to sort and retrieve information.

As we have argued in the previous section that thinking ‘flat’ (no folders) introduces some problems, common classification scheme also did provide challenges for end users. The notion of ‘common’ does not imply fits-all or having the same meaning across contexts, on the contrary, as Star (1991, p.44) explains, “no networks are stabilized or standardized for everyone”. Thus, in some contexts classification is not acknowledged:

“this metadata, it is bad… very often when you will store a document none of the provided values fit. For instance for this document I can choose from 10-12 different values… but they all do not fit… for this document I can select such values as ‘none’, ‘agenda’, ‘minutes’, ‘presentation’… this document is presentation so ‘presentation’ value fits very well [the respondent starts laughing]” (Engineer working with oil and gas production; emphasis added)
While users are quite often confronted with metadata that does not fit in their contexts, the questions remain how they cope with this situation. Sometimes users just tend to ignore the existence of metadata and use such values as ‘none’ or ‘miscellaneous’. It was planned to have a controlled vocabulary (the values are predefined in advanced) and all policies state that users should use provided metadata. Interestingly enough, MSP functionality allows to delete provided metadata values and create new ones. In turn, the second enactment strategy is to develop new values that would make more sense in local context:

“We have replaced provided metadata values with the new ones, which actually represent the activities we are working on in the project. Provided values were meaningless in relation to this project, so it would make no sense to use them” (User working in R&D)

Another challenging issue related to classification in OGC is the classification of sensitive information. Users are provided with rather simple security classification scheme to identify which documents can be available to anyone (open), to all OGC employees (internal), to specific groups (restricted distribution) or to selected individuals (confidential). Security classification is managed in team site and on the workspace level, implying that confidential document should be placed in a confidential team site rather than in an open one. The security classification scheme can be described as simple and intuitive, however, in practice, the definition of what is ‘confidential’ and how it should be handled (for the sake of convenience or additional work) is not given.

In turn, local enactments diverge from formal policies. It was, up to now, quite a ‘standard’ to store personal information on the private team site, which is not classified as confidential. In some instances, due to convenience reasons, a classified report from an external company was stored in an ‘open’ projects team site, rather than in a workspace with restricted access. Perhaps the most ‘problematic’ enactment of security classification was to store, due to unawareness and additional work, human resources related information in not confidential team sites.

Overall, classification, being an inseparable part of everyday work, is not given but has to be enacted in practice (Bowker and Star 1999). The problematic aspect, as we have illustrated above, is that imposed common classification has local variations (workarounds). While the problem of global standardization and local variation has received quite some attention (Star 1991), it is much less clear how local variations ‘travel’ across contexts. Essentially, common (shared and used across contexts) classification can be characterized as inherently having cross-contextual aspects. Such aspects are captured with writing/reading metaphor: “any reading and writing artefact that accumulates inscriptions cannot but coordinate [constrain and transform] the activities which write and which read these inscriptions” (Berg 1999, p.391). Thus, working-around common classification not only erodes the common and controlled character of classification, but also automatically imposes a certain amount of additional work for actors across contexts.

Such cross-contextual effects can be nicely illustrated with an example of planning and drilling a new well. While planning activities are conducted onshore, drilling process is to a large extent managed offshore. Planning is a collective effort of various disciplines and results in producing several documents. Central documents describe the whole drilling program, detailed drilling procedures, possible risks and a checklist, just to mention some. Drilling a new well means producing at least some 200 documents, which have to be stored in team site(s). Since they are stored in a flat structure, metadata is the primary sorting and filtering mechanisms. The problem according to one engineer is that “while the metadata values are not very bad, people sometimes do not use them or use them wrong”. In turn, incorrect use of metadata in one context, produce effects in other contexts:

“sometimes I get a call in the evening from offshore people saying that they have been searching for a specific document for an hour or so with no success… to avoid this we have developed a practice [which is unofficial i.e. a workaround] that for every new drilling program, a drilling engineer [working onshore] creates an excel document containing links to documents that are most important for drilling engineers working offshore. It is additional work as we [engineers working onshore] have to update those excel documents during
drilling, but then offshore people have much better overview.” (Drilling engineer working onshore; emphasis added).

In that sense local workaround (wrong classification of documents) produce additional work in other contexts and later trigger other workarounds (engineers developing an excel sheet with links). Thus, local workarounds can produce ripple effects. More importantly, achieving working infrastructure entails collective enactment across contexts rather than local enactment.

Another, more substantial cross-contextual effect, was enacted with security classification. Documents were not always classified as intended. Incorrect classification of documents, due to the tight integration with the corporate search engine, made possible for confidential information to be available to many more than it should be. For instance, during an interview with two system administrators responsible for technical aspects of MS SharePoint infrastructure, we were shown the possibilities of the search engine. One administrator entered the name of his colleague sitting besides to demonstrate how the search engine works in practice. Among the first results, a document containing the administrator’s work evaluation appeared. It was an embarrassing moment, for administrators in particular. Such document contains personal information, and it should be available only to few persons and certainly should have not been retrieved in that situation. Similar local enactments not only propagated across contexts but accumulated as well. The situation, according to one manager “got out of control, since too much sensitive information due to incorrect classification was available for way too many users”. Since it was not impossible to apply any ‘quick fix’, the search engine was suspended. The corporate-wide search service was not available to users for 5 months, the period that was used to ‘clean-up’ incorrect classification and develop the approach that would prohibit such incorrect enactments later. Compulsory training sessions, technical usability improvements, control routines and other initiatives are currently executed to prevent such effects. However, as it is now acknowledged in some management levels, order without workarounds is out of reach: “we may hope for altered attitudes and more care taken in the future – however, all the time search has been suspended, people have been working as before (but all the errors have been “invisible” as search was not available)” (a recent presentation on information security in OGC).

6 IMPLICATIONS AND CONCLUSIONS

We draw two sets of implications from our study of the deployment of a MS SharePoint based information infrastructure in OGC: one analytic and one practical.

Analytically, our study demonstrates the rich array of strategies and improvisational acts that go into the local appropriation of technology. Existing research (see above) has vividly illustrated process of social shaping of technology, i.e. how both flexible and rigid technologies are shaped in situated contexts. In line with the ‘practice perspective’ of Orlikowski (2000), our findings confirm that local workarounds, tinkering and ‘situated improvisations’ are not anomalies or design shortcomings but constitutive elements of working technologies.

More importantly, however, our study continues to address the nature of non-local – what we in this paper have dubbed cross-contextual – effects that are embedded in the appropriation of integrated systems. For sure, the mere existence of cross-contextual effects has been acknowledged before. Boudreau and Robey (2005) for instance point out that local practices may have ‘ripple effect’ beyond the local context. Yet they fail to develop this observation into a more systematic framework or make it subject to substantial theorising (see Hanseth et al. (2006) for an exception). One implication of our study, then, is to contribute to a higher visibility of cross-contextual effects of the use of integrated information systems. Systematic attention to cross-contextual side-effects extend the way the use of integrated information systems has been conceptualized to date. As opposed to largely local independent contexts of enacted technology, the use of integrated systems implies the interdependent enactment across contexts now linked as a result of integration. This entails considering the technological, and in particular integrative technological, detail more seriously than the previously
outlined ‘practice lens’ (Orlikowski 2000) or the application of the ‘practice lens’ in integrative environments (Boudreau et al. 2005). In this way, our study is addressing the existing overemphasis on human agency and contributes to the studies on mutual shaping of information technology and its use, to approach equally important question on how technologies constrain the trajectory of enactment (Kallinikos 2004; Doherty, Coombs et al. 2006). The empirical illustration above of how the accumulation of multiple, local appropriations added up to the effect of closing down the corporate-wide search engine service for almost half a year is difficult to arrive at with attention largely focused on local, contextual or situated ‘enactment’ of technology.

Practical implications of our study relate to the (project) management of embarking on large-scale, comprehensive infrastructure like efforts of the type reported here. At the core, the insight that due to the scale and heterogeneity of work practices and existing technological interdependencies, workarounds need to be considered as constitutive elements of working infrastructures rather than anomalies, design shortcomings or unexpected effects. More importantly, working-around integrated systems cannot be any longer considered as local phenomena, which could be black-boxed and ascribed to a specific context. Prototyping early versions of integrative technology, for instance, has more limited value than for non-integrated technologies. While the attractiveness of integrated technology is based on seamless cross-contextual information exchange, such technological platform also comes (quite often as a surprise) with inherent cross-contextual effects. Simply put integrative technology is more than often evaluated from a perspective of what positive effects it can bring, while underestimating how local (small and perhaps unimportant at first) activities can produce great effects some time later across contexts. Essentially, the benefits of integrated systems can only be realized if they are fitted across contexts rather than in some local contexts.

The second practical implication relates to the evaluation of cost and benefits (Goodhue, Wybo et al. 1992) of generic, corporate-wide integrative infrastructures. We consider the term evaluation from an interpretive perspective (Walsham 1993, p.165-186) rather than an economic one. ‘Costs’, include both, the developer’s effort to establish technological platform and users adjustments of technology to his/her needs. Generic solutions are adjustable and can fit quite well in some, but most often not in all contexts. If generic technology does not fit, additional work (‘costs’) has to be carried out. Classification is an illustrative example of this. We have exemplified how locally irrelevant classification will require additional users work to make it meaningful locally. The same happens with generic technological functionality. For instance, some users in OGC, unsatisfied with MSP functionality, voluntarily and not in accordance with existing policies, invest their time in implementing, learning and using more flexible technologies (such as Groove) or social software solutions (Wiki’s). One way to evaluate then is to consider, by percentage, to how many actors the generic technology does fit. A qualitative alternative is to consider whether generic technology fits well in specific (not excluding core) business activities. In our evaluation, we are employing the latter perspective, and have illustrated throughout the paper that both development and use costs are high, and we doubt whether they outweigh the benefits. As the IS management literature suggests, the implementation of new technology should be cost-efficient. In turn, the management should take into account not only the costs of establishing integrated technical infrastructure, but the invisible (i.e. workarounds) costs as well. In that sense, it should be made explicit who, how much and when will pay the costs of having the generic integrated infrastructure.

REFERENCES


Does Time Matter? The role of ICT in shaping temporal assumptions

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0614.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Organisational culture, Organisational Change, IT/IS management, Case Study</td>
</tr>
</tbody>
</table>
DOES TIME MATTER?
The Role of ICT in Shaping Temporal Assumptions

Abstract.
Despite its importance to temporal issues, research into the temporal impacts of information technology in organizations is still limited. On the other hand, organizational culture research shows that the way time is perceived and collectively organized reflects assumptions that are an expression of the specific organizational setting, underscoring that cultural assumptions are an important contributory factor to the strength and direction of organizational change. In this contribution, we investigate the role ICT can play in promoting changes in the temporal dimension of organizational culture, and seek to assess whether temporal assumptions can affect the way a new system is used, thus facilitating/hindering the achievement of the expected results. This research-in-progress paper presents some preliminary results of a multiple case study, showing what types of ‘temporal performance’ management expects to see thanks to the introduction of a workflow system, and showing that managers have different levels of awareness about potential temporal performance of the system. Preliminary results also seem to confirm that, after the introduction of the workflow, some temporal dimensions of organizational culture of the departments involved present some significant changes, which confirm hypothesis 1 of the study, but also some contradictory effects that seem to confirm hypothesis 2.
Introduction
It is generally accepted that, when implemented in organizations, information technology hugely speeds up business processes, thereby enabling the adopting organizations to save a great deal of time. Despite its significance in temporal terms, research on the time impact of information technology in organizations is still limited (Lee and Whitley 2002).

Empirical studies on this topic began to appear in the last decade in the journals of both Organization and IS discipline (Sahay 1998; Lee 1999; Lee and Liebenau 2000; Orlikowski and Yates 2002; Sawyer and Southwick 2002; Scott and Wagner 2003; Sarker and Sahay 2004; Prasopolou et al. 2006). Accelerating the pace of work processes, fostering workers’ polychronicity, promoting shifts from “batch” logic to “flow” logic, and improving synchronization among organizational units are just a few examples of some key objectives that can be pursued by firms when they adopt systems like Workflow or ERP. Such objectives, which we will refer to as expected temporal performances, are related to potential changes in the temporal organization of processes and activities, but also to potential changes in people’s assumptions on time and time use in the workplace.

What do we mean by assumptions on time?
Organization studies have long acknowledged time as a fundamental dimension of organizational culture (Hofstede 1991); Schein as well, in his study on organizational culture (1988), maintains that organizational culture has groups’ dominant assumptions of time and space embedded within it and describes a variety of “assumptions on time” that characterize different organizational cultures. Schriber and Gutek (1987), in their empirical research, also described rules about time as specific “dimensions of organizational culture” and found that these cultural features varied across organizations and working groups. It has been underscored that, in addition to being a condition for the coordination of activities and the production of organizational outputs, the temporal organization of work processes constitutes a framework that plays an essential role also as a template for organizing behavior: a cognitive and cultural framework that defines the activities and routines of both people and organizational units, used by people to make sense of actions and events in the workplace (Barley, 1988). In his study on temporal patterns in the organization of activities in hospitals, Zerubavel (1979) found that various types of schedules worked as “cognitive maps” used by personnel to provide a background, a “repertoire of what is expected, likely or unlikely to occur within certain temporal boundaries” [p.125]. Temporal assumptions thus represent an expression of the specific organizational and professional culture that produces them, conveying a symbolic value for the individual and the group of workers. Finally, assumptions on time can also be an expression of the specific sub-culture of different departments and professional groups within a firm (Lawrence&Lorsch 1967; Gherardi and Strati 1988; Dubinskas 1988). Therefore, the relevance of temporal assumptions as both a cultural framework and a cognitive map can be considered a factor that contributes to their strength and permanence within a given organizational context. As a result, the introduction of technologies with the potential to bring changes to this domain also involves a challenge to many of the cognitive orders and cultural values on which organizational actors rely.

Adopting the perspective of organizational culture, this research –in-progress paper seeks to contribute to the research on the temporal impacts of information technology in organizations, presenting a multiple case study that investigates the role ICT can play in promoting changes in the temporal dimension of organizational culture, and tests the hypothesis that temporal assumptions shared by people before a system is introduced can affect the way it is used, thus facilitating or hindering the achievement of the results the system is expected to convey.

Assumptions about time: a dimension of organizational culture
Schein (1985) in his study on organizational culture, maintains that organizational culture has groups’ dominant assumptions of time and space embedded within it and describes several “assumptions on time” that characterize different organizational cultures. Schriber and Gutek (1987) describe as well “norms about time” as specific “dimensions of organizational culture”, proposing scales to measure these dimensions. Besides being a fundamental condition for the coordination of activities and the production of organizational outputs, the temporal organization of work processes, as pointed out by
Barley (1988) plays an important role also as “an interpretive framework for rendering action in the setting meaningful” [p. 125]. In other words, organizational actors evaluate and make sense of events occurring during their own activity or other people’s activity using the temporal framework as a scheme of expectations to judge whether results and behaviours are appropriate.

Temporal patterns represent thus an expression of the specific organizational and professional culture which produces them, conveying a symbolic value, for the individual worker and working groups: according to Dubinskas (1988), the socially constructed character of time is such that all “times” existing within the high technology organizations of his study could be considered as “symbolic nexes around which coalesce issues of order, power, self definition and knowledge”.

As an example, in Barley’s study (1988) the different temporal organization of work of two professional groups, radiologists and radiological technicians, was also a representation of the different hierarchical and professional status of the two groups: given the “temporally unpredictable world” (p. 145) of the radiologists, technicians never knew when radiologists would be available and had to hunt for them whenever they needed one; radiologists on the contrary, given the predictable and highly scheduled “tempo” of technicians, always knew when they could summon up one of them.

Another example: Zerubavel found that a major aspect of the socio-temporal order expressed by the “schedule” of coverage in hospitals was that it functioned “as a moral order”, an expression of some fundamental organizational values like responsibility towards the patients, fairness towards staff members. Consequently, it was also a criterion to judge the appropriateness of personnel’s behaviour: some actions were considered “legitimate”, for example, only at the end of a shift, but not at the beginning. In this view, the temporal dimension appeared central for the definition of the “boundaries of norms”. The relevance of temporal assumptions as a framework and a cognitive map can thus be considered a factor which contributes to their strength and permanence within a given organizational context, consequently the introduction of technologies which have the potential to bring changes in this domain involves also a challenge to a multiplicity of cognitive orders and cultural values on which the organizational actors rely. On the other side, like all other social structures, they have a provisional nature and change over time (Bluedorn and Denhardt 1988; Ancona et al. 2001), also in association with technological innovations.

**Study Aim and Research Questions**

The purpose of this study is to investigate the role ICT can play in promoting changes in the temporal dimension of organizational culture: does the introduction of a new system contribute to change shared assumptions about time and time use in the workplace? Can ICT help transform the way people view time and the ‘appropriate’ way to collectively organize it within a given context (at organizational/department/group level)? Can temporal assumptions affect the way a new system is used, thus facilitating/hindering the achievement of the expected ‘temporal performance’ conveyed by the system? Our investigation into these issues is based on a multiple case study including three companies where a Workflow System was introduced in several organizational departments with the aim of improving cross-department processes, and addresses two main research questions:

1) Does the introduction of the workflow system transform the temporal assumptions shared by people in organizational units, thus achieving the expected temporal performance, and, if so, to what extent?

2) Do the temporal assumptions that exist in organizational units before the introduction of the system affect the use of the system, thus facilitating/hindering the achievement of the expected temporal performance?

**Assumptions about time: a framework for the analysis**

The theoretical framework adopted in this study to describe and measure the temporal dimensions of organizational culture is based on a set of concepts which have been operationalized with Likert scales and tested in previous research into the psychology of work and organizational culture (Schriber and Gutek 1987; Bluedorn et al. 1999).

---

Page 4 of 13  17th European Conference on Information Systems

Proceedings ECIS 2009 3
The framework was then integrated with concepts drawn from the ethnographic work by Zerubavel (1979) and Barley (1988) which have been adopted in order to assess more aspects of the temporal order within organizational departments which were not included in Schriber and Gutek’s model. Schriber’s work, based on data collected from 529 respondents from 51 work groups in 23 organizations, has developed and tested 15 Likert scales measuring perceptions on norms about time that have become standard measurement tools in studies at the work group and organizational levels of analysis. Bluedorn et al.’s work, based on data collection from 2 samples of 205 and 115 respondents from 2 organizations, has developed and tested a Likert scale instrument (Inventory of Polychronic Values) in order to measure the cultural dimension of polychronicity introduced by Hall (1959).

Both Schriber’s and Bluedorn’s instruments have been adopted to study changes in the temporal dimensions of organization occasioned by the introduction of information systems (i.e. Lee 1999; Lee and Liebenau). We investigated these dimensions by submitting questionnaires to employees and managers within the departments involved by the introduction of workflow systems. We also cross-checked these dimensions with the results of the content analysis of semi-structured interviews gathered within the firms. We will now explain the meaning and the relevance to our study of the dimensions considered in the conceptual framework. From Schriber and Gutek (1987) the following dimensions were drawn: Deadlines, defined as the importance of defining and meeting deadlines, temporal start-and-stop points of activities and Scheduling defined as the importance of scheduling (an activity that “concerns location in the temporal realm and gives organizations a framework for constructing temporal boundaries”). In the case of introduction of a workflow system, we could expect these dimensions to be relevant because the system could help introducing a more definite scheduling of the process, visualizing due deadlines for the deliveries and dates of completion of different parts of the process, thus strengthening perceptions and norms about the importance of meeting deadlines.

Synchronization and coordination: this dimension is related to the perceived importance of cooperating with others and working in a coordinated way or as a team. For the specific purpose of our study, this dimension was investigated because the workflow system was linking many different organizational units, thus the importance attributed to cooperation, coordination and teamwork among different organizational units was very relevant and could show significant changes.

Sequencing: the order in which activities and tasks take place; the authors distinguish this concept from scheduling stating that “scheduling is laying out a pattern of activities anchored to points in time within a specific time-measurement system, sequencing is the ordering of activities over time within that system” A specific order/sequence in time may be “inherent in the task, or prescribed by the individual who controls the process”. In our research this variable was relevant because the introduction of a workflow system can interfere with the customary sequence workers adopt in carrying out the daily activities (if any), both interrupting the sequence with queries coming from other subjects or departments, or prescribing a new sequence, embedded in the automated process.

Autonomy of time use: the perceived amount of freedom workers have in setting schedules for completing their tasks over time; The authors state that this dimension is also related to scheduling, although it is not a direct characteristic of it. It’s a secondary effects of scheduling, and reflects a more abstract level of the temporal environment. The variable was thus included in our framework because variations in schedules and deadlines, sequencing, and work pace due to the introduction of the system could bring as a consequence workers’ perceptions of enhancement or limitation of their autonomy over the way they use their time.

Awareness of time use: people’s awareness of how they use their time on the job and expectations that they know how long they take to perform activities. The concept is referred to the degree of attention and importance people pay to how they use their time at work (i.e., if they know and plan how long it will take them to accomplish a task or an activity, how worried they are about using their time well).

As in the case of autonomy, the variable was included in our framework because variations in schedules and deadlines, sequencing, work pace, or even tighter coordination with other organizational units due to the introduction of the system could bring as a consequence more awareness over the way they use their time. Speed vs. Quality: rules that people follow on trade-offs between the quality of work and the speed of work over time.
**Work Pace**: rate at which activities can be accomplished: the speed of work and people’s expectations to work fast. It concerns the perceived speed of work and people’s expectation to following their own rhythm, take breaks and so on rather then work fast. Each culture appears to have a pace that is considered appropriate for activity. This dimension was included because strictly related to speed: the introduction of ICTs is often associated to the objective of speeding up processes, which could result in increasing the work pace or, on the contrary, to slow it down because automation of phases of the process could facilitate worker’s activity, eliminate duplications of activities and so on.

**Allocation**: the amount of time devoted to a task or activity; can be considered a measure of work overload in that it defines the degree to which schedules seem too tight for activities/jobs; As for the case of work pace, this dimension was included in the framework because strictly related to speed: the introduction of ICTs, as pointed out before, is often associated to the objective of speeding up processes, which could result in increasing the perception that time to complete activities is tight or, on the contrary, that there is more time because automation of phases of the process could facilitate worker’s activity, eliminate duplications of activities and so on.

A few dimensions included in Schriber and Gutek’s model were not adopted for this study because not related to its object. They are: “time boundaries between work and non work”, “future orientation”, which is generally connected to perceptions about how much the firm invests in the future, “routine versus variety” related to the variation of job content in general, which was not relevant in our case since the system was applied to routine activities.

We included as well in our framework the dimension of polychronicity, as defined by Bluedorn et al. (1999) the extent to which people in a culture prefer to be engaged in two or more tasks or events simultaneously; and believe their preference is the best way to do things. The cultural dimension of monochronicity and polychronicity were first introduced by Hall (1959) at the level of national cultures. At organizational level a polychronic culture would value behaviors where individuals do several things at once, with short periods of time spent on each of several activities and people dealing with a number of different problems simultaneously. (Bluedorn et al., 1999)

We could expect polychronicity to be a relevant dimension to investigate since it a workflow system allows to have access and deal with a plurality of information at a time in a constant flux; moreover it allows to be connected and react in a flexible way to the flux of information, activities and queries of people working on the same processes in other departments.

In this study we included other two dimensions drawn from the ethnographic work by Zerubavel (1979) and Barley (1988). The first is the notion of Social cycle, defined by Zerubavel (1979) as a regularly recurrent pattern of activities and events, “the time intervals during which sequences of recurring successions of social activities are completed”. Zerubavel observed that the beginnings and ends of cycles are treated as discrete segments of time surrounded by rigid boundaries.

Many organizational activities may be structured in accordance with such rhythmic patterns, and the introduction of a technology can challenge the configuration of such cycles (for example allowing to work adopting a “flow”, destructured logic rather than a “batch logic” where a cycle of activity is completed before another gets started) or potentially alter the way in which cycles of different organizational units are connected to each other. We will briefly recall that this dimension is referred to coordination between the activities of individuals or groups by Barley (1988), who, building on the notions of temporal symmetry/ complementarity/ staggered coverage by Zerubavel, introduces the notions of temporal symmetry and temporal asymmetry. According to Barley the first type of temporal coordination implies that individuals or groups share a common pattern of temporal conditions (they share the same schedules, recurrences, and their working activities are synchronized). In the case of temporal asymmetry, individuals or groups operate according to different temporal patterns. The two authors point out that a condition of temporal symmetry constitute a powerful basis of mechanical solidarity among individuals and groups, while temporal asymmetry requires the development of organic solidarity. Moreover, in Barley’s study the condition of temporal asymmetry among distinct groups led to conflictual attitudes and relationships, which was reduced when their temporal patterns became more symmetrical. Though in our study we didn’t focus on the punctual mapping of the structural aspects of temporal patterns, we adopted the concept because we have different departments...
cooperating to one same process, thus temporal symmetries/asymmetries among departments might lead to cooperative/conflictual attitudes and be modified by the introduction of the system. From a methodological point of view, the two dimensions of social cycles within departments and temporal symmetry among departments were investigated through the content analysis of semi-structured interviews gathered within the firms. Ultimately, we propose the construct of “Expected temporal performance”, which we define as the whole of the expectations expressed by managers with regard to temporal issues, such as process acceleration, changes in people’s time orientations, changes in the temporal dimensions of a departmental culture etc. This aspect was investigated through the content analysis of semi-structured interviews gathered within the firms.

**Workflow introduction and time-related assumptions at Alpha, Beta, Gamma: Research Design**

The research, currently in course, is designed as a multiple case study (Yin, 2003), coherently with the descriptive and explanatory aim of the research project. Our research is based on the study of three cases. In all the three companies, which we will call Alpha, Beta and Gamma, a workflow system was introduced or is being introduced, in order to improve processes requiring the coordination of several different organizational departments. The relevance and significance of these cases to our study is confirmed by the fact that each of them: a) introduces the same type of technology solution b) applies the same type of technology to different processes c) represents different project phases. While Alpha, Beta and Gamma are three different kinds of companies, they have several key features which make them comparable for the purpose of our study: they are medium sized companies, operating in the service sector; and each of them has at least a 7 years story, which made possible to consolidate a peculiar organizational culture.

The study develops in two phases, having two different objectives: **Phase 1** Investigation of the temporal performance expected by the managers adopting the system. **Phase 2** Investigation of the temporal dimensions of culture existing in the three departments before and after the introduction of the system.

**Data collection**

Data is collected through documental analysis, semi-structured interviews with IT and Organization managers, department managers and employees of the three organizations, and questionnaires based on Likert scales addressed to both managers and employees. We first gathered background information through the analysis of company documents concerning the project and then collected data through semi-structured interviews and questionnaires with the managerial level and with employees. We interviewed so far from three to fifteen people in two out of three companies and collected 37 questionnaires in one company (one case is complete and research is still in progress in the other two companies).

**Table 1. Three cases: a preliminary comparison**

<table>
<thead>
<tr>
<th></th>
<th>Alpha</th>
<th>Beta</th>
<th>Gamma</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Project aim</strong></td>
<td>Managing the complaint</td>
<td>Managing the accounting</td>
<td>Managing projects</td>
</tr>
<tr>
<td></td>
<td>management process</td>
<td>passive cycle</td>
<td></td>
</tr>
<tr>
<td><strong>Project sponsor</strong></td>
<td>ICT and Organization dept.</td>
<td>Administration dept.</td>
<td>Organization Dept.</td>
</tr>
<tr>
<td><strong>Implementation phase of the workflow system</strong></td>
<td>Completed</td>
<td>In progress</td>
<td>Start up</td>
</tr>
<tr>
<td><strong>Number of organizational units</strong></td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td><strong>Number of workers involved</strong></td>
<td>37</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td><strong>Expectations on Temporal performance</strong></td>
<td>1. Speeding up the complaint management process</td>
<td>1. Speeding up the passive administrative cycle process</td>
<td>Research in progress</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
2. Reducing temporal misalignment among different departments
3. Shifting from a “batch logic” to a “flow logic”
4. Shifting from “indefinite urgency” assumptions to “definite deadline” assumptions

| Changes in temporal assumptions                  | Increase of temporal symmetry
|                                               | Increase of synchronization/coord.
|                                               | Increase of deadline awareness
|                                               | Adjustment of social cycles (partly)
| Increase of scheduling                          | Increase of awareness of time use
| Increase of temporal symmetry                   | Increase of temporal symmetry
|                                               | Adjustment of social cycles (partly)
|                                               | Research in progress
| Main facilitators                               | Internal workshop
|                                               | Internal technical support
| Main constraints                                | Workflow design and usability (beginning)

**Case I: Alpha - Research results**

The first case study was conducted in the Italian branch, employing 350 people, of a multinational manufacturing company. The Italian branch delivers commercial and distribution services to Italian customers. Here, a Workflow System (Lotus Notes) was introduced to improve Customer Service processes, in particular, the Complaint Management Service, the most prominent customer service activity, which requires the coordination of three different departments. As we will describe later, the workflow system was considered critical due to its potential to achieve results related to temporal issues. The company supported the change management process by organizing an internal workshop, held by the IT and Organization manager with the involvement of the department managers, which took place during the design phase in order to share the objectives and optimize the fit between the features of the complaint service process and the new system. The system had been introduced about eight months before the fieldwork started in July 2007, thus the implementation stage was sufficiently advanced to enable us to measure possible changes in the temporal assumptions within the departments.

**Expected temporal performance.** The analysis of managerial interviews showed that the main objectives inherent the introduction of the workflow system were widely shared and that expectations were very much related to temporal issues. Four expectations, relevant from the temporal viewpoint, turned out to be widely shared by all managers:

1. **Speeding up the Complaint Management Process**
   This meant accelerating the individual activities that make up the process, like the gathering of documentation on the customer’s order, collecting information about the specific problem encountered, monitoring the customer’s “complaint dossier”, and reducing the overall “lead time” of the process in order to provide faster answers to customers. “Lead time” was a very common expression used and the cross-analysis with the Workshop documentation confirmed it as one of the project’s key goals. It also meant reducing duplications and the time needed to produce and store physical documents: “less paper” was a common remark of all managers.

2. **Reducing temporal misalignment among different departments**
   Customer Point operators, who are subject to daily pressure from customers, were far more aware of the delays suffered in providing the customers with answers (“defining the dossier”) compared with other departments, which latter had other priorities and followed their own activity cycles. The workflow system was expected to facilitate departmental “alignments” on priorities and deadlines. This kind of objective can be better explained using Zerubavel’s concept of temporal asymmetry; these departments didn’t share the same “temporal order”, each of them having its own scheduling and activity cycles, which obstructed the complaint management process.

3. **Shifting from a “batch logic” to a “flow logic”**
The system was expected to make it easier for people to deal with issues instantly, as they arose and without waiting until they had accumulated a “pile of dossiers” on their desk. As the IT and Organization manager put it. “We want people to change their mentality, from a ‘batch logic’ to a ‘flow logic’, which means dealing with problems and requests as soon as they show up”. These remarks referred to the tendency to organize the activity in recursive “cycles”, occurring in some cases merely once a week (i.e. the storehouse operators were reported as checking the queries coming from the Customer Point only once a week) and also to a preference to do one type of activity at a time (monochronicity). Reduction/elimination of such cycles and enhancement of polychronicity were the objectives in this case.

4. Shifting from “indefinite urgency” assumptions to “definite deadline” assumptions

Managers reported that the appreciation of the level of urgency of a “dossier” was left to the individual operators, who, based on their experience and willingness, judged whether a complaint case was more or less urgent. There was no sharing of common deadlines for dossier definition and responses given to both internal and external clients: this resulted in a general feeling of uncertainty, well expressed by all managers when they stated that “everything is urgent here”, and that “in general, there has always been a rule that any complaint must be processed within 24 hours”. Managers admitted that it could take up to ten days in some cases.

Changes in temporal assumptions. Data from the employee questionnaires collected in the three departments were triangulated with data from ten in-depth interviews carried out with the four managers and six employees, enabling us to analyze the temporal dimensions of each department’s culture and assess the perceptions of change in these dimensions before and after the introduction of the system. For the purpose of this article we will focus on the analysis of the main changes in the temporal assumptions, following the introduction of Lotus Notes.

Data analysis showed a significant increase in three dimensions - deadlines and scheduling, synchronization and coordination and temporal symmetry among departments - after the introduction of the system. Deadlines were not perceived as so important before the introduction of the system. Although the respondents reported that there was a belief that “everything is urgent” and a general rule that “overall, all complaint dossiers should be opened on the day the complaint arrives”. Others reported that “we didn’t really think of deadlines, it was more indefinite”. According to the results of the questionnaires, the value of this dimension scored very high ten months after the introduction of the system. Respondents reported that what really made the difference was the introduction of a formalized classification of complaints based on the expected completion time and the visualization of dossiers flagged by a colored tag, which was visible to the operators of all departments, thus reminding them of the existence of a deadline and that it required alignment among departments.

The introduction of this system of deadlines and of the colored tag “artifact” represented a liaison between departments: using Barley’s concept of symmetry, the three units now share a common deadline system and have to meet aligned deadlines, which has increased the overlap among their different temporal orders.

The value of synchronization and cooperation among departments ranked very high in all three departments and was perceived as having increased significantly because the sense of “teamwork” with the other departments had also increased. Customer Point operators reported that, pre-Lotus Notes, the synchronization with other departments was very poor: other departments were reported as “having their own time”, “creating bottlenecks”, “being slow to give answers”. All the interviewed customer service operators shared the strong conviction that this situation had improved significantly with the introduction of the workflow system because their requests are now transmitted instantly through Notes and the date and hour of the request, as well as the other department’s response, are recorded. On the other hand, operators from other departments reported that the system made it “simpler and quicker to gather and send documents because the databases are now interconnected”.

The increase of these three dimensions confirmed also the achieving of three expected temporal performances: the lower level of misalignment among departments, the adherence to defined deadlines, and the speeding up of the customer complaint process.
This kind of evidence would seem to confirm the first hypothesis of the study, that the introduction of workflow systems has the potential to transform the temporal assumptions shared by people in organizational units. Nevertheless, it is important to note that the introduction of the workflow system was combined with an internal workshop involving the managerial level and that the results we report were associated by the respondents with both the innovations introduced by the system and the effectiveness of the workshop as a change management strategy.

To the contrary, the dimensions of polychronicity and sequencing showed no significant change, while the social cycles characterizing the departments revealed contradictory patterns of change. These dimensions were associated with the achievement of the objective whereby, in performing their activities, workers would shift from a batch logic to a one-piece flow logic.

Polychronicity scored low in the questionnaires gathered from all three departments. When triangulated with the interview data, the result was explained in this way: in all departments the workflow was recognized as fostering polychronicity because all data and documents were stored in a single database, making it “easier to open and close items related to different activities simultaneously, having links immediately available”.

On the other hand, employees pointed out that they didn’t like working this way and that there wasn’t yet a shared belief that this way of working was “better”. High scores of time allocation dimension cross-confirmed that there was a feeling of an increasing work overload. Further, when under pressure to deal with overloads, operators shifted back to batch logic: for example, Customer Point operators during phone call peaks (twice a day) interrupted their other activities to concentrate on answering the phone, admitting that “messages in Lotus were left on stand-by”. Another case was when they had a number of administrative tasks requiring high levels of concentration: they adjusted with colleagues in order to divide labor based on specialized activities and followed a monochronic logic. Interestingly enough, as far as Social Cycles are concerned, the interviews showed that, though the pressure toward a “flow logic” had increased, operators retained their previous cycles: at Customer Point, the activity is still organized around the “phone call peaks” and four main daily cycles are still in place. This influenced, as pointed out in the previous section, a use of the workflow system that is not yet thoroughly in line with the expected “flow logic”. The same happened in Storehouse and Logistics, where the importance of sequencing is a key feature of the departmental culture and where monocronicity is high. Here, operators reported that there had been an adjustment between the social cycles of their activity and the need for more polychronicity: they used to follow a five-six day cycle in complaint management activity: i.e. the “pile of paper” here was left to grow until, finally, they dedicated one entire day to this specific, time-consuming activity. The presence of this cycle expresses the “batch logic” they followed. In addition, this created a noticeable temporal asymmetry with Customer Point operators, who were left waiting for answers for up to a week. This department is characterized by another typical cycle, the Morning/Early afternoon cycle. In the morning, ordinary activity takes place; at one p.m. afternoon planning starts, while deliveries take place through to five p.m. At present, complaint management has been re-allocated according to this second cycle: Lotus Notes is checked in the morning and, accordingly, complaints start to be checked in a more “flow oriented” logic. This allocation of the complaint management activity to the morning/early afternoon cycle has enhanced the symmetry with Customer Point. In this case, the shift to a “flow logic” has been partly obtained through an adjustment to existing social cycles and to a still strong orientation to monocronicity in departments. To summarize, our findings suggest that the persistence of two temporal dimensions of organizational culture - monocronicity and sequencing - and the strength of the social cycles existing within the departments have influenced the use of the system and the achievement of one important expected temporal performance, the shift to a flow logic. This result seems to support our second hypothesis that temporal dimensions of organizational culture can affect the use of the system, thus having an impact on and even hindering the achievement of the expected temporal performance.

Case II: Beta – Preliminary results

Beta is a company created by a consortium of about twenty banks in northern Italy. It provides full outsourcing services to the banks belonging to the consortium, employing about 250 people.
Our fieldwork in this company, still in course, takes place in a mature phase of the project, where we can investigate consolidated outputs and changes.
In 2003 Beta introduced a workflow system (docflow) aimed at managing their accounting passive cycle (the whole process leading to suppliers payments), which is, in this company, the most prominent accounting activity since they have a lot of suppliers but very few customers (i.e. banks forming the consortium). This process requires the coordination of three different organizational units: two different offices within Administrative Department, and Purchasing department.

Expected temporal performance.
The analysis of interviews with the managerial level showed that in this case the objectives inherent the introduction of the workflow system were partly related to temporal issues. In this project attention and expectations were initially focused on reducing the production, duplication and storage of physical documents (“reducing paper”); another priority was to create a logical flow in order to decrease information redundancy, and finally to improve quality and reliability of data gathered by administrative department. In this case, compared to the previous one, managers showed a lower level of awareness about the potential of the system from a temporal point of view.
This lower level of awareness is expressed by a poorer articulation, compared to Alpha, of possible time-related objectives going beyond the “speeding up” issue.
Two objectives in this project were explicitly addressed to temporal issues:

1. Reducing the duration of passive administrative cycle process
This was explained as the need to reduce the overall time during which suppliers invoices moved around from one office to another before being paid.
The process required that invoices, once arrived via mail or fax to one of the two offices within Administrative department, were duplicated and sent to buyers operating at the purchasing department (who were responsible for specific contracts with suppliers) for technical check and authorization to payment. Another duplication was sent to the other administrative office for the checking of contractual conditions. During this double check and authorization process, invoice copies could either be left standing by until checks were completed, or move back to the first administration office with a request for keeping the payment suspended, or move back with authorization to payment. The duration of the whole process was described as “unpredictable” and it was often difficult to make out why an invoice hadn't come back and which office was keeping it suspended. For the same reason it was difficult to inform creditors about the state of their payments. Thus the workflow system was expected to speed up the process: partly because of new electronic format of dossiers related to invoices and the possibility to track them along the process, partly because it could facilitate departmental “alignments” on priorities and deadlines.
This kind of objective can be better explained if we refer to the concept of Awareness of time use: there was an overall unpredictability of duration of the process for many reasons: lack of awareness of how long each one would take to complete his own check, lack of awareness of reasons why there were delays and in which part of the process. Another aspect of this objective can be related to Zerubavel’s concept of temporal asymmetry: the different units, buyers in particular, didn’t share the same “temporal order”, each of them having its own scheduling and activity cycles, which obstructed and made the duration of the process unpredictable.

2. Respecting payments deadlines
Unpredictability of the duration of the process caused delays, which made it difficult for administrative operators to be punctual in meeting payment deadlines Even if deadlines were set within contracts, administrative operators had to work hard to compensate for delays; moreover, it was difficult to give creditors timely information about the state of their payments.

Changes in temporal assumptions.
We are not able to give a complete account of this case, nor to confirm our hypothesis yet because data collected in this company so far are not complete. Data analysis of the interviews collected so far with managers and employees who are still working in the company and can account for the changes occurred after the introduction of the system, allow us to present only some preliminary results.
A significant increase is perceived by interviewees in three time-related dimensions — scheduling, awareness of time use, and temporal symmetry among departments - after the introduction of the system. On the other hand, social cycles seem to have only partly changed.

As far as Scheduling is concerned, the workflow introduced and made visible a clearly structured sequence in the process. The operators reported that structured scheduling is now perceived as very important because all operators can locate at what point, in the temporal domain of the process, a single invoice is and what is its advancement state. As administrative operators put it “Now I can see where invoices are and who is in charge for it at any given moment” The increased importance attributed to a more structured schedule reduced as well the feeling of uncertainty about the ongoing of the process, expressed by sentences like: “I feel more in control” “Now the process is under control”.

Every action in the process being tracked, awareness of time use seems to have increased because everyone can check how long it takes to everyone else to complete his “to do list”. As an administrative employee put it, “information about ‘bottlenecks’ is now transparent”. Moreover, the system sends automatic warning mails when a due action is not completed, contributing to awareness of one’s own time use as well. As far as Social Cycles are concerned, it turned out that, before the introduction of the workflow, both buyers and administrative employees followed a similar rule: the so called “bunch rule” which meant that they would care of invoices payment process when the bunch of invoices was thick enough. This rule is similar to the “pile of dossiers” rule in the previous case. This rule gave rise to cycles which could last from one to two weeks for buyers, while administrative employees followed a one/two days cycle during ordinary periods of the year (when not under pressure for other important tasks, i.e. closing balance sheet). This created temporal asymmetry between buyers and the other units. Interviewees reported that these cycles have partly changed: administrative employees state that their activity is now generally driven by the “to do” proposed by the system. Buyers have partly modified their routines, due to the fact that access to electronic documents and data about invoices is much easier than before. Their activity has other priorities, thus many of them continue to take care of this process periodically. Interesting enough, both groups reported that they developed a sort of “bridge” between their different temporal orders based on the artefact of “urgent query”: when requested, buyers are much more flexible and quicker than in the past to give answers on specific and urgent issues blocking a payment. They recognized this “bridge” is now possible because they can have easier access to workflow data wherever they are.

**Preliminary conclusions**

Despite its importance to temporal issues, research into the temporal impacts of information technology in organizations is still limited. On the other hand, organizational culture research shows that the way time is perceived and collectively organized reflects cultural assumptions that are an expression of the specific organizational setting (at firm, departmental, group level), underscoring that cultural assumptions are an important contributory factor to the strength and direction of organizational change. In this contribution, we have investigated the role ICT can play in promoting changes in the temporal dimension of organizational culture, and sought to assess whether temporal assumptions can affect the way a new system is used, thus facilitating/hindering the achievement of the expected results. The preliminary results of this multiple case study show that managers have different levels of awareness about potential temporal performance of the system.

Case Alpha covered four types of ‘temporal performance’ management expected to see thanks to the introduction of a workflow system. In Case Beta managers showed a lower level of awareness about the potential of the system from a temporal point of view. This lower level of awareness was expressed by a poorer articulation, compared to Alpha, of possible time-related objectives going beyond the “speeding up” issue. Both cases showed that, after its introduction, the temporal dimensions of the organizational culture of the departments involved showed some significant changes, which seems to confirm hypothesis 1 of the study, but also some contradictory effects that seem to confirm hypothesis 2. In case Alpha, significant increases were seen in three dimensions - synchronization and coordination, temporal symmetry, deadlines and scheduling. The increase of these dimensions confirmed the achievement of the three expected temporal performances: the reduction of the misalignment among departments, the shift to definite deadlines, and the speeding up of the
process. This seems to support the first hypothesis of the study, that the introduction of workflow systems helps transform the temporal assumptions shared by people in organizational units. It is important to remember that the system was introduced in tandem with an internal workshop involving the managerial level and that the results reported here were associated by respondents with both innovations introduced by the system and the workshop. Nevertheless, the assumptions underlying the objective that workers would shift from a ‘batch logic’ to a ‘flow logic’ in performing their activities failed to show any significant change: data analysis shows that the persistence of two temporal assumptions - monochronicity and sequencing - and the power of the social cycles existing within departments influenced the expected use of the system and the achievement of one important temporal performance, the shift to a flow logic. This supports our second hypothesis that temporal dimensions of organizational culture can affect the use of the system, thus hindering the achievement of the expected temporal performance. Preliminary results from Case Beta also seem to confirm that significant changes were seen in three temporal dimensions - scheduling, time awareness and temporal symmetry. It seems remarkable that in this case these changes can’t be considered as planned results but rather as unexpected consequences of the workflow introduction.

References
Sahay, S. (1998), Implementing GIS technology in India: some issues of time and space. Accounting, Management and IT, Vol.8: 147-188

Objects and their participation in the interdisciplinary design and development of computer games

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0312.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Computer Games, Collaboration, Research in progress, Case Study</td>
</tr>
</tbody>
</table>
OBJECTS AND THEIR PARTICIPATION IN THE INTERDISCIPLINARY DESIGN AND DEVELOPMENT OF COMPUTER GAMES

Abstract

This paper investigates the role of boundary objects in the interdisciplinary collaborative processes found in computer games development. It draws on data from an in-depth case study in a computer games studio that explores boundary objects in relation to the compelling, sensory and entertainment-centred game-playing practices that inform computer games design and development. Sensory user experience and aesthetic considerations – of primary importance in computer games development – are becoming increasingly significant in the design and development of many other kinds of software and information systems. For this reason developments in the design and production of computer games have wider implications for other software and information systems settings and provide valuable insights into processes of collaboration that bridge cultural and aesthetic as well as technical forms of expertise. The paper seeks to provide insights into how objects contribute to such collaboration, with attention focusing especially on how game developers devise objects that span boundaries and draw on these in their collaboration. Through its focus on the material production and practices of computer games development, the research presented also seeks to contribute to the theoretical treatment of interdisciplinary collaborative working in software design and development via a critical assessment of the concept of boundary objects in the setting being studied.

Keywords: Boundary Objects, Computer Games, Design, Development, Aesthetics, Representation, Collaboration, Interdisciplinarity, Art, Programming, Practices.
1 INTRODUCTION

Most work is conducted by diverse collectives of people and things from different backgrounds and with varying levels and types of capabilities. Shared artefacts are seen by a number of studies from a range of literatures as playing a crucial role in the bridging of boundaries of human expertise and embodied skills involved in collaborative work (Gerson and Star 1986; Callon 1991; Henderson 1991; Fujimura 1992; Bowker and Star 1999; Carlile 2002; Bechky 2003; Carlile 2004; Bruni 2005; Engestrom and Blackler 2005; Levina 2005; Levina and Vaast 2005; Miettinen and Virkkunen 2005; Suchman 2005; Ewenstein and Whyte 2007; Luck 2007; Nicolini, Mengis et al. 2008). For example, a physical prototype (Carlile 2002; Bechky 2003) might help a designer communicate his/her vision to a production engineer and at the same time help the engineer understand the implications of that vision for production of that design in order to respond accordingly. Further examples from existing studies include design drawings (Bodker 1998) and engineering sketches (Henderson 1991; Bechky 1999).

More specifically, artefacts that exhibit a capacity to help different collaborating groups share representations with each other have been conceptualised by Star and Griesemer as boundary objects (Star and Griesemer 1989). Star and Griesemer's starting point for developing the concept of boundary objects is that for the success of collaborative work, cooperation and coordination need to be achieved across domains, and shared meanings and understandings established. For this to take place, information needs to "retain its integrity across time, space, and local contingencies" (Star and Griesemer 1989). This does not necessarily mean that for cooperation to take place consensus must be achieved. Instead, those involved in collaborative work have to "translate, negotiate, debate, triangulate, and simplify in order to work together" (Star and Griesemer 1989).

"In conducting collective work, people coming together from different social worlds frequently have the experience of addressing an object that has a different meaning for each of them. Each social world has a partial jurisdiction over the resources represented by that object, and mismatches caused by the overlap become problems for negotiation. … Because more than one world or set of concerns is using and making the representation, it has to satisfy more than one set of concerns." (Star and Griesemer 1989)

According to Star and Griesemer boundary objects "both inhabit several intersecting social worlds ... and satisfy the informational requirements of each of them" (Star and Griesemer 1989). They are "both plastic enough to adapt to local needs and the constraints of the several parties employing them, yet robust enough to maintain a common identity across sites" (Star and Griesemer 1989).

"These objects may be abstract or concrete. They have different meanings in different social worlds but their structure is common enough to more than one world to make them recognizable, a means of translation. The creation and management of boundary objects is a key process in developing and maintaining coherence across intersecting social worlds."

While a number of studies have made strong cases for the role specifically of boundary objects in collaboration (Gerson and Star 1986; Henderson 1991; Bowker and Star 1999; Carlile 2002; Carlile 2004; Levina 2005; Levina and Vaast 2005), most of these have been concerned with more conventional science, technology, and business settings where the forms of collaboration involved are functionally-based. There remains, however, a significant gap in the existing research regarding the role of boundary objects in relation to collaborative efforts that depend on linking technical forms of expertise with expertise in the areas of entertainment, culture, and aesthetics.

Computer games design and development, involves a highly diverse set of participants in multiple roles (Gershenfeld and Loparco 2003). Furthermore, because computer games are designed to be compelling on a sensory and entertainment level as well as to perform functionally, their development is an arena of techno-cultural production and differs from many other collaborative settings (Baba and Tschang 2001; Swartout and Lent 2003). For this reason computer games development was seen as a
good setting through which to investigate the importance of boundary objects in a situation of collaborative production characterised by a cultural and entertainment dynamic and to explore the forms of boundary objects being drawn upon. In particular, the research has sought to investigate what the role of boundary objects in the interdisciplinary collaborative processes found in computer games development is and whether the boundary objects involved reflect game-playing practices as compelling, sensory and entertainment-centred.

The article presents key boundaries and objects that span these boundaries identified so far in computer games development and describes the role of these objects in terms of how they contribute to collaboration. Attention focuses especially on how game developers devise boundary objects and draw on them in their collaboration. It also suggests ways such devised boundary objects might relate to specific performance outcomes and what the properties of boundary objects associated with successful projects might be.

2   RESEARCH APPROACH

As outlined, the study focused on boundaries of expertise encountered in an interdisciplinary computer game development setting and the use of boundary objects in the process of realising such an innovation outcome. To this end, an interpretive case study research approach was chosen (Walsham 1993) that focused on the work practices of game developers and the boundary objects they interact with during the development of these games. In particular, the research has aimed to acquire an in-depth understanding of the way in which the developers studied create, leverage, and alter boundary objects in this work and to access the meanings they create and attach to them (Ngwenyama and Lee 1997). This will ultimately form part of “thick descriptions” of the research phenomenon through engaged field experience that capture its distinctive complexity (Geertz 1973).

2.1   Data Collection

Through a combination of in-depth interviews and observations at computer games developer studios the research project presented is seeking to investigate the way the interdisciplinary collaboration involved in the development of computer games is performed by the humans and objects involved. This paper is based on interviews and some initial observations conducted at GamesDevCo, a pseudonym for a UK-based games development company, and the first of three planned case studies investigating the design and development of computer games at different development companies.

The researchers made a number of visits to the company’s studios and headquarters between September and December 2008. During these visits a tour of the five divisions was undertaken. The various teams that comprise each division were encountered and informal and unstructured conversations with team members relating to their work, their interactions with other teams, the way they work within the team, and the importance of objects to this work took place. An in-depth formal interview with the company’s director of development was also conducted and there were also a number of more informal discussions and interactions with the company’s director of business development.

A range of artefacts involved in the design and development of past games titles also formed an important part of the evidential basis for the fieldwork. Finally, a significant amount of documentation generated from a part publicly funded joint commercial and academic international research project relating to the development of reusable multimedia objects was also accessed. This provided an in-depth view of the approaches the company was exploring in terms of re-using games ‘assets’ and how to enable those from an more creative background without expertise in computer technology and programming to interact directly with the technical components of games and input graphic elements and animations without the intervention of the technical experts.
Data collection also sought to pay attention to material entities, activities, schedules, hierarchies, routines and variations, significant events, participants’ meanings, and social rules (Altheide and Johnson 1994).

While the use of other forms of evidence such as photographs of the studio and teams at work, pictures of the artefacts and objects studied, and screenshots of tools and objects used in collaboration was considered, that was not possible in this case study due to the strict non-disclosure conditions stipulated by GamesDevCo.

Interviews were recorded and transcribed, while observations were intended as a way of appreciating in greater depth and detail understandings garnered during the interviews.

2.2 Data Analysis

The analysis of the evidence has sought to counterpose theoretical formulations relating to boundary objects with the empirical data collected in order to develop an in-depth understand of how a variety of constructed boundary objects help shape, are shaped by, and represent shared meanings or interdisciplinary intersubjectivity.

From the data collected, the analysis first sets out to identify key objects used in the design and development of computer games at GamesDevCo and that were found to span the disciplinary boundaries in the company. Attention focused in particular on how these objects identified were involved in the collaboration across disciplines and what kind of interactions among the developers they were a part of.

Drawing from the original conceptualisation of boundary objects developed by Star and Griesemer (Star and Griesemer 1989), a mapping between the objects used in collaboration across disciplines in the empirical setting and the four key categories of boundary objects (repositories, ideal types, objects with coincident boundaries, standardized forms) proposed by Star and Griesemer was undertaken. On the comparison made possible by this, an analysis of conceptual concurrence and difference was undertaken, which formed the basis of the discussion that this paper presents.

3 CASE STUDY IN PROGRESS

Since its foundation in 1990 GamesDevCo has grown into a leading independent multi-platform developer employing around 250 people and comprising of five distinct divisions: family games; mature titles; serious games; downloadable games; and games technology. The company develops games under both its own brands as well as on behalf of external publishers and intellectual property rights holders.

3.1 Formal Structures and Disciplines

Formally, the various divisions of the company and the teams that comprise them are divided along the following three main disciplinary lines: art; design; and programming. The way the teams are structured in practice, however, can vary “as project needs ultimately dictate the structure and skills of the team and teams also vary in size as what [each] client needs is often very different”.

The broad structure the company adheres to was described in the following way. A project director manages schedules and interacts with clients. Then, in most divisions there will be a manager for each of the disciplines involved, with, for example, a technical manager for the programmers, a design manager for design issues, and an art manager in charge of animators and concept artists.

This structure reflects the current company view of the key disciplines involved in the games development process. In practice, however, there appeared to be much more fluidity in the disciplinary boundaries that were likely to shift and be reconsidered. An example was given regarding the
introduction of an animations manager because the animation requirements for a project were considered so demanding and there were a large number of animators involved in the work. “It is something we are experimenting with” commented a senior development executive.

Below these discipline-specific managers each team in each discipline will have a “lead” or “senior” person “to help managers organise and lead their respective departments”. These senior figures were described as “important to communications” and “senior in terms of what they have done”, with 5-10 years of experience and having “done high-profile games, either previously at the company or at other companies”. They were described as knowing “inside-out the company’s technology, its games engine, and how to make games”. Their job titles are along the lines of “lead programmer” or “lead designer”, “lead artist”, “lead animator” and so on.

Under this level are what were described as “general programmers, artists, designers, animators, and concept artists who do [technical design] concepts and visualisations”. With bigger project teams of 50-60 people, specialisations might “go a lot further”, especially with regards to art. For example, with regards to animation for large high-profile projects they now have people who are solely involved in “rigging” characters which involves “putting the skeletons on characters, but not actually animating them”. “It is quite a big job and very critical with regard to the success of the project and the success of the animations”. Similar developments are taking place with regard to 3D artists with the company now having environment artists, character artists, and “people … just dealing with lighting”. Disciplines were described as “getting more and more specialised”, with the example given of the introduction of “special effects-only artists”.

“The bigger teams definitely have the ability to have people that are specialised in a certain skill and very much at the top of their game with regards to smaller parts of the game that make the difference. The smaller teams don’t really have such needs, but that will also often depend on what the client needs”, explained the company’s development director.

The reason given for this specialisation and fragmentation of the three main disciplines were as follows:

“The stakes are so much higher regarding the visual quality of the games, the complexity of the designs, and the technology we are using. Years ago … any decent artist could get away with doing special effects and knowing what to do, because it was very much dealing with 2D assets. These days, depending on the target hardware, what the [games] engine is capable of, and the technical requirements of the [games] engine a lot of special effects artists need to have a very good technical grasp of the technical aspects of the game in terms of coding and also for coders to have a very good understanding of how the rendering technology works with regards to that rather than just have a good knowledge of coding. You need to have extra knowledge on top.”

The director of development talked about the necessity to “marry-up two different people” in one role creating “a ‘technical artist’ who knows a lot about hardware, about software, about good visualisation techniques, but also has a very intuitive mind in terms of experimenting with new and unusual ways of doing say a fire effect or a sparkling effect that goes in a totally different way than most others would do”. Such a person would do “research stuff that instead of going down the normal route … could go in a completely opposite way” by talking with people and “figuring out that there may be a really cool alternative way of doing something”.

Individuals from all the three disciplines are involved from the early stages of the origination of a game, with the creative side and designers taking a lead, but “senior technical guys and senior art guys” following them closely. “The industry is so tied together discipline-wise”, explained the director of development, who went on to add:

“Any change within an art specification could change the technical specification and the design specification definitely does [change] in terms of how many characters you can have on screen or how big the levels are and what the designer might want to do. The technical guys must look
at that and say: ‘is that possible, how is it possible, how much would it cost if we did do it?’ We try to involve those three sides of the disciplines – design, art, and technical sides – at a very early stage along with the manager. … Ultimately the project director will look at that from a costing and … scheduling point of view and say ‘we have x amount of time to do this in and these resources we can bring, what are the major risks we need to be focusing on’ and get a good estimate really of what kind of skills will be needed, what amount of people they are going to need and so.”

In the case of GamesDevCo that works quite a lot with third party intellectual property rights holders, the GamesDevCo managers dealing with the interdisciplinary team from the client side also have to consider the specifications of a project strategically and technically in terms of what the studio wants to do with regards to its styles and “where [they] want to take the engine”. “We work with that and make sure the pitch fits with what we want to do as well as meeting the clients’ needs at the same time; we kind of marry those two up and then go in to further negotiations, explained a senior GamesDevCo executive.

As can be seen in this section, even at the level of concept origination for a game, a significant amount of interdisciplinary work is already involved. At the same time, while disciplines are recognised at the corporate and organisational level in the company, in the actual work practices of the teams developing a game the disciplinary boundaries are porous and flexible.

3.2 Concept Book

An important material nexus for this early-stage interdisciplinary work as the computer games move from conceptualisation to actualization was what was often referred to by the GameDevCo developers as the “concept book” and discussed as follows:

“In the end [of the conceptualisation phase] we get to the point where we create a high-end specification document that looks great and [is presented] with good PowerPoints and videos and all sorts of madcap stuff. One thing we really have trouble doing in the industry is selling the vision of the game. … Trying to get that across is quite tough to do and quite expensive. Years ago we used to go through word documents and write stuff out and give lists and a bit of concept art here and there; nothing too major really. Now days we are dealing with producing videos and actual 3D assets which is probably the most expensive thing we do at the studio. Even [just] producing the documents you have seen takes a huge amount of resources and time to do.”

The samples viewed were from past games. They were very professionally produced and aesthetically engaging documents, styled in accordance with the theme of the proposed game. For example, one for a very successful science fiction film series had covers made out of shiny metal that was shaped and indented in such a way as to look like one of the very characteristic spaceship doors in the film. Another for some medieval adventure had the look of an old and musty book.

The documents included text that described the game and its features to the reader, the thinking behind it as well as visual representations of the main characters of the game accompanied by what can be described as imaginary biographies for them outlining their roles in the game and what they could do. The documents often – though not always – also included outline budgets and cost projections for the proposed game. It was explained that the production of these documents was undertaken by a specialist “pre-development” team that included individuals from business development and marketing as well as individuals with skills in business analysis, project planning, desktop publishing, and writing and editing text. This core team would then draw from the expertise of the different developers involved in the project from the three formal disciplines acknowledged by the company of art, design, and programming.
3.3 Game Design Document

Once a more formal agreement to proceed with the development of a title has been reached, whether this is with an external publisher or in relation to internal approval procedures, a number of new objects that are crucial to the interdisciplinary collaboration involved are assembled, as the following passage describes:

“The aim at the beginning of each project … is to create a ‘game design document’ an ‘art design document’ and a ‘technical design document’. The game design document will contain everything that is in the game. It will classify all the characters, all their moves, all the mechanics, all the animations needed, all the pickups, all the weapons, all the locations, all the mechanics. That will grow to at least a couple of hundred pages for just that. … [The game design document] is also crucial to the relationship with the client [as well as in terms of] visibility for the collaboration. The same with the art and technology design documents and what they deal with. Everything is documented in terms of meeting notes. Everything visual is designed and we obviously design everything digitally or scan it in or drawn digitally. So we keep a record of that.”

The games design document is not only a repository of all the elements that need to be assembled during the development of a game. Because the milestones agreed for the delivery of different parts of the game are also included, much of the scheduling of the tasks for a project and included in a “milestone schedule” will also flow from it.

These milestone schedules are a central coordinating device in the development process that span all the teams and disciplines involved in the project. They reach right down to the level of each individual, with printouts of graphical representations of the schedule observed fixed either side of their computer screens on to the wall or the separator screens on to which the desks of the developers abut.

The games design document is not a static high-level brief. It evolves and grows together with the games project, with minutes, drawings, and other relevant resources and assets added throughout the development process.

3.4 Drawing and Visual Representations

Drawings, within the context of both the concept book and the games design document, but also more generally in the games development process as a whole, were found to have a crucial boundary crossing role. Drawings – but also visual representations in general – were very much linked to one of the central preoccupations for a games development company: how to capture and communicate, both internally and externally, the “vision” for a game. The importance of drawings is captured in the following passage:

“We try and draw a huge amount of stuff during the project because the cheapest way of getting any visualisations is by drawing. The [art specialists] are trained to draw extremely fast as well, so we spend a lot of time drawing out the environments, drawing out some of the character moves in regards to the animations cycles, drawing out all the characters' weapons, individually style anything else we need, sometimes just drawing with regards to diagrammatic things saying: 'I want this character to move like this', or, 'here is one of the character moves and I want it to look dynamic in this kind of way', or whatever. Some quite functional things like that. So, when we hit production we've got a huge amount of material there.”

But it is not just the drawings that are so important in terms of the circulation of the vision for the game. Within the studio there was a wide-spread use of all sorts of sizes and types of models in the development process ranging from miniature mock-ups of landscapes made from the kind of modelling materials used by model railways enthusiasts, to small sculpted figures or portraits of
characters. Throughout the studio, props and objects from various well-known film and television titles could be seen on the desks of individuals and in areas occupied by different teams, but also all around the office space, giving a visually intense feel to the place.

Even external visual resources are utilised in order to convey to the individuals involved the elusive vision for the game, as illustrated in the passage below:

"With everything we have created, even if it is ‘true to original’ there is always a movie, or a book in some cases, or another game possibly, that have done something similar or have done something diametrically opposite that we can say: ‘this is really what we don’t want, we really don’t want this vision’. Or, ‘what I am trying to get to is this’, or ‘here is a movie’. Everyone watches the movie and they then hopefully understand what you mean about the emotion of the game or the visual style of the game or whatever it might be. … Using those … references, to say, 'right we really want this', then … people come up with ideas and come up with visual styles [and] that's how [the vision] works; it kind of trickles down."

3.5 Other Objects

The development of computer games at GamesDevCo also depends on a number of tools and development methodologies that have knowledge, skills, expertise, and ways of viewing and knowing the processes involved in games development built into them.

These include computer programmes for the digital manipulation of artwork, for creating high-end 3D assets, development platforms for writing computer code, and project management applications. Of particular interest in relation to understandings of collaboration built into such tools was provided in this case study by the decision of GamesDevCo to abandon Microsoft Project in favour of a programme called Hansoft, which has been developed by a number of games developers who built a specialist project management software application.

4 ANALYSIS AND DISCUSSION

Our account of games development projects at GamesDevCo has focused on three of the objects that were found to play vital boundary crossing roles because of the important consequences for interdisciplinary collaboration they were found to have in this setting.

The concept book not only crosses the boundary between the games studio and the client, but also allows the developers from a number of areas of expertise to bring together in a common space their interpretations of the vision, concept, and ideas of the proposed game.

These interpretations are then, for the purpose of the concept book, expressed in text, textures, materials, drawings, photographs, tables, and spreadsheets that have been made combinable but also accessible by all those involved in the games development process and from different areas of expertise with different representational techniques. The co-production of shared representations is a central element in the conceptualisation of Star and Griesemer (Star and Griesemer 1989), as the following passage shows:

"Because more than one world or set of concerns is using and making the representation, it has to satisfy more than one set of concerns. When participants in the intersecting worlds create representations together, their different commitments and perceptions are resolved into representations - in the sense that a fuzzy image is resolved by a microscope." (Star and Griesemer 1989)

The game design document is also crucial in this respect. It specifies a whole set of objects and relations among these objects that must come together in a game. It will specify, for example, the number of levels the game will have, what these levels will be, the storyline of the game and how that
relates the different levels to each other, who the characters are, what their role in the game is, where they can appear, how they can encounter each other, what the mechanics of their interactions are going to be, how they are going to move, what ‘things’ they can interact with or manipulate, and so on. These are all objects that whether an individual participating in the development process is a coder, an animator, a 3D artist, or an special effects or artificial intelligence specialist, will, during the trajectory of the game from concept to actualisation, continuously cross backwards and forward between individuals and groups of individuals with a particular expertise and way of working and interacting. As was explained during the research, this way, when for example the character artists come into the team, they can see straight away what characters are needed from the concept art already done in 2D and included in the document. Or a programmer can “look at the game design document, look at the pages which relate to character movement or even read the whole document with regard to how all that is going to feed into the mechanics and other stuff and then be able to think, 'OK I have got it.'”

Furthermore, the dynamic nature of this documentation means that this a continuously expanding body of material of all kinds, both textual and visual, that tracks and accompanies the game through its development, recording but also making available back to them the interactions of all those involved in the project. It also provides an updatable representation of the game being developed itself, but in such a way that this representation is a shared representation that can be used by all the different disciplines for their own purposes.

This document also includes the crucial for every project, team, and individual team member “milestone schedule”. The milestones themselves are more that just a project management device. It is through them that the performance of individuals, teams, and ultimately the company itself are judged, evaluated and rewarded, as the following passage from an interview illustrates:

“The project director gets involved … to create a ‘milestone schedule’ with a list of deliverables for the duration of the project. [The milestones] are usually monthly and those deliverables contain certain parts of the game like an x amount of characters and an x amount of levels or whatever it might be. It is important that the teams are involved in whatever they are signing-up for as it is important that the milestones that are agreed on are met, as this is how we get paid monthly in arrears from our clients. We have to be aware that our milestones are judged by the fact that if we hit the milestone – which we luckily do mostly – then we get paid. That is how the visibility works.”

Even conceptually, however, the notion of the schedule and the key stages in the lifecycle of every games project, regardless of the relative duration of each phase or the content of the work involved, shapes to a great extent the interactions of the developers with the project and with each other. They all know that there will be a pre-development phase, a production phase, and a finishing phase in every project and within these there will be certain inputs and contributions they will have to make. In turn, these broader phases are composed by other commonly understood and to a great degree inviolable key stages, like for example alpha testing, beta testing, console standardisation and certification, and the eventual graduation of the game from the “master version” to “gold master”, which is effectively the game that is played by people “in their sitting rooms” and than one can find on the shelves of the shops.

While the concept book and game design document are much more clearly discernable as boundary objects in the way proposed by Star and Griesemer, it is the non-representational capabilities of the drawings and visual representations (models, sculptures, props, books, TV series, films, other games) that enable the circulation among the developers of the “vision and emotion” of the game being built that provide the most interesting insights from a theoretical perspective.

Although the research is still at a relatively early stage and the analysis presented here is based on a preliminary assessment of the evidence collected from only one of the three sites to be studied, some important theoretical points are raised from the first iteration in confronting the material collected with the literature relating to boundary objects and their importance in collaborative work.
The central role played by visual representations, not in terms of the representational work they do but in terms of circulating among the developers involved often intangible aspects of the game such as its “vision” and the emotional response it must produce, raises an important question for the notion of boundary objects. This is because the original conceptualisation of the boundary object as articulated by Star focused on rather narrow representational aspects of these objects and the systems of representation they are part of.

Furthermore, something that became quickly evident even during such an early phase of this research project was that while objects and artefacts clearly had very important roles to play in the collaboration involved in the design and development of the computer games, it was much more difficult to discern clear-cut boundaries between disciplines or particular social groups in the setting being investigated, despite the formal categorisation of art, design, and programming used by the company itself.

While Star acknowledges that “no representation ... is either complete or permanent”, but rather that “any description is a snapshot of historical processes in which differing viewpoints, local contingencies and multiple interests have been temporarily reconciled” (Gerson and Star 1986), her formulation of boundary objects depends on a fairly stable view of roles, identities, and social groups. The notion of boundary objects depends on boundaries to have relevance, but the boundaries encountered in this research so far are not stable and ‘given’ but constantly shifting and changing. This critique of the original conceptualisation of boundary objects is made, if in a rather oblique way, by Fujimura (Fujimura 1992), who writes in a footnote:

“I do not assume that social worlds, e.g. disciplines, are stable entities in nature or society … [and] that disciplinary boundaries are also constructed and therefore can be destabilized.”

It is important to keep in mind – as Fujimura suggests – the location of Star’s work in the field of "symbolic interactionism" and the interest of this field in how people act toward things based on the meaning those things have for them. From the way individuals interacted with these key objects identified in the case study presented we can say that social groups and disciplines are not pre-existing but formed and dynamically sustained out of the interactions of individuals and their embodied knowledge and skills through these important objects.

Finally, it is important to note that it was not just objects that were found to have an important role to play in spanning the shifting and fluid boundaries encountered in this setting. Individuals in the form of the “technical artists” mentioned by one of the interviewees when describing the experienced and senior ‘leads’ of the studio, had vital boundary spanning roles to perform. As the work on interdisciplinary collaboration in IS development by Levina (Levina 2005) suggests, participants’ diverse backgrounds constitute powerful resources which help to shape the mode of collaborative actions and design.

5 CONCLUDING REMARKS

Drawing on data from a case study in a computer games studio we explored boundary objects in relation to the compelling, sensory and entertainment-centred game-playing practices that inform computer games design and development. Our analysis of this case study – the first of three such cases – highlights three set of objects found to have important boundary crossing roles in the design and development of computer games in this setting.

Paying particular attention to the importance of sensory and experiential characteristics in computer games and the need to bring together during their development diverse skills and expertise from art, design, and computer programming, the paper assessed the relevance of the notion of boundary objects in relation to the empirical evidence collected. While the importance of objects in the kind of collaborative work involved in the development of computer games was affirmed, the centrality given to the representational aspects of boundary objects in the original conceptualisation developed by Star
was found to be limiting (Nicolini, Mengis et al. 2008). In addition, the boundaries of social groups and the worlds they delineate were found to be fuzzy and shifting and far from stable in the way that the notion of a discipline would presume, with many individuals able to be part of more than one group and participate in more than one system of representations.

User experience and sensory and aesthetic considerations are increasingly taken into account in the design and development of software, hardware, and information systems (Bertelsen, Petersen et al. 2004; Fishwick 2006; Floyd, Jones et al. 2007), so acknowledging and understanding the multiplicity of roles that objects that participate in the interactions of social groups or even of the individuals involved in such forms of collaborative work is crucial. While the original notion of boundary objects developed by Star was important in terms of highlighting the role of objects in collaborative work and how common representational spaces are achieved in such situations, there is more than representation that these objects are involved in (Thrift 2007).

In a small way the games developers studied in this research and the importance to their collaboration of “the vision and emotion” of a game and how they deal with these elusive and difficult to represent entities in practice provide a glimpse of why the kind of research programme advocated by Ciborra (Ciborra and Willcocks 2006) “in which inner life is as important as surrounding circumstances, where the pre-theoretical is preserved by giving space to the moods, emotions and dispositions not linked to thinking”, is of such relevance.

Following on from these findings from the first case study, investigative effort over the next case studies will be directed towards trying to understand the techniques through which intuitive and difficult to represent aspects of the objects being realised through such processes of design and development – such as, for example, their sensory, aesthetic, or emotional appeal – are accessed and shared by those involved in their production.

Acknowledgements:

Financial support from the UK Economic and Social Research Council (Grant reference: RES-000-22-2809) for this research project is gratefully acknowledged.

References

Service Analysis - A Critical Assessment of the State of the Art

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0412.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Service oriented architecture (SOA), Information systems analysis and design, Literature review, IT Business Alignment / Value</td>
</tr>
</tbody>
</table>
SERVICE ANALYSIS –
A CRITICAL ASSESSMENT OF THE STATE OF THE ART

Thomas Kohlborn, Business Process Management Group, Information Systems Program,
Faculty of Science and Technology, Queensland University of Technology, Brisbane,
Australia, t.kohlborn@qut.edu.au

Axel Korthaus, Business Process Management Group, Information Systems Program, Faculty
of Science and Technology, Queensland University of Technology, Brisbane, Australia,
axel.korthaus@qut.edu.au

Taizan Chan, Information Technology Cluster, Information Systems Program, Faculty of
Science and Technology, Queensland University of Technology, Brisbane, Australia,
t.chan@qut.edu.au

Michael Rosemann, Business Process Management Group, Information Systems Program,
Faculty of Science and Technology, Queensland University of Technology, Brisbane,
Australia, m.rosemann@qut.edu.au

Abstract

The service-orientation paradigm has not only become prevalent in the software systems domain in
recent years, but is also increasingly applied on the business level to restructure organisational
capabilities. In this paper, we present the results of an extensive literature review of 30 approaches
related to service identification and analysis for both domains. Based on the consolidation of a
superset of comparison criteria for service-oriented methodologies found in related literature, we
compare and evaluate the different characteristics of service engineering methods with a focus on
service analysis. Although a close business and IT alignment is regarded as one of the core beneficial
promises of service-orientation, our analysis suggests that there is a lack of unified, comprehensive
methodology for service identification and analysis integrating and addressing both domains. Thus,
we discuss how our results can inform directions for future research in this area.

Keywords: business-IT alignment, service analysis, service identification, service orientation.
1 INTRODUCTION AND MOTIVATION

The emerging and accelerating trend towards business networks of collaborating business partners and the anticipation of entire service ecosystems comprising dedicated service marketplaces increase the need for enterprises to engage in service-orientated approaches to be able to gain the flexibility and agility required to succeed in this changing environment (Cherbakov et al. 2005, Sanz, Nayak and Becker 2006). It is no longer sufficient to apply the paradigm of service-orientation to the advancement of technical software architectures only. Instead, entire business architectures are about to be restructured from a services point of view in order to allow flexible and reconfigurable collaboration scenarios in business networks. While services in general can be seen as abstract resources that represent consumable capabilities (W3C 2004) offered by a service provider who performs specific actions on behalf of a service consumer at a certain point in time and place and through some channel (Dumas et al. 2001), to be more precise, we thus have to differentiate between the following two basic views of services:

- First, service-orientation on the business level enables organisations to expose and offer operations as business services to business partners in order to facilitate on-demand collaboration opportunities. A business service is the outcome of a specific “chunk of operation” that is performed by an organisation (Sanz, Nayak and Becker 2006). These services can have different levels of granularity ranging from comprehensive offerings (e.g. purchasing services) to fine granular services (e.g. address verifications) (Bieberstein et al. 2005).

- Second, to enable business services and thus support the agility of organisations, service-orientation on the technical level fosters the utilisation of software services and enables a close business and IT alignment (Cherbakov et al. 2005). Software services expose application functionalities that can be re-used and composed based on business needs. Hence, a software service supports the execution of a business service.

These two different views are both addressed by the Organization for the Advancement of Structured Information Standards (OASIS) in their definition of Service-Oriented Architecture (SOA) as a “paradigm for organising and utilising distributed capabilities that may be under the control of different ownership domains” (OASIS 2006). Since these capabilities can relate to business functions and to application functionalities as well, this definition supports a holistic SOA view.

As a consequence of the proliferation of the service idea on both the business and the software level, there is now a demand for service engineering methodologies that cover both business and software services and provide an integrated, holistic approach to ensure business and IT alignment and agility. Service engineering in general is still regarded as a research challenge in the literature about current SOA research roadmaps (e.g. Papazoglou et al. 2007, Kontogiannis 2007).

In our research, we aim at identifying an existing or developing a derived methodology for service analysis that accounts for the holistic view of service-orientation with its two domains of application, the business and the software domain. Service analysis is one of the earliest phases in a service engineering process, which covers the whole lifecycle of a service. It is of particular importance, as any errors made during this phase can flow through to and build up in the design and implementation phases, which results in increased cost due to necessary rework (Inaganti and Behara 2007). Not surprisingly, the two basic views of services introduced above are also reflected in the propositions about the scope of service analysis (see, for example, the different definitions by Marks and Bell 2006, p. 58, and Papazoglou and van den Heuvel 2006, p. 417). In this paper, we view service analysis as a comprehensive phase where the concept of service-orientation is applied to analyse the capabilities provided by an organisation and to consequently identify services that are currently supported and could be supported by IT. This notion includes the analysis of the impact of service-orientation on the
business level as well as on the technical level, which is congruent with the notion and scope of SOA introduced earlier.

As a first step towards achieving our higher research goal, the objective in this paper is to begin with a contemporary overview and comparison of existing major service analysis-related service engineering methods covering business and technical viewpoints in order to analyse their characteristics in detail and to evaluate to which extent these approaches are ready for the emerging requirements as stated above. Such a comparison and evaluation will not only point out possible gaps and necessary directions for research, it will also provide a general overview of the methods that can be chosen to start a SOA endeavour.

This paper is structured as follows. The second chapter will provide an overview of related work in the area of service-orientation methods comparison. Subsequently, we will present the criteria that have been used to compare and analyse extant approaches for service analysis and we will define the composition of our sample in the third chapter. The fourth chapter represents the actual comparison and the evaluation of the approaches regarding the selected criteria. Subsequently, we will discuss the findings and directions for further research. The final chapter will provide a conclusion.

2 RELATED WORK

The domain of service analysis is still rather young. Although the number of service-oriented methods has grown significantly, our literature review yielded only a rather narrow set of papers dedicated to the comparison of existing service analysis approaches. Moreover, these papers exclusively focus on the analysis of approaches related to software services and are very limited in regard to the number of methods compared. We essentially identified three papers on the topic (namely Klose, Knackstedt and Beverungen 2007, Kohlmann and Alt 2007 and Ramollari, Dranidis and Simons 2007) that provided sets of criteria for the comparison of service engineering methods with a focus on service analysis. For an overview of these criteria, the reader may refer to Table 1, which is presented and discussed in the next section.

Klose, Knackstedt and Beverungen (2007) not only compare nine existing approaches and, as a result, note the lack of a unified method for identifying services, but also present their own proposed method, thus adding to the set of existing methods. The objective of their comparison of different methods is limited to an overview of the width and depth of a few existing methods. Kohlmann and Alt (2007) compare six methods related to service analysis and modelling, particularly focusing on their criteria on the set of activities relevant for those methods. Finally, Ramollari, Dranidis and Simons (2007) compare ten methods, which are related to the development of SOAs in general. Thus, the analysed methodologies have typically a wider scope than just service analysis.

The authors of these publications conducted their method comparisons for different reasons, so that the employed criteria are not completely congruent. For example, Kohlmann and Alt (2007) seem to have chosen criteria that reflect the strengths of their own proposed method, while Ramollari, Dranidis and Simons (2007) provide a more general overview of service-oriented development methodologies ideally covering the complete life cycle of a service.

For this paper, we have consolidated the criteria already used in the three analysed related research publications based on a clustering process (see next section) to provide a comprehensive set of criteria for the comparison of service analysis approaches for both business and software services, thus catering to the underlying comprehensive notion of SOA and services as introduced previously.
3 SELECTION OF COMPARISON CRITERIA AND SAMPLE

3.1 Selecting the criteria

To be as neutral as possible and to increase the validity of our research by building on the existing body of knowledge, we decided to make the superset of the criteria used in the three analysed related research papers the basis for the set of criteria to be used in our more comprehensive method comparison. As there were overlaps in the original criteria, we had to consolidate the superset based on semantic similarities. This task was conducted independently by two different coders and resulted in clusters that were almost 90% identical. In a discussion process, the differences were debated and resolved, thus leading to Table 1, which not only shows all the criteria used in the three related papers, but also their grouping in the new clusters and the labels assigned to these clusters. These labels eventually represent the criteria that we use in our own comparison. In the following, we will describe our criteria, briefly explain the grouping of the criteria from the analysed papers as reflected in the rows of Table 1 and also define symbols and abbreviations used in our comparison overview in Table 2.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA concept</td>
<td>• Employed SOA concept</td>
<td>• Business- or technical-driven service identification</td>
<td>• Delivery strategy</td>
</tr>
<tr>
<td>Delivery strategy for SOA</td>
<td>• Background and starting point</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lifecycle coverage</td>
<td>• Covering of SOA design phases</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree of prescription</td>
<td>•</td>
<td>• Degree of prescription</td>
<td>• Process agility</td>
</tr>
<tr>
<td>Accessibility and validity</td>
<td>• Documentation of the method</td>
<td>• Availability</td>
<td>• Industrial application</td>
</tr>
<tr>
<td>Adoption of existing processes/approaches/techniques/notations</td>
<td>• Proposal of IT criteria for service identification</td>
<td>• Domain analysis</td>
<td>• Adoption of existing processes / techniques / notations</td>
</tr>
<tr>
<td>• Application of process models for service identification</td>
<td>•</td>
<td>• Visualisation with service landscapes/maps</td>
<td></td>
</tr>
<tr>
<td>Regard to stakeholders</td>
<td>• Regard to stakeholders for service identification</td>
<td></td>
<td>• Supported roles</td>
</tr>
<tr>
<td>Service classification and clustering</td>
<td>• Service hierarchies and classification scheme</td>
<td>• Examination of service cut</td>
<td>•</td>
</tr>
<tr>
<td>Additional characteristics</td>
<td>• Alignment with sourcing strategy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Clustering of previously used criteria to compare service analysis-related methods

**SOA concept:** To analyse whether both service concepts (business and software services) are supported, we selected SOA concept as a criterion reflecting whether an approach’s primary focus is on the derivation of business services (BS), software services (SS) or both (BS/SS). Klose, Knackstedt and Beverungen (2007) have utilised a similar criterion, but as their focus is restricted on software services only and, in addition, it is not clear what the foundation and the possible values of their criterion actually are, we abstracted from these specific values and only analyse if the SOA concept rather refers to business services or software services.

**Delivery strategy for SOA:** To address organisation-specific requirements (e.g. the need to leverage existing legacy systems), different delivery strategies for a SOA are conceivable. Hence, we need to analyse whether a particular approach supports the top-down strategy (T), where services are derived based on the analysis of business requirements (Erl 2005), the bottom-up strategy (B), which focuses on the derivation of services based on an analysis of legacy systems on an as-needed basis (Sneed...
or the meet-in-the-middle strategy (M) that combines the other two strategies (Inaganti and Behara 2007). The delivery strategy criterion was included to reflect which strategy an approach recommends to implement a SOA. To provide information about the particular starting point and focus of an approach, we included a textual comment in our analysis. Klose, Knackstedt and Beverungen (2007) call their related criterion ‘Background and starting point’, Kohlmann and Alt (2007) roughly differentiate between technical and business-driven identification and Ramollari, Dranidis and Simons (2007) just distinguish the three basic SOA delivery strategies mentioned above.

**Lifecycle coverage**: While some proposed SOA development approaches aim at offering support for the full SOA lifecycle, other approaches are more focused on a subset of the activities to be performed in the life of a SOA. This criterion has been derived by combining the criteria ‘Lifecycle coverage’ used by Ramollari, Dranidis and Simons (2007) and ‘Covering of SOA design phases’ as proposed by Klose, Knackstedt and Beverungen (2007). In our analysis, we use a trivalent scale (0, +, ++) with the following semantics: 0 stands for methods that focus on service identification and analysis only, while + represents methods with a service analysis and design focus and ++ finally marks more comprehensive approaches that include phases like implementation etc.

**Degree of prescription**: A service analysis method can be rather prescriptive and define a rigid, heavy-weight process with lots of details, or it can describe a more lightweight, flexible, less structured process that is adaptable and allows for an agile approach. This criterion is based on the ‘Degree of prescription’ criterion and the ‘Process agility’ criterion both introduced by Ramollari, Dranidis and Simons (2007), as we considered the two to be highly related. In our analysis, we use a trivalent scale (0, +, ++) with the following semantics: 0 stands for methods that are very lightweight, while + represents methods with a moderate degree of prescription and ++ marks highly prescriptive approaches.

**Accessibility and validity**: To be useful, a service analysis approach should not only be well documented, but the documentation must also be accessible, and the validity of the approach should be made clear. The documentation should provide many details, examples, ideas, case studies etc. to provide useful guidance in practice. Sometimes, methodologies proposed by vendors or industry players are proprietary. Typically, detailed knowledge about these approaches cannot be easily accessed, whereas non-proprietary approaches are openly available. Moreover, this criterion captures whether an approach has been validated or illustrated by presenting real case studies, whether it uses fictitious examples only or even remains on a purely theoretical level without any examples at all. Klose, Knackstedt and Beverungen (2007) (‘Documentation of the method’) as well as Ramollari, Dranidis and Simons (2007) (‘Availability’ and ‘Industrial application’) apply corresponding criteria for their comparisons. In our analysis, we used a three-valued tuple to describe the documentation (textual comment), the availability (trivalent scale with 0 standing for a proprietary, not openly available method, + representing a method that is at least partially documented for public use, e.g. in the form of papers about single activities that constitute the method, and ++ marks a fully open method) and the validation, e.g. in form of industry case studies etc. (textual comment).

**Adoption of existing processes/approaches/techniques/notations**: The proposed service analysis methodologies might utilise already existing techniques, procedures and notations that can serve as a foundation for the approach. Klose, Knackstedt and Beverungen (2007), for example, use their ‘Application of process models for service identification’ criterion to explicitly analyse if methods consider process models as input. Moreover, they introduce the ‘Proposal of IT criteria for service identification’ criterion to examine whether existing IT criteria (e.g. design principles, such as high cohesion and loose coupling) are proposed for service identification. Under the umbrella of their ‘Adoption of existing processes/techniques/notation’ criterion, Ramollari, Dranidis and Simons (2007) similarly analyse if methods propose reusing proven existing processes, techniques and standardised notations. Also, a number of criteria used by Kohlmann and Alt (2007) can be assigned to this cluster. For example, they analyse whether an approach comprises a ‘Domain analysis’, whether there is an ‘Alignment with process models’ and whether the emerging SOA architecture is illustrated using a...
‘Visualisation with service landscapes/maps’. In our own comparison, we describe the results for this criterion in the form of a textual comment to cope with the variety of possible result values.

**Regarding stakeholders:** As service analysis methods should address the requirements of potential stakeholders regarding services, Klose, Knackstedt and Beverungen (2007) (‘Regard to stakeholders for service analysis’) as well as Ramollari, Dranidis and Simons (2007) (‘Supported roles’) include respective criteria in their analyses. This criterion particularly reflects if the perspective of the service consumer is included or if solely the perspective of the service provider is addressed. To be flexible in our analysis, we use a textual comment to describe the value of this criterion.

**Service classification and clustering:** The criterion service classification and clustering describes if the method distinguishes different kinds of services. As the approaches do not use the same terminology for service types, we will only compare the approaches based on the number of different service types as an indication of the level of detail of each approach, similar to Klose, Knackstedt and Beverungen (2007) with their ‘Service hierarchies and classification scheme’ criterion. Kohlmann and Alt (2007) address this aspect implicitly by using the criteria ‘Service specification’ and ‘Examination of service cut’. They also introduce a ‘Service clustering’ criterion, which can be seen as a grouping mechanism above the level of individual services. In our analysis, we use a simple trivalent scale (0, +, ++) to indicate if there is only one single service concept in the method (0), if the method distinguishes between different service types but does not provide details (+) or if the method includes a detailed definition of the different service types (++).

**Additional characteristics:** This criterion is a placeholder for any other important characteristics of the analysed methods that seem important enough to be pointed out. For example, Kohlmann’s and Alt’s (2007) analysis if the respective methods postulate alignment with the sourcing strategy can be an aspect that is reflected here among others. We use a textual comment again to express aspects of interest in this category in a flexible way.

Having detailed the different criteria we employed to compare the various service engineering methods with a service analysis focus, the next paragraph will explain the selection process that was used to compile the sample of methods to be analysed.

### 3.2 Selecting the service analysis-related service engineering methods

After having defined the criteria for the comparison of extant approaches, we had to identify those prominent service analysis-related methods that were the most appropriate candidates for our analysis. We did not aim at a representative sample but rather at a comprehensive mix of approaches reflecting the broad spectrum of different characteristics. To this end, we made sure to not only include academic work, but also approaches developed by the largest providers of packaged business applications such as IBM, Microsoft and SAP (Genovese 2007). We ensured that the sample included both top-down and bottom-up approaches. Another requirement was that each approach to be included explicitly referred to the concept of service-orientation. The sources of the methods are manifold as they range from journals (e.g. Papazoglou and van den Heuvel 2006) to conferences (e.g. Klose, Knackstedt and Beverungen 2007), books (e.g. Erl 2005) and white papers (e.g. SAP 2005). The final sample of analysed extant approaches can be found in the first column of Table 2, which also shows the characteristics of each approach with regard to the selected criteria. The ordering of the list of approaches in Table 2 allows the visual identification of different classes of approaches according to the underlying SOA concept (BS, BS/SS, and SS), which are separated by thicker lines in the table. For the category of approaches that are only concerned with software services (SS), subclasses can be identified depending on the delivery strategy for SOA (T, T/M, M, M/B, or B) separated by dashed lines.
4 METHOD COMPARISON AND EVALUATION

The comparison of the 30 extant service engineering methods as reflected in Table 2 was conducted independently by two coders, whereby the second coder restricted the analysis to a random control sample. The results of the coding process of the control sample were consistent with the results of the original coder. In the following, we describe our observations that resulted from the analysis of the 30 methods for each criterion used in the comparison process. The criterion “Additional characteristics” will not be elaborated on as it is assumed that the description provided in Table 2 is self-explanatory.

SOA concept: SOAs containing primarily business services are less prevalent than SOAs for IT infrastructure. Jones (2006), OASIS (2005) and Sehmi and Schwenger (2006a) propose approaches that do not directly apply to the concept of a business service. Nonetheless, the underlying concepts can be adopted for the identification of business services. Flaxer and Nigam (2004) and IBM (2005) explicitly define business services, but a detailed approach for the identification of these services is missing. Kaabi, Souveyet and Rolland (2004) identify business services based on goal-modelling, which can then be supported by software services.

<table>
<thead>
<tr>
<th>Approaches</th>
<th>SOA concept</th>
<th>Delivery strategy for SOA</th>
<th>Lifecycle coverage</th>
<th>Degree of prescription</th>
<th>Utility</th>
<th>Adoptions of existing techniques/processes</th>
<th>Regardes to stakeholders</th>
<th>Service classification &amp; clustering</th>
<th>Additional characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>KS/BS/BS</td>
<td>(BS/BS/BS)</td>
<td>(BS/BS/BS)</td>
<td>(BS/BS/BS)</td>
<td>(BS/BS/BS)</td>
<td>(BS/BS/BS)</td>
<td>(BS/BS/BS)</td>
<td>(BS/BS/BS)</td>
<td>(BS/BS/BS)</td>
<td>(BS/BS/BS)</td>
</tr>
<tr>
<td>Jones (2006)</td>
<td>BS</td>
<td>T, domain decomposition</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>Decomposition</td>
<td>External stakeholder capabilities</td>
<td>++</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>IBM (2005)</td>
<td>BS</td>
<td>T, component decomposition</td>
<td>++</td>
<td>examples, 0, n.a.</td>
<td>Decomposition</td>
<td>Customer part of analysis</td>
<td>0</td>
<td>Focus on business components</td>
<td></td>
</tr>
<tr>
<td>Flaxer, Nigam (2004)</td>
<td>BS</td>
<td>T, business entities</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>0</td>
<td>Focus on business components</td>
<td></td>
</tr>
<tr>
<td>OASIS (2005)</td>
<td>BS/BS</td>
<td>T, capability decomposition</td>
<td>++</td>
<td>examples, 0, n.a.</td>
<td>Decomposition</td>
<td>External stakeholder capabilities</td>
<td>n.a.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seeling (2005)</td>
<td>BS</td>
<td>T, process models</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Bell (2005)</td>
<td>BS</td>
<td>T, attribute analysis</td>
<td>++</td>
<td>examples, 0, n.a.</td>
<td>Own visualisation</td>
<td>Service container addressed</td>
<td>n.a.</td>
<td>Focus on service modelling</td>
<td></td>
</tr>
<tr>
<td>Adamopoulos (2003)</td>
<td>BS</td>
<td>T, use cases</td>
<td>++</td>
<td>case study, 0, case study</td>
<td>Use cases</td>
<td>Consumer as part of use cases</td>
<td>0</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Kim, Doh (2007)</td>
<td>BS/BS</td>
<td>T, use cases</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>0</td>
<td>Source</td>
<td></td>
</tr>
<tr>
<td>Klose et al. (2007)</td>
<td>BS/BS</td>
<td>T, use cases</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Flaxer, Nigam (2004)</td>
<td>BS</td>
<td>T, process models</td>
<td>0</td>
<td>case study, 0, case study</td>
<td>Task tree generation</td>
<td>Visualization with service maps</td>
<td>0</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Bell (2005)</td>
<td>BS</td>
<td>T, process models</td>
<td>++</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Chun, Yi (2006)</td>
<td>BS/BS/BS</td>
<td>T, process models</td>
<td>++</td>
<td>examples, 0, n.a.</td>
<td>Own visualisation</td>
<td>Service container addressed</td>
<td>0</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>ERP (2005)</td>
<td>BS</td>
<td>T, process models/IT</td>
<td>++</td>
<td>case study, 0, case study</td>
<td>Use cases</td>
<td>Consumer as part of use cases</td>
<td>0</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>ERP (2006)</td>
<td>BS</td>
<td>T, process models/IT</td>
<td>++</td>
<td>case study, 0, case study</td>
<td>Use cases</td>
<td>Consumer as part of use cases</td>
<td>0</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Gold-Bernstein, Hübner (2006)</td>
<td>BS</td>
<td>T, business events/IT</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Sun (2006)</td>
<td>BS</td>
<td>T, use cases</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Sato et al. (2007)</td>
<td>BS/BS</td>
<td>T, use cases</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Ernold et al. (2007)</td>
<td>BS/BS</td>
<td>T, process models/IT</td>
<td>0</td>
<td>case study, 0, case study</td>
<td>Use cases</td>
<td>Consumer as part of use cases</td>
<td>0</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>SAP (2005)</td>
<td>BS</td>
<td>T, process models/IT</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Quek et al. (2004)</td>
<td>BS</td>
<td>T, process models/IT</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Allen (2007)</td>
<td>BS</td>
<td>T, business requirements/IT</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Zimmermann, Kraus, Güt (2006)</td>
<td>BS</td>
<td>T, domain and processes/IT analysis</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Spartanov et al. (2006)</td>
<td>BS</td>
<td>M, use cases/IT</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Marks, Bell (2006)</td>
<td>BS</td>
<td>M, multiple starting points</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Chang, Kim (2007)</td>
<td>BS</td>
<td>M, processes/IT</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Amanat (2004)</td>
<td>BS</td>
<td>M, domain analysis/IT</td>
<td>++</td>
<td>examples, 0, n.a.</td>
<td>goal modelling</td>
<td>Service provider addressed</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Papazoglou et al. (2006)</td>
<td>BS</td>
<td>M, process models/IT</td>
<td>++</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Chen et al. (2005)</td>
<td>BS</td>
<td>M, domain analysis/IT</td>
<td>++</td>
<td>examples, 0, n.a.</td>
<td>Feature analysis</td>
<td>Service provider addressed</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Nadhan (2004)</td>
<td>BS</td>
<td>B, IT analysis</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Rahmani et al. (2005)</td>
<td>BS</td>
<td>B, class diagram</td>
<td>0</td>
<td>examples, 0, n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Sneed (2006)</td>
<td>BS</td>
<td>B, source code</td>
<td>++</td>
<td>case study, 0, case study</td>
<td>n.a.</td>
<td>n.a.</td>
<td>0</td>
<td>n.a.</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Comparison of the 30 approaches
All other approaches focus on the derivation of software services, although the term business service is used to distinguish between services that encapsulate business logic and services that encapsulate application logic.

**Delivery strategy for SOA:** The delivery strategy for SOA is partially dependent on the underlying SOA concept. Approaches, which address the analysis of business services, postulate a top-down strategy for the delivery of services. Regarding the scope and depth of these approaches, none of them provides a detailed description of how to include existing (legacy) systems into the analysis. For example, Sehmi and Schwegler (2006a) propose a pure top-down approach that describes how a business model can be implemented using software services. As their method has been incorporated in Microsoft's 'Motion Methodology' (Merrifield and Tobey 2006), all details are not openly available. The starting points for the business service analysis vary widely. Jones (2006) and OASIS (2005) postulate a method that does not necessarily rely on any models or documentations, but on the collaborative analysis of the business of an organisation. IBM (2005) addresses business services as provisions of business components, whereas Flaxer and Nigam (2004) propose to analyse business entities to identify business components and business services subsequently. A bottom-up strategy for business services could not be identified.

Approaches addressing the derivation of software services postulate one of the three described delivery strategies. The top-down approach is supposed to derive a high quality SOA that is built on well-designed services and service compositions. However, depending on the size of the company and on the scope of the SOA initiative, a top-down strategy may consume significant resources, such as money and time, without showing an immediate outcome, since the upfront analysis has to be conducted before actually deriving services (Erl 2005; Gold-Bernstein and Ruh 2004).

Contrarily, a pure bottom-up strategy to deliver software services typically comprises activities that analyse existing legacy systems in order to define fine-grained services that can be linked to business processes and business requirements (Sneed 2006). Hereby, one can distinguish between two types of bottom-up analyses. Non-invasive legacy approaches encompass methods that do not change the structure of the legacy code (Nadhan 2004; Al Belushi and Baghdai 2007). They propose to build wrappers around the functionalities and components of the legacy system, so that they can be used in a service-oriented environment. Invasive legacy approaches aim at self-contained software services that encapsulate the functionalities provided by the legacy systems by restructuring the respective legacy code (Chen et al. 2005; Sneed 2006; Zhang, Liu and Yang 2005).

An interesting point is that most approaches postulate a meet-in-the-middle strategy that takes into account business requirements as well as existing legacy systems to combine the advantages of both strategies. Thus, the advantages of a high quality SOA have to be weighed against reality constraints applied by the legacy systems. Arsanjani (2004) as well as Zimmermann, Krogdahl and Gee (2004) describe what an overarching approach could look like, but they fail to go into detail, as their approaches are proprietary.

**Lifecycle coverage:** Regarding the lifecycle coverage, it is obvious that the scopes of the methods vary widely. While certain methods specifically focus on the analysis phase (e.g. Klose, Knackstedt and Beverungen 2007), others also address the service design phase (e.g. Kohlmann and Alt 2007) or even address the complete lifecycle (e.g. Erl 2005, Papazoglou and van den Heuvel 2006). However, as a unified, standardised life cycle for services or SOA is not prevalent in literature, authors addressing the life cycle of a service propose such a life cycle in the course of their publication.

**Degree of prescription:** Some of the approaches do not provide any structured guideline or process to derive services. These approaches give general suggestions about what to do, but do not provide information on how it should be done (e.g. Arsanjani 2004). Thus, they can be used or must be used in a flexible manner, as detailed application steps are missing. Nonetheless, most of the analysed approaches provide some kind of procedural model to identify business services or software services. The steps or phases for the identification are very approach-specific. For example, Rahman et al. (2006) propose an Model Driven Architecture (MDA) approach that focuses on the derivation of three
specific models (three phases) to identify services. Jones (2006), on the other hand, proposes four steps within his identification framework that can be used to identify services.

**Availability and validity:** There is a strong correlation between the documentation of the method within the respective publication and the application in practice. Typically, if the method has been applied in practice by conducting a case study, the case study will be described within the paper. If the method has not been applied in practice, typically just basic examples are presented.

**Adoption of existing techniques/processes:** Different methods relate to already existing concepts. Zimmermann, Krogdahl and Gee (2004) compare service-orientation to component- and object-orientation, as does Chang and Kim (2007), for example. Some methods specifically relate to UML, such as Stojanovic et al. (2004) to model services, which is widely applied in industry. Other methods utilise very specific languages or techniques that are tailored to their respective needs, e.g. Zhang, Liu and Yang (2005) or Quartel, Dijkman and van Sinderen (2004).

**Regard to stakeholder:** Stakeholders are addressed differently by the analysed methods. Some do not consider stakeholders at all (e.g. Arsanjani 2004), whereas others address stakeholders during the identification of services (e.g. Jones 2006, Sehmi and Schwegler 2006a, Kaabi et al. 2004 and Klose, Knackstedt and Beverungen 2007). However, the way the stakeholders are addressed varies amongst these methods. For example, Jones (2006) analyses the way in which external stakeholders interact with the services of an organisation, whereas Klose, Knackstedt and Beverungen (2007) analyse the stakeholder involvement in the service delivery process by examining the takeover and visibility potential of different process steps.

**Service classification:** Different authors propose different classification schemes based on the scope of their proposed approaches. Some approaches only provide guidelines to derive services in general, others distinguish between basic types of services and a few provide a classification scheme with descriptions of the objectives of each service. In the case of business services, no classification scheme based on the analysed approaches could be identified. For example, Arsanjani (2004) as well as Sehmi and Schwegler (2006a) propose decomposition approaches that can be used to identify services, but no classification for services is provided. If the SOA concept of software service is addressed by the approaches, different levels of details are observable. For example, Sneed (2006) and Zhang, Liu and Yang (2005) do not provide any classification and thus propose approaches that generally identify software services. Zimmermann, Krogdahl and Gee (2004) provide a rough differentiation between services, whereas Erl (2005), Klose, Knackstedt and Beverungen (2007), Kohlmann and Alt (2007) and SAP (2005) offer more detailed classification schemes. As the naming of the different services is not standardised, a proliferation of homonyms and synonyms can be encountered.

5 DISCUSSION

The number of publications that broach the issue of service engineering including service analysis has increased in recent years, but to the best of the authors' knowledge, a comprehensive analysis of the coverage and specific characteristics has not been sufficiently conducted yet, particularly not against the background of evaluating the fitness of extant methods in regard to requirements resulting from the trend towards an integrated application of the service paradigm in organisations on both the business and the technical level.

Our research has yielded a framework of criteria that was used to describe the characteristics of a significant sample of service analysis-related methods in detail. One benefit of this framework of criteria and the method comparison is that they support the process of method selection by an organisation, as they facilitate the definition of a set of desired method properties based on individual requirements and the subsequent evaluation of the methods against these desired characteristics. The second benefit of our results is that they are suitable to suggest ways how to address identified shortcomings. For example, based on our analysis we conjecture that there is currently a lack of methods that meet the requirement of a holistic service analysis method - the requirement that
motivated our research in the first place. A consolidated method should comprise business services as well as technical services. Thus, a comprehensive method needs to provide guidelines about how to identify business services that can subsequently be supported by software services. In the following, we discuss first suggestions how the results of our analysis can inform the design of such an improved method. Investigating the overlaps and complementary aspects of the analysed methods, one can apply the concepts behind method engineering (Odell 1996) to outline what a consolidated, comprehensive method may look like.

For example, we identified several commonalities and complementary aspects of the analysed methods (e.g. common starting points, different phases supported etc.). An earlier version of the method published by Jones (2006) has been used to serve as the foundation for OASIS (2005). These approaches can be complemented with the method proposed by Sehmi and Schwegler (2006a) and IBM (2005), thereby creating the foundation for the identification and analysis of business services. As processes are one of the underlying elements of a business service as defined by Sanz, Nayak and Becker (2006), methods proposing to use process models as a starting point for service analysis can be utilised. As our analysis pointed out, the majority of methods actually utilise process models for the analysis of services. For example, the methods proposed by Klose, Knackstedt and Beverungen (2007), Kim and Doh (2007) and Kohlmann and Alt (2007) could extent a detailed comprehensive method as the one developed by Erl (2005), as the different steps to derive services are complementary. Other methods use compatible starting points that could potentially be converted to be complementary to the process-driven approaches. For example, a use case corresponds to the activities and goals of an elementary business process and hence, the concepts behind the identification of services of use case-oriented approaches can be consolidated with the approaches focussing on process decomposition. Zimmermann, Krogdahl and Gee (2004) point out, however, that successful service analysis and design requires a focus on processes first, as it is necessary to analyse more than one system at a time to derive adequate services. The approach by Gold-Bernstein and Ruh (2004) can be adapted to business process models, as they propose that a table with events occurring during the operations of an organisation be used as the foundation for deriving services. These events can also be represented within a business process if they are brought in a sequential and logical order. Other approaches build upon one another and thus overlap to a certain extend.

Once the software services are identified, one needs to consider the actual application landscape that provides the needed functionality, as proposed by Erl (2005) and Inaganti and Behara (2007). The actual implementation might utilise methods for invasive or non-invasive code restructuring (Zhang, Liu and Yang 2005 and Nadhan 2004, respectively). Further research will need to show how a detailed consolidated method can be developed and how applicable such a method will be in practice.

6 CONCLUSION

In this paper, we have compiled a framework of comparison criteria that was used to provide an overview and analysis of 30 extant service analysis-related service engineering methods that can be used as a starting point for the journey into service-orientation. The comparison presented is the first step towards filling research gaps as proposed in recent research road maps of SOA and motivated in the introduction of this paper. The analysis supports the conjecture that there is currently a lack of a comprehensive service analysis method that comprises the identification and analysis of services on both the business and the technical level, thus ensuring a high level of business and IT alignment and agility. Our research has pointed out initial considerations for the consolidation of existing methods to achieve this goal through further research.
References


MEASUREMENTS, FEEDBACK AND EMPOWERMENT: CRITICAL SYSTEMS THEORY AS A BASIS FOR SOFTWARE PROCESS IMPROVEMENT

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0419.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Process Improvement, Behaviour change, Design Science, Politics</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
MEASUREMENTS, FEEDBACK AND EMPOWERMENT: CRITICAL SYSTEMS THINKING AS A BASIS FOR SOFTWARE PROCESS IMPROVEMENT

Øgland, Petter, Department of Informatics, University of Oslo, P.O. Box 1080 Blindern, 0316 Oslo, Norway, petterog@ifi.uio.no

Abstract

While organizations in software industry want to portray themselves as professional in terms of following standards and methods, they may also have needs for improvising and short-cutting when necessary. Such dilemmas of dual logics are sometimes internally resolved by evolving a false belief of what is done (practice) being in correspondence to what is said (standards), regardless of what an empirical investigation might show, something that can have poor business implications and also poor social implications. Particularly focusing on this latter point, the meta-methodology of total systems interventions (TSI) has been used for integrating critical systems theory with total quality management, improving social conditions in parallel with improving business processes. Although TSI is not designed for liberating organizations where nobody see themselves in need of liberation, the hypothesis of this paper is that it is possible to design quality management systems as “conflict machines”, causing sufficient social tension for more or less automatically changing “fake quality” into “real quality”. The hypothesis is investigated by applying design research in a Scandinavian public sector organization. The findings consist of statistical and interpretative evidence for the success of the approach, making a contribution to how TSI can be applied in the software industry.

Keywords: Software process improvement, total systems interventions, design research
1 INTRODUCTION

Selecting an effective strategy for designing software process improvement (SPI) systems can be a difficult problem, depending both on technical issues related to the complexity of production processes, software products, chosen set of SPI standards and practical engineering of the control system, and on social and cultural issues.

According to Brunsson et al (2000, p. 130), extensive research since the mid-1960s has shown that “there may be substantial differences between presentation and practice, between formal structures and actual operations, and between what people say and what they do. […] Actors have dual systems which are decoupled from each other; they may argue that they follow a standard while not doing so in practice. This is a phenomenon which standardizers seldom appear to notice, or at least seldom discuss seriously in public”.

Indeed, in many of the classic references on total quality management (TQM) and software process improvement (SPI) (e.g. Crosby, 1979; Deming, 1986; Juran, 1993; Humphrey, 1989), the main issue is on issues like explaining the importance of aligning quality control with business strategies, measuring customer satisfaction, implementing the methods in a technically sound manner, and getting sufficient management commitment for making sure that the system will be developed and sustained. The issue of organizations pretending to follow standardized principles of TQM or SPI just to win contracts or attract customers, and how this influences culture, is seldom addressed.

Primarily focusing on the internal politics of organizations along the vertical axis of management and labour, Flood (1993) has developed an approach to TQM based on systems thinking and critical theory. The idea is that TQM should be understood in a holistic perspective and it should be used for improving social conditions along with business processes. The idea of implementing TQM in this way was developed as an application of a more general framework called Total Systems Intervention (TSI) (Flood & Jackson, 1991; Jackson, 2000). TSI has been a topic for discussions in academic journals for systemic research (e.g. Tsoukas, 1993; Flood, 1994), has been used as a framework for doing action research (Green, 1992; Flood, 2006), and it remains an interesting topic among both researchers and practitioners in the area of TQM (Taiwo, 2001; Beckford, 2002; 2009).

However, as Beckford (2002, p. 301) points out, “the process of [critical thinking in TSI] is, of course, only useful where all parties are willing to contribute to the process and to make adaptations based upon the findings”. This does not present TSI as a promising approach when we deal with organizational dualism, implying primarily external motivation for implementing SPI. On the other hand, a well-known challenge in knowledge management systems design is that knowledge is political, meaning that changing the flow and distribution of knowledge may have consequences for explicitly and implicitly existing organizational power structures (Davenport & Prusak, 1998).

The hypothesis in this research is that (1) a political SPI system, exploiting political tensions induced by creation and distribution of knowledge, can be designed to fit with the ideas and framework of TSI, and (2) this type of political knowledge management approach can create a stable SPI process (improvement process) in an environment where there is otherwise little support for making SPI work.

In section two, a more detailed theoretical presentation of the revised TSI model will follow. This is followed by a section three concerned with the design of empirical research for exploring and evaluating the model in an empirical context of looking at how SPI is practiced among software engineers, managers and administrative staff in a Scandinavian public sector organization. The results of this investigation are reported in section four, pointing out aspects of the SPI project that were successful and aspects that failed. Finally, in section five, the empirical findings are used for evaluating the model, summarising the research in terms of pointing out how the political analysis leads to political guidelines all aimed at creating tension along the vertical and horizontal axes of the organization and making sure that the critical approach can be maintained.
2 PROCESS IMPROVEMENT BY SYSTEMS INTERVENTIONS

As SPI can be seen as TQM applied in the area of software engineering (Humphrey, 1989), generic quality management standards, such as the ISO 9000 standards and guidelines, can be used as framework for SPI. In the ISO 9000 series, there is a difference between control and improvement (Hoyle, 2006), that corresponds with what Argyris & Schön (1978) refer to as single loop learning and double loop learning. Quality control is the task of making sure that processes are carried out in compliance with standards and regulations (single loop learning). Quality improvement deals with challenging and improving such standards and regulations (double loop learning).

The diagram below illustrates the single loop learning of auditing a quality management system (QMS) used for defining a SPI system. The QMS contains all procedures and standards that describe how software engineering shall be carried out in a given organization, and this set of documents is used as a reference point \( r(t) \) when doing audits \( y(t) \). If there is a mismatch \( e(t) \) between practice \( y(t) \) and procedure \( r(t) \), that should function as a point of action \( u(t) \) for management.

![Diagram](image)

Figure 1. The basic configuration of a simple closed-loop control system (DiStefano et al, 1990, p. 16)

Mathematical theory of control systems started in electrical engineering, but has later been used both metaphorically and technically for studying both technical and social systems (Berlinski, 1976). In order to study the audit control system in figure 1 from a mathematical perspective, one would typically need a description of the relationship between \( u(t) \) and \( y(t) \) in terms of differential equations, and areas to be looked into would typically be observability, controllability and stability of the system (DiStefano et al, 1990).

2.1 Total Systems Intervention (TSI)

TSI evolved as a method for solving two problems; how to apply critical social theory in management problems based on a systems approach, and how to select the best systems methodology depending on the nature of the management problem to be solved (Jackson, 2000).

Doing TSI consists of going through three phases. First there is a phase labelled “creativity” that consists of using metaphors for trying to see to which extent various metaphors may give insights on problems and perspectives of the organizations (Morgan, 1980). Secondly, there is a phase labelled “choice” that deals with choosing an appropriate systems-based intervention methodology to suit particular characteristics found through metaphor analysis. As the makers of TSI experienced some difficulty in getting managers to understand Morgan’s way of linking metaphors and sociological paradigms, Jackson (2000) develop a simplified method called the “system of systems method”. Thirdly there is the phase of “implementation” which means that the conventional implementation strategies of Viable Systems Methodology (Beer, 1972), Soft Systems Methodology (SSM: Checkland, 1981), Critical System Heuristics (CSH: Ulrich, 1983), or whatever systems approach
chosen is implemented, given that the implementation of a chosen methodology or mix of methodologies is employed according to the logic of TSI (Jackson, 2000, pp. 368-370).

Although practitioners working with TSI as an approach for implementing TQM reported successful results, theoretical criticism came from academia, questioning the use of metaphors for framing problems, the way metaphors linked with methodologies, whether TSI practitioners could be expected to have the enormous amount of systems theory knowledge that seems to be needed, whether each of the methodologies catalogued by TSI are as different as the framework assumes etc. (Jackson, 2000, pp. 371-374). As a consequence of such questions, TSI theory keeps developing (e.g. Flood, 2006).

What TSI still seems to lack in its analysis, however, is that the people responsible for designing organizational interventions may neither represent the workers nor the managers, but could be considered a separate group. In his comments on the early days of Operational Research (OR), Beer (1968) points out that there were ongoing tensions between managers and scientists. The scientists developed an understanding of the organizational problems from a mathematical point of view, although with little practical understanding. The managers had a practical understanding developed through experience, but little conceptual understanding in terms of mathematical models. This resulted in mutual distrust.

Rather than trying to provoke labour and management through the kind of critical questions suggested by CSH, for an organization that is only pretending to be committed to SPI, it is an underlying assumption in the hypothesis stated in this paper that it may be better to design the SPI system as a conflict machine, a system that is not primarily designed for making people think less in order to work more efficiently but rather as system that contributes to the production and distribution of knowledge in ways that are likely to challenge current power structures. The idea is to build tension along the management/scientist and labour/scientist axes and use debate that results from this for making the organization focus on the management/labour axis.

In 1996, Flood and Romm redesigning TSI in the language of triple loop learning. Figure 2 illustrates the principle of how the nested loops are represented with different types of logic. The innermost loop of the system tries to answer the question of how to improve efficiency, how to reduce error rates, how to increase customer satisfaction etc. If a SPI problem is formulated in a clear and well-defined way, there is a large body of literature from OR that could and should be used. However, if the problem is neither clear nor well-defined, there is a body of literature on Problem Structuring Methods (PSM: e.g. Rosenhead & Mingers, 2001) that include various methods for organizations to reach mutual understanding of problems. SSM is one of the most used methodologies of this category. The aim of PSM is to address the question of what the problem is. Finally, there is a third loop that aims as questioning why are given problems focused in the first place. Whose interest are people serving by solving the given set of problems? Who is gaining power? Who is losing power?

![Figure 2. Triple loop learning as a way of describing TSI (Flood & Romm, 1996)](image)

The control loop in figure 1 gives a more detailed description of the first loop in figure 2, suggesting mathematical notation and logic that fits with the rationality of OR. The single loop learning aspect of quality management can be a highly difficult problem from a technical point of view, but from a social
point of view it is banal in the way that it simply consists of measuring whatever people are doing against the organizational standards of how it should be done.

Second loop learning in TQM is socially more challenging, as this loop consists of getting people in authority to interpret what is going on, try to make sense out of these interpretations and then choose targets for technical investigation and improvement. As Checkland (1981) argues, problem structuring methods such as SSM contain elements of critical thinking in the way that they challenge the views encoded in the single loop learning. Perhaps the problem being solved by the current OR loop is not the right problem to solve? An SSM analysis can bring in the perspectives of different people in the organization, producing a collective understanding of the organization and business, and cause radical changes in the understanding of what should be studied and improved.

For a SPI system to comply with the ISO 9001:2000 standard, the system must have elements of double loop learning. Not only do the ISO 9001 clauses require that process should fit with procedure, but procedures and all aspects of the QMS are to be audited and challenged on a regular basis (e.g. ISO 9001 clause no. 4.1; see: Hoyle, 2006). But, as argued by Brunsson et al (2000), the fact that an organization is required to practice double loop learning does not necessarily mean that it does, even if they are certified against ISO 9001 and being regularly checked.

The third loop in figure 2 is something that does not read naturally as a part of ISO 9001:2000, and something one would not expect to find in management consultant literature (Legge, 2002). From a business point of view, the aim of the organization may be issues like making a profit, or delivering services at high quality with low cost. On the other hand, there may be other agendas. The purpose of the third loop is to identify such agendas, help people to discover that they may be imprisoned by agendas they don’t know, don’t understand or may be unhealthy, and then to help in the process of liberation (Critical Systems Thinking, CST: Flood and Jackson, 1991), the socially critical aspect of TSI.

2.2 Adding the idea of a “SPI conflict machine” to TSI theory

As observed by Bénézech et al. (2001), the fact that the ISO 9001:2000 model can be interpreted as a double loop learning model means that a management system (or a SPI system) compliant with ISO 9001:2000 can be seen as a knowledge management system.

Although not explicitly stated by Flood (1993), one way of interpreting his use of critical theory as a foundation for TQM could be to say that quality management systems are not tools for managers to exercise command and control (e.g. Braverman, 1974), but they are rather tools for workers to gain knowledge and power over their own work situation, become experts and be able to challenge oppressive power structures.

In order to give people freedom, it is necessary that knowledge becomes visible and available. It is necessary that the implementation of TQM follows a strategy that is based on how knowledge about quality of processes and products affects the power structures of the organization. As pointed out by Davenport and Prusak (1998, p. 177), one of the main challenges in the implementation of knowledge management systems is that current power structures may be based on keeping people unaware, only letting an inner circle of people have knowledge that matters.

Reyniers and Tapiero (1995) discuss the design of contracts and the control of quality in a conflictual environment through the use of game theory. They model the situation between the supplier and producer as players in a nonzero-sum game, where the supplier can control the effort invested in the delivery of quality and the producer may or may not inspect incoming materials. In their paper, they state (p. 373) that the traditional approach to statistical quality control does not recognize that quality management often takes place in a conflictual environment.

The game design suggested in this paper, in order to make SPI work as a “conflict machine”, is to regard the QMS as the rules of the game, use quality indexes (measurements of compliance with QMS
standards) as payoff functions (“score”), turn the game from into a “perfect information game” (like chess, unlike the “imperfect information game” of bridge) by distributing the scores vertically and horizontally across the organization, aligning knowledge distribution strategies with power strategies as they may be identified through experience and analysis.

The most important design aspect of this SPI game design, however, is to make sure that the game is kept alive, in other words make sure that it does not challenge those who are capable of destroying the SPI system (“the owners of the system”; Checkland, 1981, p. 318).

3  METHODOLOGY

In order to investigate the effect of the revised version of TSI, the TSI method has been implemented and tested in an organizational setting, and a design research framework has been implemented (Simon, 1996; Hevner et al., 2004; Iivari, 2007). The design research approach is illustrated in figure 3, where the emphasis is on showing how the engineering cycle and the research cycle run in parallel through iterations of four stages, the engineering cycle applying theory while the research cycle aims at contributing theory. The ellipse marked “model” in the middle of the figure corresponds with the models discussed in the context of the theory in section 2 above with section 2.2 containing specifications for the design/artefact to be evaluated in the context of a TSI implementation.

In practical terms, the way the engineering cycle is used as a framework for research is by using the assumption of a successful design and statistically stable design process as hypothesis, considering the design process as an experiment, and look at the design evaluation as a test of hypothesis. This idea was originally formulated by Shewhart (1939) as a way of thinking about quality control in the field of mass production as a framework for scientific investigations, where he suggests the use of statistical process control (SPC) as a way of testing the hypothesis that the design process is statistically stable (“assignable causes have been eliminated”; Shewhart, 1939, p. 150).

As the experiment is based on a short number of life cycle iterations, the X-MR chart will be used as SPC-design (Clark, 1999, pp. 86-90). In addition, as there are controversies among statisticians as to how SPC links with statistical hypothesis testing (e.g. Woodall, 2000), the testing of whether the SPI process is stable or not will be done by visual judgements (Interocular Traumatic Test, ITT) as suggested by Roberts and Sergesketter (1993, p. 64).

If the test should indicate that the SPI process is perfectly stable during TSI, this should indicate that similar designs should be carried out under other conditions, such as other units within the same
organization or, preferably, other organizations. Otherwise, if the test should indicate problems, exploratory research needs to be carried out.

3.1 Population, research procedures, data collection and analysis

The setting for investigating the changed version of TSI is a software engineering unit within the IT department of a Scandinavian public sector organization. The unit to be discussed consisted of about 20 people, more or less equally distributed between males and females. The average age was about 40. Half of the people were systems designers who were also responsible for testing the system on large scale, while the other half were programmers that implemented the design in COBOL software and did detailed tests. The system followed an annual life cycle model, and all people were involved in writing and updating system documentation.

The initial version of the system was implemented in 1998, and the first few years consisted of making the solution fit with the original systems requirements. During the time of investigation, the system had reached a level of maturity where the life cycle work mostly consisted of making adjustments and improvements as requested by the ministry. Although the TSI implementation was carried out through annual cycles between 2003 and 2005, aligned with the lifecycle model, earlier project documentation was also used. Three evaluation reports were produced during the TSI experiment, as decision support for management at the time of the annual life cycle when the system went into production. The reports were written by the author of this paper, then functioning in the role of quality manager being assisted by quality coordinators for doing audits and tests. The final analysis has been conducted after the TSI projected was completed, depending to a certain degree on retrospective reflection.

4 CASE STUDY

4.1 Evolution of the SPI system from a technical perspective

As a part of the design research approach, the tools for evaluating the design evolved as a part of the evolution of the software process improvement (SPI) system. In the context of this narrative, the SPI system consisted of the strategies, policies and standards that made up to documented part of the quality management system (QMS) and the QMS practices that were carried out in terms of checking whether the documented system was being followed, recording problems, and taking corrective and preventive action.

Prior to taking contact with the unit in question, some interviews were first carried out with IT top management and representatives of the internal audit. In both these cases, it was pointed out that the QMS for this particular unit was one of the better systems of the organization. People were following the rules, as they should, and there was organizational learning in terms of continuous improvement.

Reading the reports from internal audit and generating a general overview of the documented part of the QMS, it quickly became obvious that there were no independent quality audits, at least not in the way required by the general QMS. Perhaps the QMS worked perfectly within the unit, people following standards as they should, but there was no way of knowing this without establishing a practice of doing quality audits.

A quality audit system was easy to implement as the procedure V10 for doing audits was written and just waiting for being put into use. Separate audits should be done along the software development life cycle, including the phases (1) threat and risk assessment, (2) quality planning, (3) requirements analysis, (4) analysis, (5) implementation, and (6) test. As a final checkpoint on procedure V10, the audits should be summarized and presented to management as input for the acceptance procedure N7.
As there were detailed standards for how to document and carry out work within each of the phases above, it was easy to develop simple checklists that was then used for ticking off and producing quality indexes for measuring compliance on a scale from 0 to 100 percent, defining $e(t)$ in figure 1.

The diagrams in figure 4 show the final results of plotting the average of all quality indexes year by year. Indexes for the years 2000 to 2002 were constructed by performing quality control of old documents and old software, while the more recent indexes were computed along with the design process and used as immediate feedback. Control limits for X-MR are based on values 2001-2003.

![Figure 4. Total quality index and improvement rate (annual increase in total quality index)](image)

As the X-MR diagram on the right hand side in figure 4 goes slightly above upper bounds in 2004, the null hypothesis of instability due to assignable causes cannot be rejected. The assignable cause for the anomaly is that the QMS consists of procedures and practice, and much of the practice was carried out by the quality department through their SPI design research rather than by the people in the unit themselves, indicating that increase in improvement rates may be a direct cause of the TSI approach.

Using a t-test with four degrees of freedom for comparing improvement rates before and during TSI made it possible to reject the null hypothesis of no difference in sample average at a significance level of 0.022. Recalculating control parameters for separating the process prior to intervention and during intervention, the average improvement rates would first be 0.6 before TSI and 3.6 during TSI. A SPC design based on X-MR only for the TSI period showed no indication of instability.

![Figure 5. Total quality index and improvement rate (annual increase in total quality index)](image)
When decomposing the total quality index into what was measured (the final year) for each of the six phases of the software cycle, the left hand side of figure 5 shows how the compliance between different phases differ. The right hand side of the figure shows the development of the column marked “assessment and risk analysis” on the left hand side of the figure, showing how the practice in this part of the life cycle got less and less compliant with the standard. Another indicator systematically evolving in the wrong direction (not included in the diagrams here) was a process capability indicator for measuring the ability of the production indicator to be within required limits.

4.2 Evolution of the SPI system from a social perspective

During the interview with the IT top manager prior to investigating this particular SPI system, the quality manager (author) was specifically being told to think about the organization and the systems development processes through the metaphor of a machine. The IT manager made some drawings in the air, indicating that he was thinking in terms of a hydraulic machine pumping information through the organization and out to the users and owners (ministry).

The systems approach used for structuring the problem of investigating the SPI was partly based on SSM-like principles like discussing with various people while drawing “rich pictures” in order to figure out how they differed in their understanding of the system. When doing audits of software engineering practice against the software engineering standards, this feedback was useful for understanding the gaps while not preventing gaps from being documented.

This approach, mixing conversations with measurements, became the implementation method that was followed consistently during the three years of the experiment.

**First year:** Reporting the findings upwards in the system caused emotional stir. People felt they were doing as best as they could, performing and documenting much better than what was required by the standards, and felt humiliated and demotivated by quality audits giving them scores in terms of numbers (“quality index”) and so on. They wrote a letter to the general director (head of the larger organization) and complained about the situation. The quality manager never got a copy of the letter, but was told by the IT manager that there were some misunderstandings and that quality audits should continue as planned.

**Second year:** Based on previous year’s unfortunate circumstances, it was important to be gentle yet sufficiently specific in order to help the software unit to improve. However, as new versions of documents were audited, the same errors and lack of compliance with standards remained. Little improvement. The people said they did not have time to improve, just to do what was absolutely necessary. Besides, they felt quality audits focused too much on process and forgot about the product. They said they took pride in producing products of high quality, and methods (means) for achieving such excellent results were of less importance as long as the results were good.

**Third year:** Trying to respond seriously to what people were saying, the same audits from the previous year were repeated but now expanded in order to include product quality indicators. Using statistical analysis, the quality management people found that the main production indicator for the system was rapidly declining, as could be seen from updates of the annual Cpk capability index (Deming, 1994). When presenting this finding to the systems and software people, they said this had to be the result of socio-geographical dynamics, not a phenomenon they could control. When asked about whether it would be a good idea to improve the prediction algorithm that caused the problems, they said they had no time for dealing with such issues. It was too difficult. They had been told to follow a given formula, and it was not their task to analyse whether this formula was optimal or not.

In 2005, after the study was completed, the people in this particular unit took part in rewriting software engineering standards to be used for further SPI. When the software engineering standards were updated, many of the changes in the methods and standards were of a kind that would give higher score according to the quality audit system but would make the method itself significantly worse. For instance, the process of updating requirement documents was eliminated.
5 DISCUSSION AND CONCLUSION

The development of the SPI system followed the three stages of TSI. **Creativity:** The quality manager was explicitly told to think about the organization as a machine. **Choice and implementation:** According to Flood (1993, p. 81), typical systems methodologies fitting with the machine metaphor are the classical management frameworks described by Taylor and Fayol. In order to carry out the technical part of quality audits, this was done according to the professional guidelines of quality engineering, i.e. in a machine-like manner.

However, this rigid design proved useful in making people respond and reflect upon the standards they had defined for their own work. When presenting the results of audits, a “soft systems approach” was used, in order to figure out what people were thinking, how one could solve the situation together, asking people what they would do if they were responsible for doing quality audits etc.

Unlike Flood’s attempts to ask critical questions for making people reflect, the method of trying to force people to follow their own standards by first “measuring the facts” and then doing SSM-like discussion, worked well for creating debate. Particularly the way audits and quality reports were widely distributed made the QMS into a “conflict machine”.

5.1 The illusion of excellence as a mental prison

It is not too difficult to see how the organization was imprisoned in a belief of high quality and continuous improvement, while the measurement of practice against standard indicated low degree of compliance with internal quality standards and continuous spiralling in the wrong direction. Similar to the cases reported by Brunsson et al (2000), the organization had locked itself in a situation where there was no measurable feedback on whether they were doing the right thing or not, only good intensions, hard work, and occasional verbal feedback if production went well.

It did not improve matters that the internal audit had investigated the unit, been impressed by standards, written procedures and system documentation without going into depth when analysing how this was used. The result of this may have been twofold; firstly adding to the beliefs of the organizational unit that they were doing things as expected, and secondly enrolling the internal audit into the same “fake quality” belief system, as it would be easier to accept the happy belief that the unit was doing fine rather than to question the quality of their own audit methods.

Although not specifically mentioned in the case study, as it happened prior to the time when the TSI experiment was carried out, the national audit had done a similar overview audit to what the internal audit later did, also contributing to the general belief that everything was fine. In other words, it would almost come as a surprise if these people were not trapped in an illusion of excellence and internal bliss, like Adorno’s interpretation of Odysseus and the lotus eaters (Sherratt, 2000).

5.2 The liberation process

As explained in the story, the “fake quality” was easily revealed through quality audits as what people were saying did not fit the results of the audits. When confronted with the facts, the response was a mixture of surprise, irritation and what they described as “unfair game” of being measured and having the measurements distributed upwards and sideways in the organization. The crash between their self-perception of what they were doing and the measurements caused emotional outbursts and intense debate.

In other words, liberation through quality management was in this case a process of agony and frustration, and it remained a difficult process during all three years of the experiment. While there were no dramatic changes during the period of experimentation, there were signs of improvement in
terms of discussion and awareness of quality management issues that seemed to have been more or less totally ignored after the QMS had been documented and made into “shelfware”.

Although some of this liberation resulted in attempts to fight the quality department and attempts to change the software engineering standards for producing better score without resulting in higher quality, the main reason for applying critical theory as a foundation for SPI should not be to force people follow standards or force them do what the quality department wants them to do, but it should be to generate debate about standards and make people take responsibility for their own software engineering processes.

The fact that the software engineers got the methodology group to remove the maintenance of the requirements documents from the software engineering methodology (end of section 4.2), despite the fact that the organization was required to keep a ten year record of requirements, was an interesting example of how SPI can be a political game of accepting and distributing responsibility. Although a strategy like this would be helpful for improving the score for the software engineers, the overall impact on the organization would be a step backwards down the maturity ladder if assessments had been done through CMM or similar SPI models (Zahran, 1998).

5.3 Theoretical implications for the TSI model in software process improvement

Although the SPC diagram and statistical reasoning in section 4.1 do not contain sufficient data for making judgements that would convince a properly trained quality engineer, the results are consistent with the interpretative narrative.

The interpretative part of the study provide the key insights on what happened, although a rich case like this contains actors and actions responding and reacting in different ways. By thinking about SPI as knowledge management rather than control, the design idea of using (1) horizontal benchmarking, (2) vertical jump-reporting and (3) improvisation for preventing those in power to destroy the TSI implementation, made the SPI system work successfully as a “conflict machine”.

Although three years was a rather short period for a case like this, and the fact that the design was particularly made to fit with the internal power struggles in a hierarchical public sector organization, the design idea (section 2.2) worked well, making a theoretical contribution on how TSI can be applied in the software industry.

The overall interpretation of the case study is that the use of critical thinking as a foundation for SPI was a fruitful idea that should lead to more empirical research.

References


MOBILE INFORMATION SYSTEMS AND ORGANIZATIONAL CONTROL: A FOUCAULDIAN APPROACH

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0422.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Mobile systems, Management control, IS Philosophy, Case Study</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
MOBILE INFORMATION SYSTEMS AND ORGANIZATIONAL CONTROL: A FOUCAULDIAN APPROACH

Leclercq-Vandelannoitte, Aurélie, IESEG School of Management, 3 rue de la Digue, 59000 Lille, France, a.leclercq@ieseg.fr

Abstract

The latest advances in the field of communication tools and networks have led to the development of ‘mobile’ technologies. Mobile technologies provide workers with almost permanent access to their company’s Information System (IS) and continual availability of information in time and space, thereby contributing to the development of “mobile IS.” Mobile IS generate a reconstruction of the relationship between time and space, and act as catalysts of deeper social and human changes, leading to ambivalent effects in the field of management. More particularly, a paradox emerges with respect to mobile IS as they can be considered in turn as instruments of autonomy and freedom or control of distance activities. To what extent are the interactions between mobile IS and organizational actors liable to change the company’s control systems? In this paper, we argue that a Foucauldian approach can enrich our understanding of the evolution of organizational control linked to mobile IS. We present the results of four case studies (based on 85 semi-structured interviews and 10 days of observation). This research shows the relevance of Foucault’s conceptual framework (linking discourses, discipline and ethics) to explore the interactions between IS, the organization and individual actors, in a political perspective.

Keywords: Mobile information systems, management control, organizational change, surveillance, case study, IS philosophy (Michel Foucault)
1 INTRODUCTION

The latest advances in the field of communication networks and data processing tools have led to the development of ‘mobile’ technologies. Mobile Information Technologies (IT) cover a wide range of terminals, such as mobile phones, laptops, personal digital assistants (PDA) and tablet PCs, linked to networks comprising numerous information resources. Lyytinen & Yoo (2002) define mobile technologies as a “network of interconnected technological, social and organizational elements”, enabling mobility which is both “physical and social” for the players concerned. Mobile ITs transcend company boundaries and are frequently used in contexts outside business, reflecting the emergence of a form of ‘multicontextuality’ (Henfridsson & Lindgren 2005). They provide us with almost permanent access to a company’s Information System (IS) and continual availability of information in time and space, thereby contributing to the development of “mobile IS.” Mobile technologies generate a reconstruction of the relationship between time and space, and act as catalysts of deeper changes that go hand in hand with the use of these tools. Both the concept of work and management and the way companies operate are directly affected by the development of this culture of mobility and the technologies that are its reflection. As they revolutionize the traditional time-space aspect of organizations, they are likely to give rise to new forms of work organization. Today, work is no longer understood as a place, but rather as an activity that can be performed outside traditional spatial and temporal frameworks.

A global discourse, conveyed by the media, IT constructors, and companies, links mobile IS to new types of flexible, responsive, dynamic and non-bureaucratic organization systems. The advent of mobile IS thus accompanies managerial and organizational discourse linked to employees’ empowerment, emancipation, autonomy, delegation and flexibility. Nevertheless, the characteristics, uses and challenges mobile IS throw up for the organization and management also engender new issues with respect to control. These technologies enable a control of distance activities (Wiredu and Sorensen 2006). In consequence, a paradox emerges with respect to mobile IS as they can be considered in turn as instruments of autonomy or control. Our aim in this paper is to analyse the ambivalent effects of mobile IS in relation to hierarchical relations and control. In addition, if the advantages companies reap from the integration of mobile IS appear undeniable, it is important to keep in mind that the systems are adopted by organizational players who must familiarize themselves with these tools before using them. Mobile IS thus raise the issue of the acceptance of a new form of work organization and re-engineering the relations with the company, reflecting a process of technological, social and human change. It is therefore also important to consider how individuals use these mobile IS, how they perceive the new working conditions that combine autonomy and control, and how they may influence the evolution of systems of control. To what extent is the use of mobile IS in an organization liable to change the company’s control systems and what are the ensuing grounds for appropriation? In this paper, we argue that a Foucauldian approach can enrich our understanding of the links between organizational control and mobile IT. This paper is structured as follows. In the second section we present a literature review about the potential links between mobile IS and organizational control. The third section explores Michel Foucault’s conceptual framework. The fourth section provides a description of our research method. In the fifth section, we present the successive results of four case studies. These results are then interpreted and discussed in a sixth section. In the final section, we present the contributions and limitations of our research and give some suggestions for future research.

2 MOBILE IS AND ORGANIZATIONAL CONTROL

2.1 Mobile Information Systems and paradoxical effects

It is now widely acknowledged that mobile IS provide an innovative answer to the challenges generated by a competitive, changing and global environment that is shaped by hyper-competitiveness, where companies are subject to cost constraints and an ever greater need for reactivity with respect to their clients and partners. As Varshney (2003) points out, mobile technologies are a way of
introducing a new form of “flexibility, in terms of time and place” into organizations, and in this sense, they offer businesses promising opportunities. Employees can log onto their company’s IS and be in contact at anytime and in any place (Robey et al. 2004). The benefits are numerous, the first being an increase in individual productivity through a decrease in work constraints, greater flexibility, and reduced coordination costs. With enhanced communication and knowledge exchange, these technologies also allow information to be accessed immediately, provide improved decision-making performance and consequently greater reactivity (Davis 2002). More and more companies now provide their employees with mobile ITs, primarily field workers such as sales representatives, consultants and technicians, but also more sedentary workers.

However, the advent of mobile IS and their use within companies can lead to negative side effects which recent studies have brought to light (Robey et al. 2004, Cousins and Robey 2005, Besseyre des Horts and Isaac 2007). Demands for almost permanent availability and responsiveness appear to have developed alongside the use of mobile IS by organizations. When employees use mobile technologies, companies have access to a potential form of “digital traceability” (Robey et al. 2004), which can give rise to a certain degree of stress. Information and cognitive overload is also linked to the use of mobile technologies within business organizations. Often encompassing a notion of continual availability, the utilization of these technologies raises a number of issues with respect to infringement of private life and the breakdown of borders between private and professional life (Cousins and Robey 2005). It also leads to issues of fragmentation and interruptions at work (Davis, 2002), fostering distraction rather than time for reflection. Similarly, employees may feel oppressed by the emergence of a culture of speed and instantaneousness and a sense of permanent urgency, which obliges them to make over-hasty decisions or decisions in contexts unsuited to decision-making. In addition, as Lytinen and Yoo (2002) pointed out, several levels of analysis are affected by what they term “nomadic computing,” not only at individual level, but also at the level of the team, and, more widely, the organization. Consequently, it is likely that evolutions in employees’ space-time norms, and the lack of face-to-face interaction will impact on cooperation, cohesion, trust between colleagues, group decision-making and, more generally, interpersonal relations.

Notwithstanding, mobile technologies may be considered as particularly equivocal tools, whose effects cannot be predicted in advance, either in terms of social interaction or of company management (Arnold 2003, Jarveenpa and Lang 2005). To this effect, Cousins and Robey (2005) highlight the contrast between the expected benefits and the unexpected social impact resulting from mobile system environments. At a time when management practices are being reinvented around mobility via the porous nature of time and space, it is important to consider the issues involved in mobile IS and the renewed management style they engender.

2.2 The evolution of control systems associated with mobile IS

More particularly mobile IS raise a challenge concerning the evolution of control systems. Control can be defined as “the effort exercised by managers, not just to collect and share information, but also to use information for directive purposes with their units: the aim is to encourage or provoke a general reaction from the people who report to them” (Mintzberg 1994). A focus on the informational dimension of organizational control indicates that the latter is enabled by data processing and storing. In consequence, given their capacity to save, store and analyse information flows, information technologies are far from being neutral elements with respect to control. Furthermore mobile technologies offer both continuity and discontinuity in comparison with other generations of technologies. They perpetuate certain practices and management methods, such as task allocation, process standardization and activity control (Zuboff 1988), but the liaison opportunities they offer also pave the way for new means of communicating, exchanging information and working outside the traditional corporate space-time framework. Mobile IT revolutionize the traditional time-space aspect of organizations. For instance management today is no longer confined to the company premises but may potentially be practiced anywhere, at anytime and in unexpected contexts. Management is no longer limited to the traditional idea of time-space and no longer necessarily occurs in a context of shared action.

A paradox emerges with respect to mobile IS as they can be considered in turn as instruments of autonomy or control, freedom or servitude. While mobile IS represent a means to promote flexible, responsive, dynamic and non bureaucratic organization systems, they may also be perceived as
instruments that reinforce control and demands in terms of availability and responsiveness, as well as employee traceability, acting as a sort of “electronic lead”, that goes way beyond organizational boundaries (Sorensen and Gibson 2005, Jarvenpa and Lang 2005). Mobile IS appear as tools to enhance the independence and mobility of the workforce, but they are also symbolic of the preservation of the “hierarchical line,” even beyond the company boundaries. More specifically, mobile technologies reflect a dichotomy through the autonomy they offer and their potential as instruments of control (Zuboff 1988). The same technology can thus be considered in two different ways that are entirely contradictory. Jarvenpa and Lang (2005) highlight the different paradoxes that result from the use of mobile IS, some of which are directly linked to control and autonomy, the individual’s decision-making freedom and the constraints that influence their activities (Besseyre des Horts and Isaac 2007). The paradoxes between freedom and servitude, independence and dependency, improvisation or planning, engagement or disengagement, are all conflicting consequences that arise from the use of mobile IS, which reflect both autonomy and control (Besseyre des Horts and Isaac 2007). According to Wiredu and Sorensen (2006), mobile IS raise the question of organizational control, insofar as they play a major role in the control of distance activities.

2.3 Important issues at stake

Mobile IS are far from being neutral tools of communication and information transmission. On the contrary, they may influence systems of control and take place in political contexts, constituted of hierarchical relationships, interactions between organizational actors, power games, negotiations, and conflicts (Markus 1983). Given that mobile IS influence the very foundations of collective action, the space-time dimension of human experience and interpersonal relations, it seems necessary to analyse the political dimension, the power balance and the issues at stake during mobile IS implementation. Different stakes arise, such as the political dimension of these interactions and also attitudes when confronted with control and the technologies that sometimes support it. Given the control potential of mobile IS, it is important indeed to hone in on the attitudes and practices that individuals adopt, and the different interactions possible between individuals and mobile IS. Mobile IS implementation not only represents a technological change but also social and human ones. It is therefore also important to consider how individuals perceive the new working conditions that combine autonomy and control. Even if the individual cannot always choose to accept or reject the technology in an organisational context, he can still choose the manner in which he appropriates it, via varying levels of engagement or involvement. Individual attitudes thus have a necessary influence on systems of control linked to mobile IS implantation. Despite the fact that IS are directly involved in control issues, their effects are in no way predetermined. We should therefore avoid any “determinist perspective”, adopting instead an “emergent perspective”, which dwells on the interactions of system and context of use and offers a means of identifying the political dimension of dialectics between control and IS (Markus 1983). An emergent perspective highlights the interactions between the technological, individual and organizational choices, together with their integration in a political context involving the interactions between players and their relation to power (Markus and Robey 1988). This analysis highlights the importance of placing our research in an overarching framework that enables us to analyse the ambivalent effects and the paradoxes of mobile IS in terms of control, with a focus on the interactions between the players, the organisation and IS. It seems necessary to develop a conceptual framework to explore the relationships between individuals, the organisation and technology in the light of control issues, at the same time taking the political dimension of these interactions into consideration. As shown in the following section, our attention was drawn to the thinking developed by a particular theorist, Michel Foucault. We believe that his work provides tremendous potential for developing a new approach to technological and organizational change.

3 CONCEPTUAL FRAMEWORK: THE RELEVANCE OF A FOUCALDIAN PERSPECTIVE

This section provides an analytical framework based on Michel Foucault’s work, whose main concepts can provide in-depth insights when applied to research in management and IS.
3.1 Three conceptual entities: discourse, discipline and ethics

Foucault’s genealogical method, which focuses on three conceptual entities - “discourses”, “discipline” and “ethics” - appears highly relevant in analyzing some emerging forms of organization, closely linked to the effects of information control promoted by IS.

3.1.1 Discourse:

Foucault (1971) examined the social effects of the knowledge produced by discourses. According to Foucault, discourses both create and control the objects they claim to know. The social world is organized and normalised in specific ways through discursive practices. The Foucauldian perspective shows that discourses are far from neutral, and constitute the “crucial way” to the exercise of power. For example, in his early writings, Foucault shows how madness, prisons, the body, life, death, and above all human beings, progressively became the objects of observation and new scientific discourses, which offer “an insidious form of social control.” In this way, organizations can be seen as political arenas where discourses are manipulated to influence individuals. Several studies in organization theory focus on the discursive practices that constitute organizations as regimes of truth and discipline, and act as a powerful constraint on the organizational members (Sewell and Wilkinson 1992, Barker 1993, Knights 1997). Organizations are considered as political spaces where discourses are constituted (Barker 1993) to better control organizational members. Foucault emphasizes the deeply relational nature of power and its incarnation in discursive practices which convey representations of the organization and technology.

3.1.2 Control mechanisms and discipline:

This focus on discourse, truth and knowledge enables Foucault to develop the image of a disciplinary-based modern society. In this disciplinary society, various means, technologies and practices - such as “hierarchical observation”, “normalizing judgment”, and “dressage” - are used to govern men. Hierarchical observation, combined with division and classification practices, enable to closely supervise organizational members. Moreover, a deep process of normalization enables to establish goals, to compare individuals and to make distinctions between modern subjects. Finally, “dressage” practices render bodies and minds obedient, docile and useful. Foucault (1975) explains how the soul, conscience and thought progressively became the primary objects for punishment and rehabilitation. These practices enable to discipline and to correct abnormal behaviours. Every aspect of human life is controlled through the construction of a “micropower.” Foucault uses the metaphor of the “panopticon” developed by Bentham (1791) to represent this disciplinary power. In panoptic architecture, observers can observe all prisoners without the prisoners being aware that they are being watched. The panopticon is characterized by invisible surveillance, a depersonalization of power, an embedding of controls, and subtle coercive mechanisms. In this way, technologies can be seen as object of disciplinary power, while the development of specific discourse legitimizes its adoption and use. Many authors relied on this panopticon metaphor to show the role of IT in surveillance, though the concepts of “electronic panopticon” or “electronic eye” (Willcocks 2004; Lyon 1994).

3.1.3 Ethics and resistance:

Because of his developments regarding disciplinary power, Foucault was accused of developing his own “iron cage,” where the human subject appears passive and subject to an entire disciplinary society. This led Foucault to develop his consideration of “ethic” for and “care of the self”. Foucault (1976-1984) describes a responsible individual actor who is able to resist disciplinary practices. “Here Foucault focuses on a more active, individual subjectivity, less imprisoned in and less constructed through scientific discourse and power relations, more geared to self-knowledge supporting work of self on the self, to constitute a self-stylization able to separate from subjectification practices” (Willcocks 2004:248). Foucault considers that individuals are able to find satisfaction in constraining situations through ethics. Foucault identifies different kinds of “technologies of the self” which allow individuals to work on themselves by regulating their bodies and thought and by constructing their identity (McKinlay and Starkey 1997). Modern subjects can subvert the conditions of their own subjectivity and constitute themselves as moral agents through ethics (Willcocks 2004).
3.2 The Foucauldian perspective: toward a renewed overarching theoretical framework

The ideas developed by Foucault provide us with a conceptual framework with powerful heuristic possibilities. At the centre of Foucault’s thinking, the concept of “power-knowledge” provides us with a tool to grasp the links between his conceptual entities (discourses; power and control; and human agency). Power produces knowledge, and discourse and knowledge have power and truth effects. “Power and knowledge directly imply one another…there is no power relation without the correlative constitution of a field of knowledge, nor any knowledge that does not presuppose and constitute power relations” (Foucault 1975). This concept implies that disciplinary power permeates the social body through power-knowledge relationships. But it also reveals that power relations are not merely negative but productive as well. “Power must be analyzed as something that circulates…Power is exercised through networks, and individuals do not simply circulate in those networks: they are in a position to both submit to and exercise this power. They are never the inert or consenting targets of power; they are always its relays” (Foucault 1975). In consequence, there are no relations of power without resistance. Foucault identifies the existence of norms that condition behaviours, but meanwhile recognizes that individuals are able to resist such norms.

Foucault’s main concepts can be usefully harnessed to management research, particularly in IS research. Foucault’s ideas encompass the issues of discourse with respect to a phenomenon, control and resistance within a single conceptual framework, at the same time placing the individual at the heart of the question. Foucault endeavoured to bring out the relative character of “discourse” and truths by identifying “power/knowledge” games that both engender and drive them. ‘Man’ is not simply an object of knowledge, but is also an object of power, which is expressed in micro-physics and disciplinary technologies. Beyond these disciplinary mechanisms, however, ‘Man’ also appears as a moral agent, subject to a certain form of behaviour and motivated by a deep “care of self” ethic.

The Foucauldian approach moreover provides a novel definition of technology, allowing us to consider IT as an “electronic panopticon,” and also as technologies embedded in the micro-physics of life, power relations, discourse, and resistance moves. Willcocks, for example, develops an analysis of “behavioural and social technologies encoded in material technologies” (Willcocks 2004, p.289), while Knights and Murray (1994), developing a Foucauldian perspective, consider IS as “a set of human and non human artefacts, processes and practices, ordinarily directed toward modifying or transforming natural and social phenomena in pursuit of human purposes.” Technology mirrors a vision of the organization and the intentions of decision-makers, but its effects can never be predicted in advance. Developing a Foucauldian approach, Bloomfield suggested: “Technology does not impact on organizations or society: a change in social relations, tasks, skills and knowledge is already prefigured in the way that the technology is conceived and constructed. Machines do not control social relations: they presuppose, mediate and reinforce them” (Bloomfield 1995, p.497).

As regards our research question, the introduction of mobile IS into the workplace has given rise to certain arguments concerning the transparency provided by these technologies, the increase in autonomy and the potential for new forms of work organization to emerge. A Foucauldian perspective enables us to see beyond these discourses and understand how such technologies are employed in systems of control. The Foucauldian approach thus provides a means to analyze how certain technological resources may be used in a hierarchical relationship and what changes arise as a result of these tools in terms of organizational control and autonomy. Moreover this approach enables us to explore individual reactions and attitudes towards technology and underlying effects on systems of control. Thus, at a time when management is built around discourse, a Foucauldian perspective provides an interpretative framework that helps us understand how certain technological resources may be used in a hierarchical relationship via power-knowledge relations, and how control systems can be influenced by individual interactions. This approach informs the development of a political model of technological and organizational changes in organizations, that we will confront to reality in a qualitative empirical study (Figure 1).
4 RESEARCH METHOD

Our study has adopted a “dialogical perspective” (Leidner and Schultze 2002), which allows us to study the nature of a constructed reality as well as the impacts of power-knowledge and discipline. We developed a qualitative study of four cases, representing high diversity, to explore the empirical contributions of a Foucauldian approach to the analysis of interactions between mobile IS, individuals and organizations from the perspective of organizational control systems. This empirical research study includes different kinds of data: primary data consist of semi-directive interviews and direct field observations. Secondary data include internal documentation, meetings, and press reviews. We interviewed different levels of respondents: managers and top managers (CEO, CIO, human resource managers, operational managers, middle managers, area managers) and field workers (equipped with mobile IT). Our qualitative analysis was developed from both a deductive and an inductive perspective: on the one hand and in accordance with the deductive principle, we identified a priori the main topics (based on Foucault’s conceptual entities which informed the interview guide) related to organizational discourses linked with mobile IT deployment, impacts on the ways of governing employees (control systems, discipline, authority versus empowerment, flexibility, autonomy, decentralization, and delegation), and individual reactions (acceptance of change, resistance moves, adoption and appropriation of mobile technologies). At the same time, in accordance with the inductive perspective, other themes, as expected, emerged from the corpus of transcriptions. Different interview guides were prepared beforehand based on our literature review and in accordance with the different functions performed by respondents. Every interview began with general questions about the respondent (his or her role and responsibilities) and then covered wide-ranging, open topics linked to mobile IS implementation, impacts on systems of control and individual reactions.

A total of 85 semi-structured interviews of approximately 90 minutes were conducted in four French companies between January 2006 and November 2006. Ten days of direct in the field observation were also carried out in order to enrich our analysis. These interviews were tape-recorded and fully transcribed. We first conducted an infra-case analysis, in other words, an in-depth analyses of cases, followed by inter-case analyses, that is to say pure comparative analyses, which enabled us to identify consistencies between the cases, and to isolate recurring elements that are common to several cases and identify disparities. The interviews were subjected to a qualitative thematic analysis, using Nvivo software to reduce and codify the data. We applied a mixed and rich thematic coding system and performed textual and Boolean searches. We also employed double-coding as a means to check the reliability of our analysis.
5 RESULTS

The first case, Technoplus, involves a French family company with 160 employees specialized in retail of industrial products. In 2004, suffering from fierce competition in its sector, the company decided to improve its customer responsiveness by equipping its technical-sales staff with mobile technologies. Every sales representative was given a Blackberry, directly connected to the company’s commercial IS. This technology was presented by management as a means of increasing the employees’ responsiveness and productivity, while reinforcing their autonomy and responsibility, as all necessary information was now at their disposal from a distance. The sales manager thus expected the technical-sales representatives to become “autonomous entrepreneurs” in their respective area. The introduction of mobile IT was justified by a climate of urgency and a demand for greater customer-centred involvement. According to the CEO, the technical-sales representatives were largely involved in the development of this technological project. Nevertheless, we should stress that the technical-staff representatives were only allowed to give their opinion on minor aspects of the technological development. Furthermore, the fact that representatives were associated with the technological deployment was used by the CEO as an argument to prevent any resistance from employees. As they had been party to improving the device, they were no longer in a position to resist or express deviant opinions. In fact, mobile IT deployment was unilaterally and rapidly introduced, characterized by a highly directive style of use. Furthermore, initially legitimized by an imperative of responsiveness, mobile IT use progressively turned into an obligation, a duty of efficiency and performance for these representatives. The CEO explained that they had no other choice than to be responsive, available, efficient, and high-performing. Moreover, the introduction of mobile technologies led to some changes in the control methods. Before the development of mobile IT, control methods relied on management by objectives and commission, and was essentially based on trust. The introduction of mobile IT reinforces surveillance over technical-sales representatives, with increased visibility and follow-up, and a normalization and homogenization of behaviours. The discourse deconstruction we carried out for example reveals a metaphor of visibility in the management interviews. Representatives must fill in a sales report in real time, just after their visits. These reports directly inform the sales manager, who knows exactly how many visits have been made by representatives during a certain period of time. The sales manager thus has the possibility to compare the representatives’ performances, to identify anomalies and gaps, and to correct abnormal behaviours. This first case thus demonstrates the panoptic evolution of control systems. Mobile IT give the management control over the behaviour of their mobile staff that the company could not manage physically. The technical-sales representatives nonetheless adopt different attitudes to mobile technology, depending on their profile and experience, and have various representations of the underlying changes in control methods. Some uses are in line with senior management expectations, while others are more or less deviant. For example, some sales representatives take advantage of mobile IT to structure and organize their activities; others use mobile IT as instruments to demonstrate to their managers their involvement in the company; on the other hand, the more experienced representatives claim they don’t need such tools and deliberately neglect their sales report. In these specific cases, the company management accepts such practices as these representatives are generally good professionals. These disparate uses may reflect adherence to objectives, hostility to organizational goals, or recourse to habits. The case finally shows the capacity of representatives to resist and the identification of sources of satisfaction in a set of social relations that are constrictive by nature.

The second case, ABConstruction, involves a very large French building company (38,500 employees), where a mobile IT project entitled “Sesame” has been developed to improve data management and optimize processes. Site foremen have been equipped with Tablet PCs, directly connected to the company’s IS, in order to enter the data relative to the building site directly into the system. The Sesame project enables more reliable and rapid data tracking and is supposed to avoid double entries. The project generates other effects linked to the equipping of site foremen with new mobile IT, such as enhancing their role and repositioning them at the heart of company. Top management moreover relies on these induced effects to promote the project to site foremen who are unfamiliar with new IT. Major discourses around the site foremen’s empowerment and role enhancement are developed to promote their acceptance of the technology. A satisfaction survey
carried out among a few site foremen was widely disseminated within the company, for example, to gain the support of other site foremen, and even to generate a sense of shame among those who were on the verge of rejecting the project. This project led to major changes in processes, practices, and control systems. Before the Sesame project, site foremen had to fill in a paper report, which was then transmitted to the accounting services for control and validation. Because they knew there were multiple checks, site foremen didn’t usually pay a lot of attention to these reports. Moreover, they were directly supervised by their operators, who regularly visited building sites to check that procedures were respected. Many site foremen would therefore offload their report onto their operators. The company’s top management team progressively considered these controls and checks by administrative services and operators as a waste of time and efficiency. As the site foremen’s position in closer to the building site, they hold crucial operational information for their company, so management decided to give them the means to enter their data and report directly in the IS. The site foremen’s responsibilities and relationships with their operators consequently evolved, as they now had to manage the building site expenses. The technologies at their disposal became “disciplinary” technologies, in the sense that they introduced rigour to the practices and provided a form of distance surveillance. This technological deployment led to a transition from controlling execution to controlling objectives/results (project profitability). The control systems changed from close, direct supervision to self-control procedures for site foremen who were firmly invited to check the data they entered in the system themselves, including the amounts for expenses. The empowerment and greater autonomy of the site foremen thus engenders new obligations and constraints for them that represent an evolution which is at times difficult to accept. As a consequence, the site foremen adopt different attitudes. Some of them seem to appreciate their management team’s initiative and feel valued and recognized by the hierarchy. Others tend to resist this organizational change which they consider as a top management initiative that goes against their own identity. They resist through deviant uses of the technology, by making voluntary mistakes in their report for example. Although they are in a subordinate position, they know they hold power (linked to operational information relative to the building site and to their position on the site foremen recruitment market). Such information constitutes resources which make them less disposed to obey. They therefore develop a strategic use of the technology at their disposal and finally manage to overturn managerial intents.

The third case, GammaCom, involves a large French telecommunications company (7400 employees), which decided to install a system to optimize interventions on its telecommunication network sites. Every day, different teams of technicians have to maintain or repair network sites. The new technological system is composed of a smartphone, connected to software which plans the interventions of technicians every day, depending on the needs of the network sites and the localization of the technicians. Every technician is equipped with a smartphone which dictates the list of jobs he has to do every morning and the approximate time he should spend on each site. The software is based on a geo-localization system, which enables the management to know exactly and in real time where the technicians are. Top management argues that this technological deployment should improve the technicians’ profitability and security. Nevertheless, the technicians immediately considered this system as a means of controlling their activities. It is true that some problems occurred in certain teams in the past, when technicians were accused of putting their own interests first when planning their interventions, instead of optimizing their round. To solve such problems, the new system makes the technicians follow a given round. This system thus leads to changes in the control systems, which rely on structuring behaviours. Mobile IT appear in this case as direct supervision tools. It’s all the more important to note that the management stresses the discipline potential, and uses the notion of surveillance as an argument to structure technicians’ behaviours. The upshot of this discourse around potential surveillance is that most of the technicians expressed strong reticence linked to concerns about loss of freedom. They therefore try to take advantage of the technology by adopting different attitudes. Some of them, for example, circumvent the system by switching off the geo-localization device in their car when they want to cut themselves off (at lunch or in the evening for example). Moreover, most of them reinvent the sense of technology and use it as an instrument of proof for their hierarchy, arguing that the technology is a means for them to demonstrate their involvement to their managers, for example. They thus develop a defensive use of mobile devices.
Finally, we should note that such actions lead the technicians to legitimize the technology and underlying control system.

The fourth case, Eurobank Consulting, involves a consulting firm composed of 12 consultants. These professionals work in a context characterized by fierce competition from leading consultancy agencies. A general ambient discourse conveys the idea of urgency and hyper-reactivity in this specific sector, where time appears to be a key resource. In order to increase efficiency and optimize time, the company has moved towards an agile and flexible organizational form. Consultants are increasingly allowed to work at distance, from home, for example, and to limit their movements to visits to customers. Surprisingly, no mobile technologies have been deployed in this company, except in the case of associates who are equipped with modern and sophisticated devices (in a statutory logic). On the contrary, the consultants are generally left to find their own equipment for professional purposes. They are fully aware of the need to be available and reactive. According to the managers, no particular control system exists within Eurobank Consulting. They claim that the company is characterized by an extremely flat management system, absence of control, and a relationship of trust with the consultants. They apply a management by objectives system, so that control is focused on “deliverables” (services provided to customers). In fact, the consultants create their own rules and use a form of self-control. The consultants consider that working outside traditional hours is a moral obligation and they remain available through mobile IT use. Following a long socialization process, they associate specific values with their job of consultant (involvement, responsiveness, discipline linked to time pressures). They finish by accepting implicit organizational rules as their own rules, and believe that they respect values they have developed on their own. In the end, their flexibility and freedom appear to be a disguised form of “coercive autonomy.” This case indicates the emergence of a subtle, invisible form of control via the mobile technologies used and the emergence of a permanent availability. As they consider involvement to be their duty, the consultants construct their own control system via mobile technology which is reinforced by temporal pressure.

6 DISCUSSION

6.1 Unanimous discourse regarding efficiency to recognition of diverse concepts of control

A summary of the cases studied shows the existence of unanimous discourses concerning efficiency and reactivity, which legitimize and justify the introduction of mobile IS in organizations. This discourse gives a dominant place to the notions of liberty, emancipation, autonomy and empowerment of mobile populations, together with a democratization of access to information (especially in the cases of Technoplus, ABConstruction and Eurobank Consulting). There is a counterpart to this culture of transparency, however, as it applies to the individual’s behaviour with efficiency consequently becoming a duty. A Foucauldian perspective provides a counterview to the idea of individual empowerment by showing that the delegation of responsibility generated by these technologies gives rise to new obligations which are just as constricting. Nonetheless, these constraints and, to a larger extent, the issue of control, are often neglected in the discourses. To be more precise, they are exploited in different ways by the senior management teams in the companies we studied. Our cases show the existence of discourse manipulation tactics, illustrating different visions of control and trust, depending on the populations in question. Discourse is thus manipulated to orient individual behaviours (for example, threat of surveillance is upheld to structure the technicians’ behaviours in the case of Gammacom while, on the contrary, the panoptic evolution of control systems at Technoplus is completely hidden behind the official discourses, so as to promote acceptance by the sales representatives, who are used to having a certain amount of power in their commercial area).

6.2 Changes to control systems

One of the main changes to control systems in the cases studied is the emergence of a time-related discipline system, enabling people who are by definition mobile in space and visually ungovernable to be controlled. Through the use of mobile IS, time effectively becomes an instrument for locating people, an organizational norm, and a governance technique. The organization thus uses time to
control individuals who cannot be compartmentalized in space. This shift may be likened to the move towards “control societies,” characterized by continuous control and instant communication (Deleuze 1990). The structuring of mobile populations in space-time leads to different and flexible forms of control, as they adapt to the populations in question through surveillance and bureaucratic control, target-based management and control via the shared values (concerted control). Let’s highlight that the different forms of control enabled by mobile IT, observed in our case studies (bureaucratic control, target-based management, and control via the shared values), correspond to the various disciplinary technologies identified by Foucault himself (hierarchical supervision, normalization, and dressage). The case studies also show a shift in the location of exercise of authority, as individuals take on an increasingly active role in the control process. (Some actors are led to legitimize the technology and underlying control system, while others directly participate to the construction of their own control). Lastly, this analysis enables us to highlight the subjective dimension of control, which depends on the vision that individuals have of their autonomy and control, as well as the use that is made of the technology by the manager. Control depends on the information needs of the activity, and the perceptions developed with respect to the mobile populations in question. These findings explain the emergence of different kinds of control. At the end of the day, technology does not necessarily imply ‘de-bureaucratised’ forms of control, as more subtle forms of control are involved via technology.

6.3 Individual reactions or the importance of power-knowledge relations

The cases studied highlight different reactions and forms of appropriation, identified by applying a Foucauldian perspective. In some cases, individuals develop highly proactive and positive attitudes with respect to technology, whether they are aware of their control potential or not. Self-discipline, the search for a certain degree of satisfaction and equilibrium in the professional and private sphere, and the emergence of personal reflexivity, are all reflections of the trend towards the adoption of technology. Other cases provide us with insights into the resistance of some individuals, with some forms of resistance motivated by the notion of identity and the vision of job function and autonomy. A Foucauldian approach also provides insights into the interactions between the individual, technology and the organization in the framework of a power-knowledge relationship that is highly evolutive, and that should be repositioned in a wider institutional and societal context.

7 CONCLUSION

As mobility becomes a central feature of society, exploring the evolution of control systems in relation to mobile IS is a key issue in the field of management and IS research. It is also a key issue for practitioners in the field of human resource and IS management since it has a direct bearing on change management and the adoption of technologies. Our research shows the relevance of Michel Foucault’s conceptual framework (linking discourses, discipline and ethics) to explore the interactions between IS, the organization and individual actors, particularly from the viewpoint of hierarchical relations and control systems. we need to take a critical look at our research approach given its anchoring in the Foucauldian perspective: this research itself contributes to a regime of truth. We therefore need to emphasize the very relative aspect of the validity of knowledge produced in this study, as well as the highly structuring aspect of its framework discipline. We nevertheless encourage other researchers to broaden the frame of our model and to develop a Foucauldian model applied to IS, which provides a political analysis of technological and organizational change.

References


Managing uncertainty in service production with mobile systems - case waste management company

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0464.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Services operation and management, Mobile systems, Resource Based View, Resource Dependence Theory</td>
</tr>
</tbody>
</table>
MANAGING UNCERTAINTY IN SERVICE PRODUCTION WITH MOBILE SYSTEMS – CASE WASTE MANAGEMENT COMPANY

Juntumaa, Miira, Helsinki School of Economics, POBOX 1210, FI-00101 Helsinki, Finland, Miira.Juntumaa@hse.fi

Lauraeus-Niinivaara, Theresa, Helsinki School of Economics, POBOX 1210, FI-00101 Helsinki, Finland, Theresa.Lauraeus@hse.fi

Tuunainen, Virpi Kristiina, Helsinki School of Economics, POBOX 1210, FI-00101 Helsinki, Finland, Virpi.Tuunainen@hse.fi

Öörni, Anssi, Helsinki School of Economics, POBOX 1210, FI-00101 Helsinki, Finland, Anssi.Oorni@hse.fi

Abstract

Today’s growth of the service sector as a whole has created demand for more efficient service production. Many services require interaction between customers and service personnel, whereas some can be automated into self-services. In this study, we focus on services, that are neither purely human facilitated, nor purely automated, and contain uncertainty in the production process. Based on resource centric theories of strategy and research on uncertainties in service production, we introduce a research framework to evaluate efficient solutions for service production. Our research framework looks at environmental and informational uncertainties, and how an organization can adapt to these by utilizing technology or skilled labour. Illustrated with a case company, we show how mobile information systems can be used to manage service production related uncertainties, which are also typically barriers to standardization. The case study demonstrates how informational uncertainty could be more easily controlled using the new system. The job satisfaction of the workers was increased and their turnover and training time was decreased. Additionally, customer complaints were reduced and invoicing became more efficient. These enabled the company to enhance the efficiency of the service production processes further, moving closer to standardizing and automating the service production process within an uncertain environment.

Keywords: Service production, Mobile systems, Resource centric theories, Uncertainties
INTRODUCTION

Service has been characterized as something that is perishable, immediate, subjective, intangible, and inherently variable (Chen & Chien, 2008); services are consumed but not possessed (Berry, 1980). Today, the service sector comprises roughly 75 percent of the gross domestic product of developed nations and employs more people worldwide than either agriculture or manufacturing (Maglio et al, 2008). The stream of activity in western economies flows from manufacturing of goods towards service production (see e.g. Penttinen 2007). Yet, there is also an undercurrent of conceptual change in the opposite direction: service production is increasingly turning into “service manufacturing”, or standardized service production, and information and telecommunication technologies (ICT) have a salient role in this development.

Not all services will be affected alike, though. Standardized service production requires automation and current technologies apply differently to different categories of services. Many services close to the conventional services marketing paradigm, such as hairdressing, require interaction between the customer and the service personnel, and hence are ill suited for automation at the current level of technological development. At the opposite end of the continuum are those services, such as personal banking, that can be – and in many cases have been - automated into self-services. In between these two opposites lies a host of services, for example environmental management or health care services, production of which requires combining the best parts of both technology and service personnel: efficiency and predictability of technology with human skills and flexibility. The focus of our study is on these services, that are neither purely human facilitated (like hairdressing), nor purely automated and digitized (like e-banking).

Furthermore, our specific interest lies in the possibilities of mobile technologies in supporting efficient delivery of these services. These mobile business-to-employee solutions, such as wireless field force automation are used by mobile employees, remote from their base of operations, utilizing wireless technologies to perform their specific business tasks in, for example, services such as facilities maintenance and insurance claims appraisal (Barnes et al, 2006). Mobile technologies are expected to enhance business efficiency by distributing information to the workforce, as well as by offering new communication channels with customers (Leung & Antypas, 2001). For example, in an action research study on mobile technology in field customer service, Rossi et al. (2007) claimed that with a mobile information system, the case company was able to realize efficiency gains through simplified processes and at the same time the company could serve the customers better through more transparent stock and sales information.

Managing services is different from managing goods, because of uncertainties related to the characteristics of services. Intangibility implies that service cannot be inventoried or easily displayed; heterogeneity arises because service often depends on labour, which is inherently more unreliable than machines; simultaneity of production and consumption (inseparability) means that the customer participates in the transaction, and therefore service is not easily centralized, and; perishability means that for many services, once the time of potential service passes, the opportunity to sell that service perishes (Roland & Chung, 2006).

In this study, we look at how mobile information systems can be used to manage uncertainties in service production. Based on resource centric theories of strategy and research on uncertainties in service production, we introduce a research framework to evaluate efficient solutions to service production, in terms of environmental and informational uncertainties. Furthermore, we are interested at how the uncertainties affect the levels of standardization or flexibility in the process of producing services. To illustrate the usefulness of the framework, we present a case study in a major Finnish waste management company. Even if the industry under study may be lacking in glamour, it is one of the future growth industries in the service sector and serves as a good example of an industry where companies operate in distributed
production environments. It is also a good example for describing the opportunities mobile technology offers for services standardization.

2 THEORETICAL BACKGROUND

Our study builds on resource centric theories of strategy, namely the resource based view (RBV) (Wernerfelt 1984; Barney 1986), the resource dependency theory (RDT) (Pfeffer 1981), and transaction cost economics (TCE) (Williamson 1975, 1992, 1996), focusing particularly on the treatment of uncertainty. Our perspective follows the tenets of service-dominant logic: we conceptualize service as a process, rather than a unit of output; we focus on dynamic resources, such as knowledge and skills, rather than static resources, such as natural resources; and we view value of service as a collaborative process between providers and customers, rather than what producers create and subsequently deliver to customers (Lusch et al, 2008).

Resource based view, resource dependency theory, and transaction cost economics as streams of strategic thinking, differ in their motivation: transaction cost theory has been conceived to explain the existence of different organizational structures (i.e, markets vs. hierarchies) in exchange systems. The resource dependency theory builds on research on the bases of power within organizations (Weber 1947) and seeks to explain organizational survival. The resource based view is motivated by the heterogeneity of organizations even when discounting the differences across industries.

What is common to these streams of strategic thinking is their focus on resources and, emphasis on uncertainty as the most important target of management function. Uncertainty, defined as “the degree to which future states of the world cannot be anticipated and accurately predicted “ (Pfeffer & Salancik, 1978, p. 67), is a central concept in management literature for it is the ultimate source of both opportunities and risks.

RBV (Wernerfelt 1984) emphasizes uncertainty related to intra firm factors of production. Barney (1986) suggests that internal resources and the related uncertainty are more firm specific than environmental uncertainty and hold higher promise for unique, sustainable advantage. TCE is motivated by environmental uncertainty which, combined with bounded rationality (Simon 1991) of decision makers, subject the firm to opportunistic behaviour of its transaction partners. The position TCE holds on uncertainty is close to that of the RDT, which posits that organization’s success is tied to managing environmental uncertainty related to supply of strategic factors of production (Pfeffer 1981). Thus, while RBV focuses on resources and related uncertainties internal to an organization, both TCE and RDT focus on external dependencies on strategically important factors of production.

The RBV suggests that firms should examine internal resources related uncertainty in an attempt to learn how to extract unique, sustainable competitive advantage. The other two schools of strategic thinking concentrate on inter-organizational relationships in resource acquisition and, as a result, emphasize environmental uncertainties. TCE suggests that organizations are born out of desire to optimize over production and transaction costs (Williamson 1975; 1992; 1996). RDT posits that the central management function is minimization of organization’s dependence on external resources and maximization of others’ dependence on resources of its own (Pfeffer 1981).

3 MANAGING SERVICE PRODUCTION

The Unified Service Theory defines a service production process as one that relies on customer inputs; customers act as suppliers for all service processes. Non-services (such as make-to-stock manufacturing) rely on customer selection of outputs, payment for outputs, and occasional feedback, but production is not dependent upon inputs from individual customers. (Sampson & Froehle, 2006) The increased service production requires more efficient ways to cope with the
variance that customers bring into service production process. From the service provider’s perspective, a major element of service production is human capital. When customers introduce a high degree of variability (in other words, uncertainty) into the service production process, service organizations may be able to address this variability and successfully satisfy customer needs, when their employees are proficient at diagnosing problems, thinking creatively, developing novel solutions, and so on; that is, when they possess high levels of skill, knowledge, and expertise (i.e., human capital) (Skaggs & Youn, 2004). Although, employees’ personal skills as well as the service production environment vary every time the service is produced. Standardizing the service production process or making it more flexible can be solutions to improve the service. Nowadays technology plays an increased role in the process. Next, we will define environmental and informational uncertainties, discuss the standardization and flexibility in service production, and introduce the role of mobile technologies in this context.

3.1 Uncertainties in service production

Environmental and informational uncertainties have an effect on the possibilities to standardize a given service or a service process. Environmental uncertainties are related to complexity and variability of the environment where the service is produced. Environmental situations and conditions cannot be fully controlled by the organization, and thus increase the uncertainty that the service provider faces. Environmental uncertainties are considerably lower for services produced and delivered at the service organization’s premises than for those produced and delivered outside. Furthermore, the environmental uncertainties are lower for services produced and delivered in a single or permanent location, than services produced in multiple or varying locations.

Informational uncertainty is an internal uncertainty within the organization. In producing non-standardized services, the experience and information held by personnel plays an important role. When a customer demands customized service, the employee needs to hold enough information to be able to produce the service in the best possible way. This involves informational uncertainty for the whole organization as the information might not be distributed within the organization, while it may also flow outside of it.

3.2 Technology-enabled standardization or labour-enabled flexibility in service production

To adapt to the environmental and informational uncertainties presented above, the organization has two options: either to standardize the service production process or to enable more flexible process. Levitt (1972) suggests that the standardization of a service can take three forms: (1) substitution of technology for personal contact and human effort; (2) improvement in work methods; and (3) combinations of these two methods. Standardized services do not, however, necessarily mean mechanical services, although many service tasks are routine, allowing specific rules and standards to be easily established (Järvinen et al., 2003). If the environmental and informational uncertainties are high, standardized service production process might be unrealistic, as it would require the organization to be able to control the uncertainties. Clearly an organization cannot fully control its environment. Standardization works best on routine type of services that are produced in a fixed location.

Increased flexibility in the service production process, regardless of high uncertainties, requires highly skilled labour, who can creatively adapt to changing needs. This is a solution with high environmental uncertainties. Still, the knowledge is rather personal and leaves the organization with a risk of losing some of the information along its labour.
3.3 Role of mobile technologies in service production

Mobile technology offers new possibilities to overcome service production related uncertainties. The crux of employing mobile technology is that it enables companies to collect and maintain real time information about the variable production environment and to share this information on the on-demand basis to the employees going about in the distributed production environment. The environmental uncertainties can be brought under control through more complete real time information. Changes in the environment can be instantly, and often automatically, fed into the system, and they are immediately available to others in need of this information. From the resource planning point of view accumulation of production related information in the company databases diminishes the informational uncertainties caused by the stickiness or leakiness of information. Information necessary to efficient resource planning will not remain the property of individual employees, rather, it is instantly available to all employees and the mobile systems can be used to automatically feed this information to those needing it to diminish the informational delays. This information will also remain in the company even after the employee who discovered it leaves, decreasing the dependency on particular individuals’ attributes.

The more dynamic the environment, the more important is timely flow of information. The more complex and idiosyncratic the service production process is, the more information is needed to properly automate or standardize the process. Especially those services produced and delivered outside of the company premises are likely to benefit from the opportunities provided by mobile technology to standardize the work flow and the service production process.

4 RESEARCH FRAMEWORK

Our research framework builds on the environmental and informational uncertainties, and two extremes for an organization to adapt to the situation: standardization with technology and flexibility with highly-skilled labour (see Figure 1). As explained, typically low uncertainties in service production enable standardization as technology can easily be utilized but increased uncertainties require flexibility in the process and highly-skilled labour that can cope with varying requirements that the uncertainty creates.

The upper left corner in our framework demonstrates a situation with both low environmental uncertainty and low informational uncertainty (L,L). Both the service environment and the information related to service production are stable, making it relatively easy to standardize the service production using technology instead of labour. In fact, not standardizing it would increase service production costs since, skillful workforce would be used for tasks that do not require workers to assess each situation separately and make situational decisions. An example of this type of service could be a retail store cashier. Standardizing service production in low uncertainty situations decreases the significance of particular workers’ personal learning abilities and importance of having long term loyal workers that have learned from practice and stay within the same organization to use and share the knowledge. If and when these workers leave the organization their knowledge leaves with them. However, as both uncertainties are low the necessary knowledge can be stored within the organization and share it as it is needed. This is particularly easy in routine services but also possible in some more complicated type of service, if the knowledge is such that it can be saved in a database. Then other workers can be provided with access, so they will be able to produce the service according to general guidelines without needing to assess each situation separately. This leads to increased homogenization of organizational roles at the same organizational level and tasks, but heterogeneity of the roles at the different organizational levels.
Figure 1. Service production and related uncertainties

The lower right corner shows the combination of high environmental uncertainty and high informational uncertainty (H,H). High environmental uncertainty demands high flexibility in the service production. As the service production environment and the required information varies, workers need to be flexible and skilled to assess the service production situations separately using their knowledge and decision making rights. An example of a service producer in an uncertain environment could be a travel guide. Greater flexibility that is achieved through highly trained service personnel would mean encouraging learning. Simon (1991) offers that organizational learning is profoundly influenced by the organizational roles. Distribution of decision making rights and the related expectations decree the organizational role structure. The mandate of the role is largely dependent on the variety of situations facing the employee. Clearly, if the service production process is characterized by high environmental uncertainty, the service varies highly depending on its environment, the customer, and the service producers. Thus, the workers need to be able to make quick decisions in each case. The role within an organization requires ability to learn, assess the particular situation and make a decision. As people vary on their ability to learn, wider mandates typically lead to increased heterogeneity in de facto organizational roles and produced services.

The lower left corner, on the other hand, suggests low environmental uncertainties but high informational uncertainties (L,H). In this, having highly-skilled and flexible labour handling tasks that have low environmental uncertainties basically means high costs and thus inefficient service production process. Based on the stable service production environment, the service could be automated or standardized utilizing technology, but now skilled-labour is occupied in producing this service. If the information held by the labour would be stored within an organization and the service would be automated, the production process would shift to the upper left corner making it much more efficient. The upper right corner demonstrates high environmental uncertainty but low informational uncertainty (H,L). This is not an efficient solution for service production, as high environmental uncertainty requires skillful and flexible employees. If the internal informational uncertainty would be low, it suggests that employees do not possess critical knowledge and are therefore not able to respond the need set by the uncertain environment. Thus this means that the company has high opportunity costs as it cannot produce the service in a way that would meet the demands set by highly uncertain service production environment. Having high environmental uncertainty and need to operate efficiently, an organization must move to lower right corner and invest in skilled labour.
Utilizing technology to automate or standardize service production process in a situation of both low environmental uncertainty and low informational uncertainty or investing in highly-skilled and flexible labour in high environmental uncertainty and high informational uncertainty are thus both applicable and sustainable solutions for an organization. This area is demonstrated with white colour in our framework. However, having a situation where environmental uncertainty is low but informational uncertainty is high, as in the lower left corner of the framework, means high service production costs: if the service production environment is stable, investments made on flexibility and ability to adapt to environmental uncertainty are needles. On the other hand, if environmental uncertainty is high but informational uncertainty is low, as in the upper right corner of the framework, the changing environment requires flexibility and adaptation to new situations but the organization does not meet these requirements lacking the required information and flexibility.

Thus, when facing a situation of high environmental and informational uncertainty, the service production requires flexibility created and information produced by skilled labour. To adapt to this situation an organization might gain greater control over the informational uncertainty if technology can be utilized to maintain more up-to-date information about the service production environment and other service production related issues, and distribute it on a need-to-know basis. As mentioned having highly trained service production personnel increases organizations dependency on them which increases informational uncertainty. The solution is to have more up-to-date information stored within the organization databases that is shared in service production. The stored information makes it possible to decrease the impact of environmental and informational uncertainties that the service producer is facing and move the production process closer to automation also within an uncertain environment. In terms of our framework, this would mean that moving from lower right corner closer to upper right corner is still efficient, as the required information or environment has not objectively changed. This requires an extensive use of mobile technologies to be able to share the information to service producer living within an uncertain service production environment.

5 EMPIRICAL STUDY

In the empirical part of our study, we look at the presented ways to manage service production related uncertainties in a company specializing in environmental management. In particular, we focus on a new mobile system introduced to facilitate the field work of the service personnel in waste management, and the ways the system has helped to adapt to both environmental and informational uncertainties. Furthermore, in order to illustrate the usefulness of our framework, we discuss the case company solution in terms of informational and environmental uncertainties.

5.1 Research methodology

We chose to look into one case to be able to acquire more in-depth understanding of the phenomenon. The empirical material was collected through in-depth, semi-structured interviews with four employees of the case company. In order to get as multifaceted view of the phenomenon as possible, the interviewees were at different levels and different positions in the company: the development director, the Auvo system developer, the project manager, and finally a field worker. All of them have been closely involved in the development and introduction of the mobile logistics system. The interviewed field worker uses the system on a daily basis, and he has been with the company long enough to know both the earlier system as well as the new mobile system.

Each semi-structured interview, lasting from half an hour to one hour, was conducted by two researchers. All the discussions were recorded and transcribed for analyzes. The topics covered the old service process, the new system, and its implications. Of particular interest was how the new system has changed the work process, what kinds of benefits had been created by it, and
the impacts the new system has had on the processes and the company. In addition to the interviews, publicly available information related to the organization and its services was collected and analyzed.

5.2 The Case Company: Lassila & Tikanoja

The case company, Lassila & Tikanoja (L&T) specializes in environmental management and support services for properties and plants. L&T has business operations in Finland, Sweden, Latvia, Russia and Norway. L&T’s net sales in 2006 amounted to EUR 436.0 million and it employed 8,328 people at the end of the year. The company’s shares are quoted on the Helsinki Stock Exchange. L&T’s services are split into three divisions: Environmental Services, Property and Office Support Services, and Industrial Services.

L&T employs a vast number of maintenance personnel and similar, whose work consist of a collection of maintenance and support tasks in customers’ premises. Previously each worker received a list of work tasks and their planned schedule as they arrived to the office. They got all the needed information, tasks, schedules, and maps on paper format, with which they started their work day. After accomplishing the tasks, they returned to the office to report the finished work tasks and to receive a new set of instructions. Changes to the schedule were possible only as the workers physically visited the office. This process involved a lot of driving, causing high fuel as well as car maintenance costs. Salaries and fuel and car maintenance costs account for the top three costs for the company. Furthermore, in the old paper based system, the field workers had to learn the routes and tasks, which took a considerable amount of time for new workers.

5.3 The Mobile System: Auvo

L&T launched the mobile logistics control system, named Auvo, to improve its efficiency. Auvo collects information related to service production and distributes it on demand see Figure 2.). For waste management field workers, whose job includes a lot of driving, the information related to the locations and the routes is crucial.

![Diagram](image)

**Figure 2. The new mobile system: Auvo**

Each maintenance worker has a mobile device with him, providing the work tasks, schedules, as well as maps that were all previously delivered on paper format. Workers do not need to return to the office to sign for new tasks, but the updates and changes are delivered automatically on their mobile devices. As the maps and routes are provided directly to the devices, the routes do not need to be learned. As a task is completed, the worker signs it by pressing Enter on the device, possibly adding some other information, if needed. The process moves from the
customer order to the service production and finally to billing without manual paper handling. The field workers have their mobile devices with which they operate, and office workers and customers can access the systems using a web browser.

5.4 Benefits of the new system

According to the development director and the developer of Auvo, the new mobile system has had a significant effect for the company in form of reduction of production costs, consisting of salaries, fuel costs, and the maintenance of special vehicles. Fuel consumption has decreased, since there is less unnecessary driving back and forth to the office, and the field workers know exactly where the target customer sites are. Maintenance of the vehicles is now cheaper because the control system lowers driving speed, causing less damage to the vehicles. Also the shortening of the entire production process from order to invoice from earlier 30 days to only 3 days with Auvo is seen as a significant advantage.

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Director</th>
<th>Developer</th>
<th>Project Manager</th>
<th>Field worker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large savings from less driving to and from the office.</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire production process from order to invoicing has become faster.</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decreased production costs, as the field workers can serve more customers in the same time.</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPS positioning in vehicles making it possible for the supervisor to know in real time where the vehicles are.</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New customer orders relayed directly to field</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waste management production life cycle shortened from 1.5 to 1 day.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Availability of real time information on production.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Automation of invoicing has decreased manual work from 100 to 4 invoices.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Fuel consumption has decreased.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Maintenance of vehicles is cheaper.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Invoicing increased 10 %, because there is less unforgotten invoices.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>The quality of customer service has risen, because of less mistakes by the field workers.</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The number of customers has increased.</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPS maps and work lists in cars, letting the field worker know exactly where the customer site is and how to drive next.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Field workers’ duties are easier than before.</td>
<td>+</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Less mistakes with customer orders by the filed workers.</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acknowledgement of customer order/target is easier.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>It is quick and easy to train a new field worker.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Field workers job satisfaction has increased, and worker turnover has decreased from 60% to 4.5%.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Table1. The benefits of AUVO system according to the interviewees (+ denotes a mention by an interviewee)

For the project manager, the biggest benefits of Auvo are created by GPS positioning, shorter production life cycle, and automation of invoicing and customer orders. Waste management production life cycle has shortened from 1.5 days to 1 day, because now the management can reschedule orders and change the order of customer sites to be visited, if needed. The field workers can serve more customers in the same time, and there is no more waiting for the next assignment at the office. New customer orders are relayed directly to the field workers’ task lists and to the electronic maps, reducing the number of mistakes made. GPS positioning in vehicles makes it possible for the supervisors to know in real time where each vehicle is driving. Automation of invoicing has decreased manual work from 100 to 4 invoices, as there are no
more paper orders to handle. Furthermore, invoicing has increased for 10 %, because there are fewer orders and invoices forgotten. The decreased number of mistakes has also increased the level of customer satisfaction. Also the number of customers has risen lately.

In the point of view of the field workers, the biggest advantage is that the work is, in general, easier than before, and there are less risks of making mistakes. GPS maps and work lists in cars allow the field workers to know exactly where the next customer site is and how to drive there. The field worker we interviewed also valued the fact that acknowledgement of orders and customer sites target is easy and swift. The job satisfaction of field workers has increased, and worker turnover has decreased dramatically from 60% to 4,5 %. A significant advantage is also that it is quick and easy to train new field workers. Our interviewee has already trained 100 workers to use the new Auvo system.

Summary of all the benefits mentioned by the interviewees is presented in Table 1.

5.5 Discussion

Service production that requires high flexibility has conventionally been labour-intensive and more costly than standardized service production. Being able to standardize the production process enables the company to decrease production costs, in our case through simultaneous management of environmental and informational uncertainties.

The service production environment involves uncertainties as the service is produced in customer premises. The routes can change and it is crucial for new employees to learn the environment related information. Conventionally, the employees have been the data depositories possessing the idiosyncratic information needed to efficiently navigate the distributed production environment. This makes the informational uncertainty also relevant for the service production. Ownership of production related information commands increased influence. The knowledgeable employees have been a valuable asset to the company and, hence, enjoyed considerable negotiating power over the terms of employment. Information uncertainty is most clearly reflected in the inability of the employer to prevent information from leaking outside the company as employees change their employment. The mobile logistics control system enabled the collection of service production related information and its distribution on demand decreasing informational uncertainties that the company is facing in its service production process.

Figure 3. Service production and related uncertainties in L&T
With the old paper-based system, the company was situated in the lower right hand corner of our research framework, having high environmental and informational uncertainties. Before the emergence of mobile location-aware information systems, environmental uncertainty could only be combated with specialized labor, as technology fell short of meeting the requirements of the work and its employment would have led to high opportunity costs in the guise of low service quality. The mobile production system allowed simultaneous combatting of both environmental and informational uncertainty. Even if the environment cannot be changed, its idiosyncrasies can be tracked in real-time. This information ends up being stored in the company instead of being in the possession of employees only, which decreases informational uncertainty. This decrease in informational uncertainty is depicted in our framework (see Figure 3) as a move closer to the upper right corner. The mobile system thus has changed the preferred organization for efficient service production: the upper right corner had previously been an inefficient solution but utilization of mobile technology has pushed back the frontier of efficient service production, increasing the ability to automate service production within an uncertain environment.

The mobile logistics control system also adds value for the workers. All the necessary information is stored in a database and accessible on demand. Especially the electronic maps and routing makes the work easier and less stressful. Previously having all the information on paper was risky in the sense that losing some of the papers, or ordering the papers enabled more mistakes from the workers. Now, it is easier and quicker to train new workers, and the work in general is easier. Initially, when the system was introduced, change resistance occurred, as the workers could not see the added value. However, after its launch, the system has been constantly updated and improved, the workers have been encouraged to put forward any wishes concerning its improvement. After the adoption, job satisfaction has clearly increased and consequently, the turnover of workers has fallen considerably.

The impact of the system in an organizational level is easy to recognize. Customer service quality has been improved as the number of mistakes has fallen. This can be seen from the radically decreased amount of customer complaints. The invoicing is more accurate as the system does not loose orders or forget to send invoices. Previously the financial department used to go through approximately 100 orders, in comparison to 4 nowadays, as 96 are invoiced automatically. This improvement is dramatic. Also, unnecessary driving back and forth is minimized which decreases fuel consumption. Overall, the lifecycle of the production process has fallen from 1.5 days to one day.

6 SUMMARY AND CONCLUSIONS

In this study, we looked at how a mobile information system can be used to manage uncertainties in service production process. We introduced a research framework based on resource-centric theories of strategy and literature on uncertainties in service production. The research framework looks at environmental and informational uncertainties, and the efficient ways of service production within the framework. The framework was illustrated with a case company in the field of waste management. The case company had previously used a paper-based system in distributing work tasks, which has now been replaced with a mobile information system to improve and rationalize the information flow within the service production process.

The efficiencies gained with the new mobile system are clear, as the length of the service production life cycle as well as the invoicing life cycle have fallen together with the number of customer complaints and lost orders. The improved information handling and its on-demand delivery were not possible before the launch of the new mobile system. This efficient information processing and sharing are required to enable production of standardized services. Thus, with soundly implemented mobile system service production can be standardized to an activity reminiscent of goods manufacturing. Standardization of service production process
improves cost efficiency, as the company does not have to train and support highly diverse and skilled workforce, and some parts of the process can be automated.

This case study also demonstrates how ICT changes the composition of the efficient solutions. Utilizing ICT applications companies can change their service production processes from highly flexible to standardization, which would not be possible with the traditional paper based systems. The next step in our research is to further evaluate the framework with other companies in different industries and environments, with different informational and environmental uncertainties.

7 REFERENCES


A preliminary laddering analysis on mobile services usage

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0727.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Adoption, Means-end analysis, Mobile communications, Mobility</td>
</tr>
</tbody>
</table>
Abstract

This paper proposes the use of laddering technique to determine the decision making process to adoption. Means end theory reveals the human values that are fulfilled for the individual by using various m-technology and services. Preliminary findings of a bigger study are presented. IT shows that mobile technology and services often fulfil such basic needs as self-esteem, achievement, individuality, belonging and well-being. Exploring the realization of values as a theoretical framework offers researchers a way forward in environments characterised by individual technology decisions.

Keywords: Technology Adoption, Means-End Analysis, Mobile Communication, Mobility
1. Introduction

Mobile phones are one of the fastest adopted technologies in history. In 2002 the number of mobile phones worldwide surpassed the number of landlines, with 4 billion mobile phones forecasted for 2010 (Nystedt, 2006; Srivastava, 2005).

The mobile phenomenon goes beyond its unprecedented speed of adoption. Many factors that have contributed to this adoption include the technology’s ability to be used at any time in any place, its simple operation and compatibility and the fact that it connects to any phone system (Rogers, 2003). Also, mobile technology is continuously being upgraded and reinvented (Anckar, 2002; Rogers, 2003). Mobile phones have evolved from being a tool for business people always on the move to becoming an integral part of people’s personal communication. Mobile devices and services keep continuously presenting new features and application that are leaving its core function behind. Mobile phones importance is such that there are. Sociological studies turn to mobile phone as an integral tool to understand social behaviour in the 21st century (Srivastava, 2005).

The aim of this paper is to present the preliminary findings of a study of m-technology and services adoption and usage. Using means end chain and laddering framework and methodology. It will begin by explaining the growth and significance of mobile-communication. It will then assess the applicability of means ends chain for researching m-technology and services adoption and use. The paper explains the benefits of using laddering to identify the reasons underpinning different consumer value choice perceptions. Finally, the significance of values to explain the decision making process with be discussed.

2. Theoretical Framework for examining Mobile adoption and Use

Many scholars and industry professionals are attempting to understand what influence individuals’ to adopt and use mobile services. A number of theories have been used to study consumers’ decision making process and to determine the “how” and “why” of adoption of technologies. People adoption of new technologies depends on several of factors, for example, how useful they perceive the technology to be, the product availability and price, how user friendly it is, social-economical motivation (benefits), culture, demographics and psychographics, and past experience.

A number of theories relating to the adoption of new technologies by consumers exist in the literature:

Some authors such as Barnes and Huff (2003) have used Roger’s (1962, 2003) innovation and diffusion theory to examine the diffusion of Internet access via mobile phones (iMode). Roger’s developed a number of characteristics that explain innovation diffusion: 1) Relative advantage: The degree to which the technology provides an advantage over other methods, 2) Compatibility: The degree to which the technology is compatible with how people work or behave, 3) Complexity: Whether people perceive the technology as easy to understand and use, 4) Trialability: The degree to which a technology can be trialled before being adopted, 5) Observability: The level of visibility of the product to the other members of the adopter’s social group.

Studies related to advertising and marketing are closely associated with consumer attitudes and built around the Theory of Reasoned Action (TRA) (Ajzen & Fishbein, 1980; Fishbein & Ajzen, 1975)and its applications to IT settings. The theory provides a framework to understand why people behave as they do when they are making decisions. TRA proposes that the use of technology can be predicted by a person’s behavioural intention and that this is determined by a person’s attitude towards using the technology. A person’s attitude is shaped by their positive or negative feelings towards performing a specific behaviour (or using a technology) and whether people who are important to that person, typically in the workplace, think that the person should or should not perform the behaviour.
Using a form of TRA (Tsang, Ho, & Liang, 2004) examine the link between attitude, intention and behaviour in relation to m-marketing. In their model entertainment, informativeness, irritation and credibility are seen to shape attitudes in addition to permission which has a major impact. The availability of incentives, such as free calls, impacts on intention to receive m-marketing for certain attitudes. Intention is directly related to behaviour in relation to m-marketing. Their study findings indicate that consumers have a negative attitude towards mobile advertising unless they have consented to it. All four attributes of mobile advertising impact significantly on attitude towards mobile advertising. Attitude was significantly correlated to intention with incentives also positively impacting on intention. There is a strong correlation between intention and behaviour. This study was conducted in Taiwan with a large percentage of respondents regularly using SMS, although it is unclear from the results presented the extent to which respondents had received mobile advertising.

The Technology Acceptance Model (TAM)(Davis, 1989) is tailored to information systems contexts and is designed to predict IT acceptance and usage in the workplace. It focuses on perceived usefulness of the technology and perceived ease of use. The consumer behaviour literature shows that utilitarian, in the sense of instrumental value, or hedonic benefits determine the intention to consume. In some m-commerce studies hedonistic factors including entertainment value have been considered as significant (Bauer, Barnes, Reichardt, & Newman, 2005; Bouwman, Carlsson, Pirkko, & Francisco, 2008).

As there are a number of adoption models available to researchers, (Venkatesh, Morris, & Davis, 2003) synthesized the main models in order to provide a unified view of user acceptance. The unified model identifies determinants and moderators related to intention and suggest intention is a predictor of use behaviour. Four factors impact on intention and usage: performance expectancy, effort expectancy, social influence and facilitating conditions. The key moderators are gender, age, experience and voluntariness of use. Interestingly, attitude was considered to be in overlap with performance and effort expectancies. The non-significance of attitude in the presence of these two other constructs has been supported in a number of other studies (Standing, Benson, & Karjaluoto, 2005). The unified model was found in empirical studies to be a substantial improvement on any of the other earlier models. Work on technology acceptance is still evolving with for example studies that integrate user satisfaction constructs with technology acceptance constructs (Venkatesh & Brown, 2001).

Standing, Benson and Karjaluoto (2005) used a version of the Unified Theory to determine significant factors in the decision to participate in m-marketing schemes and found that granting permission, financial savings and highly relevant information were significant factors in the decision to participate but the time and effort involved in processing m-marketing messages were not considered important.

Consumer adoption related factors can be summarized as including the consumer’s general attitude toward the technology, level of involvement, innovativeness, response to stimuli, trust and perceptions of utility, choice, control and risk. Demographic factors (age, gender, income, education) have also been found to be important control variables to consider when looking at consumer acceptance of m-services (Barnes & Scornavacca Jr., 2003; Bauer et al., 2005; Davis, 1989; Gebauer, Tang, & Baimai, 2008; Jonathan, 2008; Se-Joon, James, Jae-Yun, & Kar-Yan, 2008).

Although it is widely recognized that younger consumers have embraced mobile technology it is being increasingly recognized that factors beyond age or gender may be important. It can be argued that segmenting people on the basis of their acceptance and use of technology as well as their lifestyle motivations is more representative of their actual behaviour (Sultan & Rohm, 2005). One study by Forrester Research (Forrester, 2001) looked at factors such as technology attitude, income, career, family and entertainment. One older group identified in the study characterized by high income, interests in
entertainment and a positive attitude toward technology may be high mobile device users. Cultural factors can also impact the type and nature of service used. For example, in high power distance cultures such as Korea, text messaging to supervisors would be seen as a serious offence (Sarker & Wells, 2003).

Consumers’ adoption of new technologies/services depends on a number of factors, for example, the type of service to be offered, how comfortable people feel using the technology, how user friendly the service interface is, socio-economic factors, motivations (benefits), culture, demographics and psychographics, time that the customer expects to use the service and past experience (Daghfous, Petrof, & Pons, 1999; Sultan & Henrichs, 2000). Sarker and Wells (2003) provide a framework for understanding the adoption and use of mobile devices that includes most of these factors. Their model considers not only the decision made in the initial adoption but also how users appropriate the technology and services through exploration and experimentation. They argue that users assess their experiences on at three dimensions: functional (e.g. time savings), psychosocial (e.g. safety, elevated self-worth, sense of freedom) and relational (building relationships). We argue that the benefits are consequences of using the technology but the real psychosocial implications are the drivers – in other words all three are not on the same lane but part of a hierarchy.

3. Means-End Chains

The means-end chain concept concentrates on the relationship between product/service attributes, consequences and values (Gutman, 1982; Judica & Perkins, 1992; Leao & Mello, 2001, 2002; Reynolds & Gutman, 1988; Woodruff & Gardial, 1996). Attributes relate to product characteristics. Consequences are defined as the physiological or psychological results acquired directly or indirectly by the consumer from his/her behaviour (product or service use). This model represents how the consumption of a product enables the individual’s realization of his/hers ends desires. The central aspect of this theory is that “...consumers choose actions that produce desired consequences and minimize undesirable consequences” (Reynolds & Gutman, 1988).

It is implicit to the means end chain model different levels of abstraction to demonstrate their depth, the levels of abstraction are a way to identify different subcategories of attributes, consequences and values (Gutman & Reynolds, 1979). Attributes are described to be the physical aspects of the product or its abstract properties. For example mobile phone size is a physical, tangible attribute. However more subjective properties like cute or flash are abstract attributes (Fouskas, Pateli, Spinellis, & Virolo, 2002; Reynolds & Gutman, 2001). For consequences the levels of abstraction are functional and psychological. Functional consequences relate to direct results of using the product. For example SMS saves money. Psychological consequences are those instrumental to the achievement of psychosocial results. (Reynolds & Gutman, 2001) For example make an impression, have fun, feel powerful etc. Finally Values, they represent end states of existence. The value construct in this model is drawn from the concept used in psychology and sociology and relates to Rokeach, (1973) classification of human values. (Rokeach, 1973) identified two types of values: instrumental and terminal. Instrumental values relate to those values that act like tools in achieving end-state behaviours (values like courage, honesty, ethics, etc.). Terminal values, also used by (Gutman, 1982) refer to “Preferred end-states of existence” (Gutman, 1982 p 61) for example: accomplishment, happiness and satisfaction. Gutman’s, (1982) model (1982) has two basic underlying assumptions: 1) Values are connected to consequences as long as the consequences have positive or negative connotations and 2) Consequences have a direct relationship with product attributes as long as, consumers obtain the product which may cause the desired benefits. It is not uncommon for researchers to have some level of difficulty distinguishing between psychological consequences and instrumental values.
4. Methodology

The ladder technique is the method used to reveal the means-end hierarchy (Leao & Mello, 2002; T. J.; Reynolds & J. Gutman, 1988; (Zanoli & Naspetti, 2002). The term ladder refers to the relationship between attributes, consequences and values. It is a representation of the connection between the actual product and the consumer’s cognitive process that leads to a direct and useful understanding of his/hers perceptual orientation in relation to the product or service. The ladder technique is an in-depth individual interview seeking the understanding of consumers’ decisions. It translates product attributes into associations relevant to consumer’s “self”. (T. Reynolds & J. Gutman, 1988)

The data analysis using ladder has four steps. The first one is to specify the elements of the means end chain (MEC) by separating into chunks of meaning. Then a content analysis of the interviews and coding is done in order to combine and generalise the meaning across subjects. The third step is development of the implication matrix by quantifying the relations between the content codes. Finally the there is the creation of hierarchical value map to illustrate the connections between the different levels of abstraction (Gengler & Reynolds, 1995)

For this research 60 people were interviewed but only 57 used. Due to the exploratory aspect of the research the only factor that was fundamental in terms of sampling was that the subject used a mobile phone. Gender distribution was random but the total sample has resulted in 23 males and 34 females. We used a convenience sample and ages varying from 16 to 59. During the analysis the respondents were divided for demographic purpose in gender and age group. The age groups identified were: 16-25; 26-40; 41-54 and over 55. Although due paper size constrains, in this paper we will only discuss the findings in the general hierarchical value map (HMV) and a gender segmented map. The results of the HMV segmented by age will be presented in further publications.

The study aimed at identifying the underlying reasons for mobile phone usage. The ladder interview began with a question to allow respondents to elicit the distinctions they perceive as different features/services offered on their mobile phones. One of the most used question to elicit distinction was “Why did you get a mobile phone”. This has elicited direct services associate with the mobile phones as well as extended services. The interviewees did not seem capable to differentiate between “core” and extended product. Neither the literature seems to be looking at this particularity. What is a core service and what is an embedded object. However this does not seem to impact and in the end the most relevant attributes were voice and text

For this research the software Mecanalyst, (specifically for Laddering) was used. The software was particularly useful in the generation of the implication matrixes and the HMVs. Macanalyst is commercial software. There is an open source ladder software called LadderMAP developed by Chuck Gengler. However this software is MS DOS based and there is very little or inexistent technical support. Mecanalyst was designed by two Italian professors and is commercialise by an Italian SME called Skymax-dg. This software is windows based and is quite useful in the way it generates the tables and the maps. However the first step is the content analysis is still a manual effort. The second step in the data input is the grouping of the many different terms used by respondents to refer to the different levels of abstraction, where Mecanalyst was quite useful. It also generates a final list of codes. This research had 45 terms (code and synonyms) distributed in the following way: 07 concrete attributes, 03 abstract attributes, 13 functional consequences, 15 psychological consequences, 03 instrumental values and 02 terminal values. The codes table used in this study are depicted on table 1. As explained before the table is the starting point of the ladder, it is the probing method to distinguish differences between products in this case, mobile phone characteristics.
5. Findings

The attributes identified here refer mostly to handset features instead of services offered by carriers for example news alerts. This indicates that those services are not yet a major determinant to the use of mobile services. Reinforcing that voice and text are what consumers are really interested in. Three abstract attributes were identified then asynchronous nature of text message, MMS and the alternative of consumer having only one device to take pictures.

<table>
<thead>
<tr>
<th>Code</th>
<th>Type</th>
<th>Freq F</th>
<th>Freq M</th>
<th>F %</th>
<th>M %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Camera</td>
<td>Ac</td>
<td>10</td>
<td>9</td>
<td>29%</td>
</tr>
<tr>
<td>2.</td>
<td>Design/Looks</td>
<td>Ac</td>
<td>6</td>
<td>8</td>
<td>18%</td>
</tr>
<tr>
<td>3.</td>
<td>Internet</td>
<td>Ac</td>
<td>3</td>
<td>5</td>
<td>9%</td>
</tr>
<tr>
<td>4.</td>
<td>PDA (organiser)</td>
<td>Ac</td>
<td>5</td>
<td>6</td>
<td>15%</td>
</tr>
<tr>
<td>5.</td>
<td>Ringtone MP3</td>
<td>Ac</td>
<td>3</td>
<td>3</td>
<td>9%</td>
</tr>
<tr>
<td>6.</td>
<td>SMS</td>
<td>Ac</td>
<td>28</td>
<td>16</td>
<td>82%</td>
</tr>
<tr>
<td>7.</td>
<td>Voice calls</td>
<td>Ac</td>
<td>25</td>
<td>20</td>
<td>74%</td>
</tr>
<tr>
<td>8.</td>
<td>Asynchronous communication</td>
<td>Aa</td>
<td>8</td>
<td>2</td>
<td>24%</td>
</tr>
<tr>
<td>9.</td>
<td>Get a camera</td>
<td>Aa</td>
<td>2</td>
<td>2</td>
<td>6%</td>
</tr>
<tr>
<td>10.</td>
<td>Picture Message</td>
<td>Aa</td>
<td>1</td>
<td>3</td>
<td>3%</td>
</tr>
<tr>
<td>11.</td>
<td>Capture memories</td>
<td>Cf</td>
<td>6</td>
<td>1</td>
<td>18%</td>
</tr>
<tr>
<td>12.</td>
<td>Convenient</td>
<td>Cf</td>
<td>28</td>
<td>19</td>
<td>82%</td>
</tr>
<tr>
<td>13.</td>
<td>Easy to carry around</td>
<td>Cf</td>
<td>3</td>
<td>3</td>
<td>9%</td>
</tr>
<tr>
<td>14.</td>
<td>Effective cost savings</td>
<td>Cf</td>
<td>11</td>
<td>6</td>
<td>32%</td>
</tr>
<tr>
<td>15.</td>
<td>Express myself</td>
<td>Cf</td>
<td>4</td>
<td>1</td>
<td>12%</td>
</tr>
<tr>
<td>16.</td>
<td>Feels Impersonal</td>
<td>Cf</td>
<td>6</td>
<td>2</td>
<td>18%</td>
</tr>
<tr>
<td>17.</td>
<td>Get Organise</td>
<td>Cf</td>
<td>6</td>
<td>13</td>
<td>18%</td>
</tr>
<tr>
<td>18.</td>
<td>Get things done quicker</td>
<td>Cf</td>
<td>18</td>
<td>2</td>
<td>53%</td>
</tr>
<tr>
<td>19.</td>
<td>Healthier</td>
<td>Cf</td>
<td>1</td>
<td>3</td>
<td>3%</td>
</tr>
<tr>
<td>20.</td>
<td>It is more personal</td>
<td>Cf</td>
<td>8</td>
<td>6</td>
<td>24%</td>
</tr>
<tr>
<td>21.</td>
<td>Keep in contact</td>
<td>Cf</td>
<td>10</td>
<td>8</td>
<td>29%</td>
</tr>
<tr>
<td>22.</td>
<td>Keep informed</td>
<td>Cf</td>
<td>3</td>
<td>7</td>
<td>9%</td>
</tr>
<tr>
<td>23.</td>
<td>Record / Proof</td>
<td>Cf</td>
<td>4</td>
<td>4</td>
<td>12%</td>
</tr>
<tr>
<td>24.</td>
<td>Unique</td>
<td>Cf</td>
<td>1</td>
<td>3</td>
<td>3%</td>
</tr>
<tr>
<td>25.</td>
<td>Acceptance</td>
<td>Cp</td>
<td>6</td>
<td>2</td>
<td>18%</td>
</tr>
<tr>
<td>26.</td>
<td>Always available</td>
<td>Cp</td>
<td>1</td>
<td>4</td>
<td>3%</td>
</tr>
<tr>
<td>27.</td>
<td>Bonding</td>
<td>Cp</td>
<td>3</td>
<td>2</td>
<td>9%</td>
</tr>
<tr>
<td>28.</td>
<td>Considerate of others</td>
<td>Cp</td>
<td>13</td>
<td>7</td>
<td>38%</td>
</tr>
<tr>
<td>29.</td>
<td>Control</td>
<td>Cp</td>
<td>13</td>
<td>8</td>
<td>38%</td>
</tr>
<tr>
<td>30.</td>
<td>Convey emotion</td>
<td>Cp</td>
<td>7</td>
<td>4</td>
<td>21%</td>
</tr>
<tr>
<td>31.</td>
<td>Entertainment</td>
<td>Cp</td>
<td>8</td>
<td>6</td>
<td>24%</td>
</tr>
<tr>
<td>32.</td>
<td>Feel good</td>
<td>Cp</td>
<td>14</td>
<td>12</td>
<td>41%</td>
</tr>
<tr>
<td>33.</td>
<td>Freedom</td>
<td>Cp</td>
<td>2</td>
<td>4</td>
<td>6%</td>
</tr>
<tr>
<td>34.</td>
<td>Image</td>
<td>Cp</td>
<td>7</td>
<td>11</td>
<td>21%</td>
</tr>
<tr>
<td>35.</td>
<td>less misunderstanding</td>
<td>Cp</td>
<td>3</td>
<td>3</td>
<td>9%</td>
</tr>
<tr>
<td>36.</td>
<td>Productivity</td>
<td>Cp</td>
<td>8</td>
<td>9</td>
<td>24%</td>
</tr>
<tr>
<td>37.</td>
<td>Safety</td>
<td>Cp</td>
<td>10</td>
<td>3</td>
<td>29%</td>
</tr>
<tr>
<td>38.</td>
<td>Social Interaction</td>
<td>Cp</td>
<td>7</td>
<td>6</td>
<td>21%</td>
</tr>
<tr>
<td>39.</td>
<td>Successful</td>
<td>Cp</td>
<td>3</td>
<td>8</td>
<td>9%</td>
</tr>
</tbody>
</table>
Table 1. The number of times a particular attribute was mentioned by a subject in total and per gender.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Cp</th>
<th>3</th>
<th>2</th>
<th>9%</th>
<th>9%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time consuming</td>
<td>Cp</td>
<td>3</td>
<td>2</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>Achievement</td>
<td>Vi</td>
<td>11</td>
<td>11</td>
<td>32%</td>
<td>48%</td>
</tr>
<tr>
<td>Independence</td>
<td>Vi</td>
<td>4</td>
<td>4</td>
<td>12%</td>
<td>17%</td>
</tr>
<tr>
<td>Belonging</td>
<td>Vt</td>
<td>26</td>
<td>12</td>
<td>76%</td>
<td>52%</td>
</tr>
<tr>
<td>Quality of life</td>
<td>Vt</td>
<td>13</td>
<td>9</td>
<td>38%</td>
<td>39%</td>
</tr>
<tr>
<td>Self-esteem</td>
<td>Vt</td>
<td>19</td>
<td>12</td>
<td>56%</td>
<td>52%</td>
</tr>
</tbody>
</table>

Observing the frequencies on the table above it becomes quite noticeable males’ preference for voice calls and females’ preference to SMS. However further details and relationship within the ladders is better represented in the hierarchical value map.

The hierarchical value map or HMV enable researchers to have a better understanding of the results from the laddering study without having study each ladder separately (Grunert, Beckmann, & Sorensen, 2001). “HMV is an estimate of cognitive structure for that group” (Grunert et al., 2001). When drawing a HMV a cut off level is determined to generate maps that will include only the most significant links (Grunert et al., 2001). In the HMVs illustrated in this paper the different levels of abstraction have been colour coded for easier identification. Concrete attributes are in orange; abstract attributes are yellow; functional consequences are in green, psychological consequences are in white, instrumental values are represented in salmon and terminal values purple. The arrows indicate the direction and strength of the relationships. The thicker the arrow the stronger is the connection.

The strongest relations in the ladder voice call/convenience/ get things done quicker/feel good /self esteem. However there is only a marginal difference to the ladder: SMS/convenience/ get things done quicker/feel good /self esteem Overall the preliminary findings from the general map (cut off 7) seem to confirm what is known about the key services consumers use on their mobile phones. Looking at isolated relationships, besides those already pointed, the other significative relationships are between conveniences and get things done quicker is expected. The relationship between convenience and considerate of others leading has considerable strength and seem to lead to belonging and self esteem.

Figure 2 is the HMV of males (cut off 5) and figure3 HMV (cut off 6) females. When confronting those two maps interesting results and relations are identified. The male’s map has six out of the seven concrete attributes identified indicating a more evenly use of the features and services. This can be related to the fact that most males use of their phones primarily for businesses purpose. In general, whenever talking about their personal usage of their mobiles the cost factor was raised. Another interesting finding relates to the ladder: design /looks (aA) – image (pC) - self-esteem (tV). Many would expect this chain to be stronger within the female group. However the attribute Design/looks is not represented at all in the female map contrasting with the strong relationship between designs, image and self-esteem ladder on the men’s map. Also by contrasting both maps it is noticeable that men have stronger ladders from voice. This indicates that they tend to talk on their mobiles more than women. This is visible but the strong relation voice calls/convenience- leads to achievement; self –esteem and belonging generating 3 ladders of similar strength. SMS is a strong attribute did not generate any real significant relationship between consequences and values in the males HMV. However the female HMV (cut off 6) has in SMS / convenience/ its strongest relation. The only attributes presented in the female’s map are SMS, voice call and camera. Incomplete ladders indicate a very weak relation. This is the case between camera and capture memories. The key ladder in the female map is SMS/convenience/get things done quickly / belonging. However the SMS / convenience relation lead to a series of ladders. This finding seems to contradict some popular belief that man uses more text message than woman.
Figure 1. HMV general cut off 7. This map was generated with no demographic filter. Concrete attributes are in orange; abstract attributes are yellow; functional consequences are in green, psychological consequences are in white, instrumental values are represented in salmon and terminal values purple. The arrows indicate the direction and strength of the relationships. The thicker the arrow the stronger is the connection.
Figure 2. HMV cut-off 5 filtering by male gender

Figure 3. HMV cut off 06filtering by female gender
6. Discussion and conclusion
This paper presented a summary of the preliminary findings to study mobile service adoption using means end chain and laddering. Many carriers and content creators are forecasting the change in mobile phone users expanding their services usage beyond calls and text messages. The initial findings of this study do not indicate that the existing scenario will change soon. Voice and text (SMS) are the key attribute consumers used because of its utmost convenience. This finding seems to be supported by the literature. For example (Bouwman et al., 2008) found that the usage patterns in Finland were reasonable stable in terms of preferences. The authors have indentified flexibility as a contributor factor between present and future usage.

The cognitive association maps presented may not be a definitive group for the m-services area. Rather we propose them as an illustration on the value of this approach. Nor are the maps mutually exclusive. Mobile services have a functional value as they are convenient but the decision to adopt may often be combined with the desire to feel part of a group or community of use. Indeed, the erosion of boundaries between work, home, leisure, learning and education, partly brought about by mobile technology, means that people may have multiple reasons for adopting or using services. Mobile technology can serve multiple needs including, family, friends, work and curiosity or learning.

7. Study limitations and further research
The results presented here are preliminary and partial. For a better indication of the patterns of mobile services use additional maps need to be studied. Also due to the qualitative nature of this research the aim only to gain deeper insights and understanding of consumer use of mobile services.
References


Impacts of Information Technology (IT) Outsourcing on Organizational Performance: A Firm-Level Empirical Analysis

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0234.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Offshoring / Outsourcing, Panel data, Empirical study, Business value of IT</td>
</tr>
</tbody>
</table>

Proceedings ECIS 2009
IMPACTS OF INFORMATION TECHNOLOGY (IT) OUTSOURCING ON ORGANIZATIONAL PERFORMANCE: A FIRM-LEVEL EMPIRICAL ANALYSIS

Ozcelik, Yasin, Fairfield University, Dolan School of Business, Fairfield, CT 06824-5195, USA, yozcelik@mail.fairfield.edu

Altinkemer, Kemal, Purdue University, Krannert Graduate School of Management, West Lafayette, IN 47907-2056, USA, kemal@purdue.edu

Abstract

We investigate the impact of Information Technology (IT) outsourcing on firm performance from several dimensions, including changes in labor productivity, improvements in financial and operational performance variables, and stock market valuation of IT outsourcing initiatives as measured by Tobin’s q. While our main objective is to better understand the economics of IT outsourcing, we also aim to contribute to the literature on the business value of IT in general. Our research contributes to the relevant literature from the following perspectives: (i) the change in the performance levels of firms due to IT outsourcing is measured against that of firms not outsourcing at all, (ii) panel data regression model is utilized in order to capture both cross-sectional and time-series differences among firms, (iii) the diversity of IT outsourcing initiatives is explicitly considered in the model, and (iv) a comprehensive data set covering the period between 1984 and 2007 is used.

Keywords: IT outsourcing, business value of IT, performance, panel data regression.
1 INTRODUCTION

Proliferation of the Internet and advancements in the Information Technology (IT) brought new opportunities to companies to conduct their businesses more efficiently than the past. The use of the Internet and IT not only changed the way firms do business, but also improved their existing processes. Today, firms can do business either by using their own resources and expertise, or by outsourcing some of the internal functions to outside contracting firms that specialize in certain functions.

Outsourcing, in its most succinct form, can be defined as the delegation to another party of the authority for the provision of services under a business contract that incorporates service-level agreements related to cost, quality, and the timeliness of deliverables. Given a diverse nature of business processes a firm has to manage today, it is nearly impossible for a firm to manage all of its processes by solely depending on its own expertise. Even if it is feasible, the firm may lose its focus and efficiency. Outsourcing some or all of non-core business processes can enable a firm to focus on core competencies, rather than services that fall outside of expertise. It will not only improve function effectiveness and flexibility by accessing a support network with highly qualified and specialized workforce, but also help firms control their costs and business risk by transforming high fixed costs to predictable expenditures.

Firms have been outsourcing various functions for years, ranging from assembly lines to Research & Development (R&D), from office documenting services to litigation. Although the very definition of outsourcing has not changed, its nature has evolved over time, expanding both the range and depth of services being outsourced. Firms today prefer outsourcing their business processes to firms that are highly specialized in using IT for business purposes. IT outsourcing, in this sense, is defined as “involving a significant use of resources (either technological or human resources) external to the organizational hierarchy in the management of IT infrastructure” (Loh & Venkatraman 1992a).

Despite its potential benefits, IT outsourcing may not be a straightforward decision for a firm because of potential risks involved, such as loss of control on outsourced activities, sharing critical company data with third-parties that may be used without consent of the firm, dependence on a firm whose internal operations may not be transparent generally due to autonomous nature of outsourcers, lack of knowledge on outsourcing process if the firm has not outsourced before, and existing firm culture that may resist to change.

Given the scale and popularity of IT outsourcing among firms today, the major organizational changes they entail, and the risk of failure, it is reasonable to expect that IT outsourcing has a significant and measurable effect on firm performance. Quantifying the impact of IT outsourcing on firm performance will not only help corporate managers make effective decisions on IT outsourcing, but also shed light on the prevailing debate over the outsourcing of U.S jobs to other countries (White House Joint Economic Committee Study 2001).

In this research, we investigate the impact of IT outsourcing on key firm performance variables by using a comprehensive data set compiled from primary and secondary sources. While our main objective is to better understand the economics of IT outsourcing, we also aim to contribute to the literature on the business value of IT in general. Our research contributes to the relevant literature from the following perspectives: (i) the change in the performance levels of firms due to IT outsourcing is measured against that of firms not outsourcing at all, (ii) panel data regression model is utilized in order to capture cross-sectional and time-series differences among firms, (iii) the diversity of IT outsourcing initiatives is explicitly considered in the model, and (iv) a newer data set covering the period between 1984 and 2007 is used.
2 LITERATURE REVIEW

Our research is strongly related to two streams of previous literature: the work on the business value of IT, and the more specialized and limited literature on IT outsourcing. In this section, we briefly survey previous studies in each of these areas that are most relevant to our research.

2.1 Business value of information technology

There is an extensive body of literature examining the business value of IT investments at the firm level. The roots of the debate can be traced back to 1990s when available data from 1980s and 1990s failed to show evidence of improved firm productivity due to investments in IT in the manufacturing sector (Morrison & Berndt 1990). This result, later called the “Productivity Paradox of IT,” was found to be even more pronounced in the service sector that had used over 80 percent of IT products during 1980s (Roach 1991). Researchers attempted to resolve the paradox by pointing out that the inability to show significant returns may be because of (i) measurement errors of outputs and inputs due to rapid price and quality changes in IT equipment, (ii) the time necessary for learning and adjustment, and (iii) mismanagement of IT resources by firms due to insufficient expertise to take advantage of the potential of using IT in traditional business environments (Brynjolfsson 1993). Other researchers rejected this paradox by providing empirical evidence to show a positive relationship between IT investments and firm performance (Brynjolfsson & Hitt 1995, Brynjolfsson & Hitt 1996, Lichtenberg 1995, Dewan & Min 1997, Bharadwaj et al. 1999, Kudyba & Diwan 2002, Anderson et al. 2003).

Not all studies, however, were able to show a clear payoff from IT investments. For example, Barua et al. (1991, 1995) both find that even though IT spending improves intermediate variables of organizational performance, such as capacity utilization, inventory turnover, or relative price, it does not necessarily lead to improvements in higher-level performance variables, such as Return on Assets (ROA) or market share.

Researchers also proposed innovative methods for measuring the business value of IT investments. For example, Brynjolfsson et al. (1994) show that effects of IT on firm productivity variables are substantially larger when measured over long time periods, since long-term returns represent the combined effects of related investments in organizational change. Devaraj and Kohli (2003) emphasize the importance of actual usage in driving the impact of IT on firm performance. In order to correctly measure the business value of IT, Kohli and Devaraj (2003) recommend that future studies explicitly report which complementary changes in business practices, such as Business Process Reengineering (BPR) and Enterprise Resource Planning (ERP), accompanied the IT investments. Barua and Mukhopadhyay (2000) emphasize that such analyses will help isolate and identify the effectiveness of complementary changes that lead to IT payoffs.

2.2 Information technology outsourcing

The main motivation for IT outsourcing is found to be cost reduction (Altinkemer et al. 1994, Gilley & Rasheed 2000). This is because paying for outsourcing generally costs less than maintaining equivalent services in-house. According to Malhotra (1995), factors that affect IT outsourcing decisions are reduction in operating costs, cost predictability due to fixed contract, sharing risk on technology investments, access to specialized expertise, political reasons that hinder internal IS efficiencies, and perception of efficiency of internal IS function. Clark et al. (1995) identify the changes in information technology, business trends, and technology management as the major factors that favour outsourcing. Outsourcing decisions may also be due to internal influence or imitative behaviour (Loh & Venkatraman 1992b).

The degree of IT outsourcing is found to be positively correlated to business and IT cost structure, and negatively related to the performance of the existing IT infrastructure (Loh & Venkatraman 1992a).
On the other hand, Wang et al. (2008) find that the level of business value created by IT outsourcing is contingent on firms’ core IT capability. That is, firms with superior core IT capability have an advantage in leveraging their outsourcing initiatives to enhance firm value. Jiang et al. (2006) find empirical evidence for improved cost efficiency as a result of IT outsourcing, but no change in the productivity and profitability of the outsourcing firms.

Researchers also investigated the effects of IT outsourcing versus insourcing on firm productivity. For example, Lacity and Hirschheim (1995) argue that most cost reductions achievable through outsourcing can equally well be achieved by the in-house IT function if it is given freedom to reorganize. They present eleven generic cost reduction strategies that internal IT departments can implement to reduce costs.

3 HYPOTHESES

Many potential benefits of outsourcing have been identified in the literature. Outsourcing firms often achieve immediate cost advantages (Jiang et al. 2006, Lei & Hitt 1995). Thus, outsourcing may be an attractive method of improving a firm’s financial performance, especially in the short run. Outsourcing firms may also achieve long run advantages compared to firms relying on internal production. In-house production increases organizational commitment to a specific type of technology and may constrain flexibility in the long run (Harrigan 1985). On the other hand, outsourcing firms can switch suppliers as new and more efficient technologies become available. In addition, outsourcing allows for quick response to changes in the environment (Dess et al. 1995). As a result, we expect firms engaging in IT outsourcing to experience an improvement in their performance variables during the period that starts when the initial outsourcing contract was signed. Thus, our first hypothesis is stated as follows:

Hypothesis 1. Firm performance improves as a result of IT outsourcing.

IT outsourcing initiatives vary considerably across firms; while some firms focus only on a single process, others may find it more profitable to outsource several business functions over the years. Diversification of outsourcing activities in this sense provides a measure of outsourcing intensity, and can affect the level of impact on firm performance (Gilley and Rasheed 2000). Besides, utilizing a diversity variable in our model provides the necessary linkage between IT outsourcing, which is implemented at the business unit level, with our performance variables, which are measured at the organizational level. Our second hypothesis is thus:

Hypothesis 2. The effect of IT outsourcing on firm performance increases with the diversity of outsourcing.

4 DESCRIPTION OF DATA RESOURCES

We used two leading online news sources, ABI/INFORM and Lexis/Nexis, to search and compile all press releases and news about firms announcing their IT outsourcing projects. We recorded the name of the outsourcing firm, the year of the initial outsourcing contract, and the type of the outsourcing projects conducted by each firm. We should note that our observations are unavoidably limited to those IT outsourcing projects that have been publicly announced. We may therefore have missed some of the projects that have not been announced, and consequently miscoded some companies as non-implementers when, in fact, they have undertaken an outsourcing project.

We observed that most of the projects announced were undertaken by large U.S. firms. Arguably, some of these firms may realize a higher level of performance benefits from IT outsourcing than others can do, which may affect the decision to undertake such a project. Therefore, in order to better understand the gains from IT outsourcing, we included in our data all of the firms in the Fortune 1000 list, regardless of whether they adopted IT outsourcing or not. Since firms frequently enter and exit the Fortune 1000 list every year, we took 1998 as the baseline year and included every firm that was listed...
in that year in our analysis. We used the Standard and Poor’s COMPUSTAT database to collect most of our data, which covers the period between 1984 and 2007. We also utilized the Information Week magazine’s Top 500 lists that publish total annual IT budgets of large companies.

5  ECONOMETRIC MODEL

Since our data has both cross-sectional and time-series components, we utilize a panel data regression analysis in order to test our hypotheses. We use the STATA statistical software package to run our regressions. By indexing firms by $i$ and years by $t$, we can express our panel data regression model as follows:

$$y_{it} = \mu_i + \sum_{k=1}^{m} \beta_k x_{itk} + \epsilon_{it} \quad i = 1, ..., N \quad and \quad t = 1984, ..., 2007.$$  

where $y_{it}$ is the dependent variable, $\mu_i$ is a separate constant term for each firm, $x_{itk}$ are independent variables, $m$ is the number of independent variables, $\beta_k$ are regression coefficients, $N$ is the number of firms, and $\epsilon_{it}$ is a classical disturbance term with $E[\epsilon_{it}] = 0$ and $\text{Var}[\epsilon_{it}] = \sigma^2 \epsilon$. The general form of our econometric model is as follows:

$$\log(\text{performance measure numerator})_{it} = \text{intercept}_i + \log(\text{performance measure denominator})_{it} + \text{implementation dummies}_i + \text{diversification dummies}_i + \text{firm controls}_i + \text{industry controls}_i + \text{year dummies}_i + \epsilon_{it}$$

We will perform the following regression diagnostics to avoid obtaining biased estimators. In order to see the level of possible multicollinearity in the regression model, we will create the correlation matrix among our independent variables, and visually examine its entries. As another check for multicollinearity, we will obtain the Variance Inflation Factors (VIF) for all of our independent variables, and check whether their values are less than the threshold level of 10. We will run our regressions by using robust covariance matrix for heteroskedasticity according to White’s (1980) procedure to prevent inconsistent covariance matrix estimates. Finally, we will utilize the Hausmann Test in order to decide whether to use the Fixed Effects or Random Effects panel data model to run our regressions. In what follows, we describe our regression variables and explain the rationale for including them in our model by providing support from the literature.

5.1  Dependent and independent variables

We construct various measures to calculate labor productivity, financial and operational firm performance, and stock market valuation (all serve as dependent variables) using standard approaches found in literature (Brynjolfsson & Hitt 1996, Bharadwaj et al. 1999, Brynjolfsson & Hitt 2000, Hitt et al. 2002).

Labor productivity is calculated by dividing total sales by the number of employees. Financial firm performance is measured via three well-known variables: Return on Assets (ROA), Return on Equity (ROE), and Return on Sales (ROS). Operational performance is measured through two variables: asset utilization that is calculated by dividing sales by total assets, and inventory turnover that is the cost of goods sold divided by inventory. Finally, the stock market valuation of IT outsourcing initiatives is measured by using Tobin’s $q$.

Tobin’s $q$ was first introduced as a predictor of a firm’s future investments (Tobin 1969, 1978). Since then, it has been extensively used as a measure of a firm’s intangible value, which is based on the assumption that the long-run equilibrium market value of a firm must be equal to the value of its
assets, giving a \( q \) value close to unity. Thus, a value of \( q \) greater than one implies an unmeasured source of intangible value generated by the firm.

Regression analyses featuring labor productivity or financial performance ratios have the advantage that they can capture different aspects of firm performance, and are commonly used in the literature. Their primary disadvantage is that they cannot adequately incorporate future gains from IT outsourcing, which could substantially exceed current or past gains. The use of Tobin’s \( q \) analysis mitigates these concerns because it is based on the expectations of future benefits that the firm may receive. For our purposes, we use the Tobin’s \( q \) specification adopted by Hitt et al. (2002), which relates market value of a firm to its total assets. The derivations of our performance measures are outlined in Table 1 below.

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>Numerator</th>
<th>Denominator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor Productivity</td>
<td>Sales</td>
<td>Number of Employees</td>
</tr>
<tr>
<td>Return on Assets</td>
<td>Pretax income</td>
<td>Assets</td>
</tr>
<tr>
<td>Return on Equity</td>
<td>Pretax income</td>
<td>Equity</td>
</tr>
<tr>
<td>Return on Sales (Profit Margin)</td>
<td>Pretax income</td>
<td>Sales</td>
</tr>
<tr>
<td>Asset Utilization</td>
<td>Sales</td>
<td>Assets</td>
</tr>
<tr>
<td>Inventory Turnover</td>
<td>Cost of goods sold</td>
<td>Inventory</td>
</tr>
<tr>
<td>Tobin’s ( q )</td>
<td>Market value</td>
<td>Assets</td>
</tr>
</tbody>
</table>

Table 1. Construction of Performance Measures.

Following Hitt et al. (2002), we use the logarithm of the numerator of each performance measure as a dependent variable, and the logarithm of its denominator as a control variable. This specification provides flexibility in the relationship between numerator and denominator, while retaining the interpretation as a performance measure.

Our key independent variable is the implementation dummy that indicates whether a firm conducts an IT outsourcing project during year \( t \). For each firm, it takes a value of one during and after the years an outsourcing project takes place, and zero otherwise. In order to examine possible lagged effects of outsourcing, we lag the implementation dummy up to three years. For example, if a firm started its outsourcing project in the year 2000, its implementation dummy with a 1-year lag will take a value of one for the year 2001 and beyond, and zero otherwise. Similarly, its implementation dummy with a 2-year lag will take a value of one for the year 2002 and beyond, and zero elsewhere. In order to test the second hypothesis, we use another dummy variable that indicates whether a firm diversifies its IT outsourcing activities over the years. This dummy variable takes a value of 1 for firms that are engaged in several IT outsourcing activities, and zero otherwise.

### 5.2 Firm-level control variables

Our firm-level control variables are firm size, IT budget, advertising expenditure, and market share. First, we use the natural logarithm of the number of employees as a proxy for firm size, as is traditional in the literature. Second, in order to observe the effect of IT outsourcing on firm performance across firms with varying degrees of IT investments, we utilize IT budget as another firm-level control variable. Third, there is ample evidence in the economics, marketing, and strategy literature supporting a positive relationship between advertising expenditure and firm performance (Comanor & Wilson 1974, Nelson 1974, Schmalensee 1978, Aaker 1991, Megna & Mueller 1991). Finally, market share is included as a control variable given that efficiency theory (Day & Montgomery 1983, Buzzell & Gale 1987), market power theory (Smirlock et al. 1984, Martin 1988), and studies on product quality assessment (Smallwood & Conlisk 1979) provide evidence for a
relationship between market share and firm performance. Jacobson and Aaker (1985) and Jacobson (1990) emphasize further that market share can serve as a proxy for other firm-specific assets (such as managerial skill) not specifically captured in this study. Based on the results of previous research, we expect a positive relationship between market share and firm performance.

5.3 Industry-level control variables

The literature on industrial organization economics supports the view that the structure of an industry impacts the performance of firms within the industry (Porter 1980). We, therefore, utilize three variables frequently used in previous research to account for variation in firm performance due to idiosyncratic characteristics of different industries at the 2-digit Standard Industrial Classification (SIC) level: industry concentration, industry capital intensity, and industry average Tobin’s q.

Consistent with the literature, industry concentration in our research is proxied by the four-firm concentration ratio, which is the total market share of the four largest firms in an industry. Industry capital intensity is calculated as the sum of all capital expenditures divided by the sum of all sales in an industry. It is included in our model to capture potential effects of entry barriers on firm performance. Since capital intensive industries are likely to face fewer competitors, incumbent firms could earn higher profits (Capon et al. 1990). This implies a positive relationship between industry capital intensity and firm performance. On the other hand, the relationship could be negative because high capital intensity requirements could take away resources from intangible investments, thereby reducing firm performance (Bharadwaj et al. 1999). Finally, the industry average Tobin’s q is utilized in our regressions regarding the stock market valuation of IT outsourcing projects. Inclusion of this variable enables us to justify pooling of data from multiple industries, as well as to control for idiosyncratic industry characteristics not adequately captured by the other industry control variables (Dess et al. 1990, Bharadwaj et al. 1999).

5.4 Time control variables

Following Hitt et al. (2002), we use separate dummy variables for each year to capture transitory, economy-wide shocks that may affect firm performance in our data set.

6 RESEARCH STATUS

We have finished compiling our primary data from the ABI/INFORM and Lexis/Nexis, where we have recorded the following information: (i) the names and stock tickers of all firms that publicly announced their IT outsourcing projects, (ii) the year of their initial outsourcing contracts, and (iii) the scope of their outsourcing projects. We have also finished collecting our secondary data from the COMPUSTAT database and the Information Week magazine. Currently, we are in the process of analyzing the data, and we expect to obtain our initial regression results by the time of the ECIS 2009.

References


ANTECEDENTS OF SUCCESS IN IS OFFSHORING PROJECTS – PROPOSAL FOR AN EMPIRICAL RESEARCH STUDY

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0388.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Offshoring / Outsourcing, Global software development, IT Strategy, Partial Least Squares</td>
</tr>
</tbody>
</table>
ANTECEDENTS OF SUCCESS IN IS OFFSHORING PROJECTS – PROPOSAL FOR AN EMPIRICAL RESEARCH STUDY

Westner, Markus, Dresden University of Technology, 01062 Dresden, Germany, markus.westner@tu-dresden.de

Abstract

The paper presents a research model and a measurement instrument for a research-in-progress study on the antecedents of success in IS offshoring projects. In this empirical-confirmatory study, we intend to analyse the impact of the constructs “offshoring expertise”, “trust in offshore service provider”, “project suitability”, “knowledge transfer”, and “liaison quality” on offshore project success. Constructs and indicators are derived from an extensive literature review. We plan to formulate a structural equation model and to test it using partial least squares (PLS) as an analysis technique. Our research model addresses the paucity of research that quantitatively examines offshoring success.

Keywords: offshoring, outsourcing, project success, partial least squares.

1 INTRODUCTION

Information systems (IS) offshoring describes the transfer of IS services to an offshoring service provider (OSP) in a near or far away country. This OSP can be an internal subsidiary (so-called “captive offshoring”), a partially-owned unit, or an external service provider (so-called “offshore outsourcing”). The services themselves are partially or totally transferred. (Carmel and Agarwal 2002, Hirschheim et al. 2005, Jahns, Hartmann and Bals 2006/7, Mirani 2006, Niederman, Kundu and Salas 2006, Rajkumar and Mani 2001, Srivastava, Teo and Mohapatra 2008)

High labour cost differentials in comparison to western countries and the resulting cost savings are the main reasons why companies engage in IS offshoring. Accordingly, the market volume for offshoring of IT services has been growing fast in the last years, with India being the most popular offshoring destination (Knapp, Sharma and King 2007, Metters and Verma 2008, Poornima 2008). Application development and maintenance activities, where labour constitutes a significant share of total costs, are especially likely to be performed offshore (Bitkom 2005, Boes, Schwemmle and Becker 2004, William, Mayadas and Vardi 2006). However, recent studies among companies worldwide indicate that a large number of companies that engaged in IS offshoring are not fully satisfied with their engagements’ performances (Bright 2008, Computerwoche 2008).

The situation is especially noticeable in Germany. There, offshoring levels are rather low: only 6% of all companies source IS services from abroad in contrast to 64% that already use domestic IS outsourcing (SchAAF and Weber 2005, ZEW 2007). Additionally, German companies also experience difficulties in performing IS offshoring successfully (Prehl 2008). This seems to be due to language and cultural barriers (Dibbern, Winkler and Heinzl 2006, Mertens 2005, Wiener 2006).

IS offshoring is worth being researched as a domain of its own because it has specific characteristics that distinguish it from the well-researched field of IS outsourcing. In IS offshoring, service delivery occurs under the additional condition of “distance” between service provider and consumer in terms of physical distance, time zone differences, or cultural differences. Additionally, complexity increases due to the higher degree of geographical dispersion among team members. Finally, IS offshoring arrangements often create additional organisational challenges because offshore staff partially replaces domestic onshore staff. This increases the importance of knowledge transfer, knowledge absorption, project management, and HR management to ensure successful service delivery. (Chua and Pan 2008,

Research in IS offshoring has been growing in the last years and journals such as the “MIS Quarterly” (vol. 32, issue 2) or “Information Systems Frontier” (vol. 10, issue 2) have published issues addressing the phenomenon. IS offshoring research, in contrast to IS outsourcing research, is primarily case study based and qualitative, which shows that it is still in its initial, theory-building stage (Dibbern et al. 2004, King and Torkzadeh 2008). The research situation is furthermore characterised by studies that employ a project or organizational level of analysis, focus on India as an offshoring destination, and investigate “success/outcome factors” or “economic value” as research topics (King and Torkzadeh 2008).

Derived from the special characteristics of IS offshoring, we investigate how “project suitability”, “knowledge transfer”, and “liaison quality” as well as how the constructs “trust in OSP” and “offshoring expertise” impact offshoring project success at German companies. We employ a confirmatory-quantitative research approach to address this objective. In this sense we follow the current state in IS offshoring research to focus on “success/outcome factors” with “projects” being the level of analysis. However, we add original content through our research model that partially builds upon recent research results but also incorporates new aspects. We ensure methodological originality by gathering a broad empirical dataset and by analysing our research model with structural equation modelling as a tool for analysis. Finally, we address the paucity of research that quantitatively investigates offshoring in the context of German businesses.

We focus our research along four dimensions: (1) our regional focus is Germany, (2) we focus on the offshore consuming client’s perspective, (3) our unit of analysis is offshoring projects, i.e., not the arrangement or relationship between service consumer and provider in total, and (4) we focus on application development or maintenance projects.

2 THEORETICAL FOUNDATION AND RESEARCH MODEL

2.1 Research model overview

Our proposed research model argues that offshoring expertise has a direct positive effect on offshore project success. Additionally, it is positively associated with project suitability, knowledge transfer, and liaison quality which act as mediators for offshore project success. Trust in the OSP is positively associated with knowledge transfer, liaison quality, and offshore project success. Figure 1 illustrates the model. The subsequent sections develop and describe its constructs and their relationships.

![Figure 1. Research model on antecedents of offshore project success. A plus (+) symbol denotes a positive relation between constructs.](image-url)
2.2 Measuring offshore project success

Offshore project success is the dependent variable in our research model. As Erickson and Ranganathan (2006) show, success can be understood and measured in multiple ways, including “the organization’s satisfaction with the results of outsourcing (Grover, Cheon and Teng 1996), an expectations fulfillment view (Lacity and Willcocks 1998), a cost/benefit approach (Wang 2002), a psychological contract perspective on fulfilled obligations (Koh, Soon Ang and Straub 2004), and a strategic fit view of success (Lee, Miranda and Kim 2004)” (Erickson and Ranganathan 2006).

Several studies measure success as the satisfaction of outcomes, sometimes calibrated by initial expectations (Balaji and Ahuja 2005, Grover et al. 1996, Dahlberg and Nyrhinen 2006, Wüllenweber et al. 2008). In their extensive review of IS outsourcing success definitions and measures, Dahlberg and Nyrhinen (2006) find that satisfaction with outcomes can be evaluated along four categories which are “strategic factors”, “economic factors”, “technological factors”, and “social factors”. Additionally, overall satisfaction forms a part of their success definition.

Strategic, economic, technological and social outcome factors may also apply to projects but they are not applicable in all cases. For example one might think of projects that completely lack a specific strategic proposition. Since a project is by definition an effort bound by “schedule”, “budget”, “functionality”, and “quality” (Erickson and Ranganathan 2006), it rather makes sense to use these dimensional factors together with overall satisfaction as an operationalization of offshore project success.

Therefore, our paper interprets the dependent variable offshore project success as the perceived satisfaction with the outcome of the offshore project in total, and with the dimensions of schedule, budget, functionality, and quality.

2.3 The role of project suitability for project success

We define project suitability for offshoring as the sense that a project’s attributes and its task characteristics make it more amenable for delivery in a dispersed, inter-cultural environment, i.e., in an offshoring setting.

The identification of suitable project candidates for offshoring is one of the first activities before engaging in an IS offshoring arrangement. Once identified, these offshoring candidates then represent the core objects in the subsequent implementation of IS offshoring. Therefore, research and practice indicate that the identification of suitable project candidates is a main step in pursuing an IS offshoring endeavour. (Aron and Singh 2005, Chua and Pan 2006, Kumar and Palvia 2002)

Research in IT outsourcing has shown that there is a link between the function being outsourced and arrangement success (Fisher, Hirschheim and Jacobs 2008). They suggest focusing on routinely performed and non-core functions. Applying the lens of transaction cost theory and operations management models, Stratman (2008) finds that well understood, standardized service processes that are non-core are best candidates for successful offshoring. Stringfellow, Teagarden and Nie (2008) show that it is more challenging to offshore complex, loosely defined and non-standardized tasks that require complex judgments and implicit knowledge. If projects or tasks show these characteristics, offshore delivery incurs additional costs which might threaten project success. King (2008) suggests a framework for determining whether an IS activity should be considered for offshoring. He posits that activities should be kept in-house if they require proximity and the risk of offshoring is too great, or if the activity is too business-critical. Schaffer (2006) develops a similar framework that suggests refraining from offshoring projects which are very short, require a tremendous amount of personal interaction, are of high security and extreme criticality for the business. Mirani (2006) states that small applications or components of low complexity, for which specifications can be communicated completely, and whose development process is highly structured, are more likely to be successfully delivered in an offshore arrangement.
Since most of these studies are conceptual in nature or rely on a small set of empirical data, we carried out a qualitative pre-study with 47 German offshoring experts from different companies to find out whether project suitability is actually important for project success and what the respective evaluation criteria could be (Westner and Strahringer 2008). In the interviews, these experts confirmed that a project’s characteristics and its suitability for offshoring have a strong impact on later project success. Criteria such as project size, project duration, operating language, degree of codification, and business specificity were most frequently mentioned as determining a project’s suitability for offshoring with regard to successful delivery. If projects have a certain size and duration, the project language is English, the degree of codification is high, and business specificity or required domain knowledge is low, it takes less time and effort to make OSP staff fully productive. Therefore we hypothesize:

\[ H1: \text{Project suitability is positively associated with offshore project success.} \]

### 2.4 The role of knowledge transfer for project success

Following Davenport and Prusak (1998) and Lee, Huynh and Hirschheim (2008), we define knowledge as “a fluid mix of experience, values, contextual information and expert insight that provides a framework for evaluating and incorporating new experiences and information” (Davenport and Prusak 1998). Knowledge transfer as an outcome is the result of (1) the exchange of knowledge as a systematic activity between individuals and organizations (Chua and Pan 2008, Wang, Tong and Koh 2004) and (2) the ability to absorb the knowledge, to apply it and to use it in project delivery (Orlikowski 2002, Oshri, Kotlarsky and Willcocks 2007). Common terms to describe these two aspects are “knowledge transition” for (1) and “knowledge integration” for (2).

Application development and maintenance represent knowledge-intensive work. Knowledge pertinent to applications can either be explicit, such as software documentation, technical specification, or standardized development processes, but it can also be tacit, such as practices like norms of communication or non-specified processes and activities. (Chua and Pan 2006, Nicholson and Sahay 2004)

To profit from the economic benefits of offshoring, offshore staff must actually replace more expensive onshore staff (Chua and Pan 2008). Accordingly, all project-relevant explicit and implicit knowledge needs to be transferred to offshore staff. This knowledge transfer happens at the beginning of an offshoring project but is also a continuous activity during the whole project. Correspondingly, offshoring process models used in the industry and proposed by research recognize knowledge transfer as a specific activity (Bugajska 2007, Oshri et al. 2007, Voigt, Novak and Schwabe 2007).

Applying this perspective, the importance of knowledge transfer becomes obvious and has also been addressed in research. A case study by Chua and Pan (2008) examines how a financial institution transferred knowledge within a captive offshoring arrangement, and highlights knowledge transfer’s importance for successful service delivery. Another case study, by Oshri et al. (2007) from the OSP perspective, investigates best practices for managing dispersed knowledge among on- and offshore sites and acknowledges that knowledge transfer is a key part of successful offshoring. Previously, Ganesh and Moitra (2004) identified knowledge transition and the absorptive capacity of the OSP as one of the critical success factors for successful service transition in the context of business process outsourcing and offshoring. Rottman and Lacity (2008) develop best practices to ensure success in IS offshoring. Most of these best practices are fundamentally linked to facilitate and ensure successful knowledge transfer. Finally, in one of the few recent empirical-confirmatory studies, Lee et al. (2008) examine IS outsourcing arrangements between Korean firms and find significant support for the hypothesis that knowledge sharing is positively related to the success of outsourcing.

Thus, we can conclude that if knowledge transfer is successful, (1) offshore staff is more productive because it has the required know-how to perform project tasks and (2) onshore staff can be replaced as initially planned because it does not hold exclusive knowledge anymore. Based on this understanding we hypothesize that:
H2: Knowledge transfer is positively associated with offshore project success.

2.5 The role of liaison quality for project success

We define liaison quality as the degree of connectedness between onshore and OSP staff in the aim to achieve specified goals, i.e. in our case, a project’s objectives (Winkler et al. 2008). Liaison between staff should incorporate reciprocity and closeness (Xu and Yao 2006).

The environmental circumstances of IS offshoring delivery have negative impacts on liaison quality. Due to distance, communication frequency between team members decreases, collaboration is aggravated, and individuals tend to feel themselves not as equal parts of a team (Herbsleb and Mockus 2003, Xu and Yao 2006).

Therefore, research in IS offshoring emphasizes the importance of liaison quality on offshoring success. Erickson and Ranganathan (2006) highlight the need for clear roles, responsibilities, communication mechanisms, and conflict resolution in the management of global virtual teams. Rottman (2008) recognizes liaison quality’s impact on success and suggests building personal connections between OSP staff and client staff, for example by regular site visits and face-to-face meetings. Furthermore, he proposes to integrate offshore staff into onshore staff and synchronize training of offshore employees with internal training efforts. Similarly, Heeks et al. (2001) find that a high degree of congruence between provider and client improves chances for project success regarding schedule and budget. They recommend building bridging relationships between involved team members and using “straddlers”, i.e., dedicated individuals who are responsible for facilitating and moderating the interaction between on- and offshore staff. Levina and Vaast (2008) mention that liaison quality lessens negative effects of distance and thereby improves performance. They mention good onshore middle managers, frequent communication, constructive communication, and the efficient usage of technology as practices to improve liaison quality. Other research shows congruent findings and mentions the positive effect of liaison quality on performance achieved by liaison engineers and personal relationships (Kobitzsch, Rombach and Feldmann 2001), facilitation of informal communication (Herbsleb and Mockus 2003), the presence of expert intermediaries, and supplier presence on-site (Carmel and Nicholson 2005).

Achieving satisfactory levels of liaison in an offshore project setting seems to be challenging due to the negative effects of cultural and physical distance. However, liaison between on- and offshore staff is vital for collaboration, working efficiency, and productivity. Thus liaison quality directly impacts offshore project success and we hypothesize:

H9: Liaison quality is positively associated with offshore project success.

2.6 The impact of offshoring expertise on project suitability, knowledge transfer, liaison quality, and offshore project success

We define expertise as a certain degree of individual or organizational experience in managing or conducting offshoring in a more efficient and thus successful manner. In organizational research this is commonly referred to as “learning curve effects” or “experience curve effects” (Day and Montgomery 1983, Ghemawat 1985).

As mentioned in the introduction, delivery in an offshoring context raises multiple challenges for all involved parties. Individuals as well as organizations can benefit from best practices and experiences they have had in past engagements. Thus they can cope better with offshoring-specific challenges.

The positive impact of expertise on diverse activities of the offshore process and directly on offshore project success has already been addressed in research. Carmel and Agarwal (2002) develop a maturity model for companies engaging in offshoring and give recommendations how to move along this maturity curve. In a study of an eight-year offshore outsourcing alliance, Kaiser and Hawk (2004)
describe how the alliance evolved towards a more beneficial co-sourcing model, for both the consumer and the supplier. Similarly, Mirani (2006) shows how increasing expertise leads to a change in the offshoring relationship from rather simple to more sophisticated arrangements. Rottman and Lacity (2006), in their study of offshoring practices at 21 U.S. companies, also find positive effects of expertise on offshoring success.

Higher levels of organizational and individual expertise help to cope with the potential challenges of offshoring and thus increase the probability of project success. Thus, expertise has a positive impact on all three mediating constructs because based on past experiences it is rather likely a company selects projects which are most suitable for offshoring. Additionally, the organization and the individuals know how to manage knowledge transfer and improve liaison quality based on their expertise. Thus, we hypothesize:

H4: IS offshoring expertise is positively and directly associated with offshore project success.
H5: IS offshoring expertise is positively associated with project suitability.
H6: IS offshoring expertise is positively associated with knowledge transfer.
H7: IS offshoring expertise is positively associated with liaison quality.

2.7 The impact of trust in the OSP staff on knowledge transfer, liaison quality, and offshore project success

We define trust as the “expectation that an actor (1) can be relied on to fulfil obligations […], (2) will behave in a predictable manner, and (3) will act and negotiate fairly when the possibility for opportunism is present” (Zaheer, McEvily and Perrone 1998). Trust can thereby take the form of interpersonal or inter-organizational trust. Inter-personal trust is trust placed by the individuals in their individual opposite member. Inter-organizational trust is trust placed in the partner organization by the members of a focal organization (Lee et al. 2008, Zaheer et al. 1998).

Trust is important within an IS offshoring context because it is a facilitator and precondition for activities such as knowledge transfer, but also for collaboration among team members in general. Trust thereby increases the room to manoeuvre within an arrangement beyond the specifications of a contract. If individuals or organizations trust their counterparts, they are more willing to cooperate and to put in extra effort if needed. (Lee et al. 2008)

In IS outsourcing research, the role of trust as an important arrangement attribute has been widely recognized. Higher levels of trust seem to positively influence the relationship between client and vendor (Grover et al. 1996, Lee and Kim 1999, Winkler et al. 2008). Recent empirical-confirmatory studies show that trust is positively related to the extent of knowledge sharing (Lee et al. 2008) and that trust, mediated by cooperative learning, has a significant positive influence on knowledge transfer (Park and Im 2007). With respect to IS offshoring research, trust is mentioned as a critical success factor regarding the interface between offshore consumer and supplier (Jennex and Adelakun 2003). Kaiser and Hawk (2004) confirm this in a case study and perceive the creation of trust as a best practice for successful offshoring because it facilitates collaboration between on- and offshore staff. Thus, offshore staff becomes productive in a short time and projects progress faster. Winkler et al. (2008) show that trust positively influences the degree of connectedness between an offshore consumer and a service provider in their aim to achieve specified goals. Rottman (2008) illustrates how trust facilitates the knowledge transfer within an offshoring arrangement because it increases the willingness to share knowledge and collaborate.

Apparently, trust seems to influence knowledge transfer because individuals are more likely to share knowledge if they trust each other. This is especially important when it comes to implicit and thus sticky knowledge. Additionally, trust fosters and facilitates collaboration, communication, and – more generally – increases liaison quality among team members. Thus we hypothesize:
H8: Trust in OSP is positively associated with liaison quality.

H9: Trust in OSP is positively associated with knowledge transfer.

Similar to our hypotheses concerning the construct “offshoring expertise”, we could hypothesize a direct effect of trust on offshore project success. However, the studies mentioned above and other non-IS research (c.f. literature overview by Lee et al. 2008) do not support such an association. They do not link levels of trust directly to success or outcome but rather examine the impact of trust on constructs such as relationship or partnership quality, thus assuming a fully mediated effect. We nevertheless include a potential direct impact of trust on offshore project success in our model. However, because the theoretical backing is weak, we treat it with the necessary prudence regarding its direct impact in the model on success and later interpretation of results:

H10: Trust in OSP is positively and directly associated with offshore project success.

3 ANALYSIS

We follow an empirical-confirmatory research approach. Our units of analysis are offshore application development or maintenance projects at German corporations. Section 3.1 explains our measurement instrument. Section 3.2 outlines the intended analysis method using Partial Least Squares (PLS) as a measurement technique.

3.1 Measurement instrument

We measure each construct by an indicator set, with each indicator being assessed on a 7-point Likert scale from “1 – strongly disagree” to “7 – strongly agree”. As far as possible we used indicators that were developed and applied in previous research and adapted them if necessary. Only for project suitability we had to define indicators ourselves since there is no known existing instrument from previous studies. We constructed these indicators on the basis of a previous qualitative study that examined factors determining a project’s suitability for offshoring (Westner and Strahringer 2008).

All constructs are measured reflectively, with the exception of project suitability which is operationalized by means of a formative multi-item scale. The reason for using a formative measurement model is that the indicators to measure project suitability are causing the construct instead of being caused by or reflecting it. Thus, the construct needs to be measured by a formative approach. (Diamantopoulos and Winklhofer 2001)

At this stage, we are still in the progress of conducting pre-tests of the instrument with academic staff at different universities and with industry experts and practitioners. The pre-test intends to ensure validity, quality, and comprehensibility of the questions and its presentation. Table 1 illustrates our preliminary measurement instrument. Column 1 contains the research model’s constructs, Column 2 the respective label, Column 3 the corresponding indicators, and Column 4 respective references to literature, if applicable.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Label</th>
<th>Indicator (“1 - fully disagree” – “7 fully agree”)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offshoring expertise</td>
<td>EXP1</td>
<td>At the time the project was started… most project team members had already gathered work experience in offshore arrangements.</td>
<td>Carmel and Agarwal 2002, own</td>
</tr>
<tr>
<td></td>
<td>EXP2</td>
<td>…our company had already performed other projects in an offshore arrangement.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EXP3</td>
<td>…our company had dedicated processes and organizational structures in place to plan, manage and execute offshore arrangements.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EXP4</td>
<td>Overall, at this time, we considered our level of offshoring expertise as being high.</td>
<td></td>
</tr>
<tr>
<td>Construct</td>
<td>Label</td>
<td>Indicator (“1 - fully disagree” – “7 fully agree”)</td>
<td>References</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------</td>
<td>---------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Trust in offshore service provider</td>
<td>TRU1</td>
<td>Makes beneficial decisions to us under any circumstances.</td>
<td>Lee et al. 2008</td>
</tr>
<tr>
<td></td>
<td>TRU2</td>
<td>Is willing to provide assistance to us without exception.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TRU3</td>
<td>Reliably provides pre-specified support.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TRU4</td>
<td>Is honest.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TRU5</td>
<td>Cares about us.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TRU6</td>
<td>Overall, we had the impression that we could trust the offshore service provider staff.</td>
<td></td>
</tr>
<tr>
<td>Project suitability</td>
<td>SEL1</td>
<td>The offshored project’s volume in terms of workload was rather large.</td>
<td>Own</td>
</tr>
<tr>
<td></td>
<td>SEL2</td>
<td>The offshored project’s duration was rather short (reversely coded).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEL3</td>
<td>Most of the project communication between staff was done and documentation was available in English.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEL4</td>
<td>Most of the information and knowledge concerning the project was well codified and documented.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEL5</td>
<td>The project required business-specific know-how of all staff members (reversely coded).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEL6</td>
<td>Today, we would say the project was suitable for offshore delivery.</td>
<td></td>
</tr>
<tr>
<td>Knowledge transfer</td>
<td>KNO1</td>
<td>Business proposals and reports.</td>
<td>Lee et al. 2008, Simonin 1999</td>
</tr>
<tr>
<td></td>
<td>KNO2</td>
<td>Manuals, models, and methodologies.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KNO3</td>
<td>Know-how from work experience.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KNO4</td>
<td>Each other’s know-where and know-whom.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KNO5</td>
<td>Expertise obtained from education and training.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KNO6</td>
<td>The offshore service provider staff had learned a great deal about the project-related technology/process know-how.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KNO7</td>
<td>The offshore service provider staff had greatly reduced its know-how related reliance or dependence upon us since the beginning of the project.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KNO8</td>
<td>Overall, we were satisfied with the knowledge transition from us to offshore service provider staff within the project.</td>
<td></td>
</tr>
<tr>
<td>Liaison quality</td>
<td>LIA1</td>
<td>Communicated frequently and openly.</td>
<td>Erickson and Ranganathan 2006, Xu and Yao 2006</td>
</tr>
<tr>
<td></td>
<td>LIA2</td>
<td>Developed a mutual understanding of the respective ethnic and corporate cultures.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LIA3</td>
<td>Members each perceived themselves as equal and recognized members of the project team.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LIA4</td>
<td>Formed close individual working connections with each other.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LIA5</td>
<td>Overall, we were satisfied with the working liaison between our staff and offshore service provider staff.</td>
<td></td>
</tr>
<tr>
<td>Offshore project success</td>
<td>SUC1</td>
<td>We were satisfied with the project performance regarding time schedule.</td>
<td>Erickson and Ranganathan 2006, Grover et al. 1996, Wüllenweber et al. 2008</td>
</tr>
<tr>
<td></td>
<td>SUC2</td>
<td>We were satisfied with the project performance regarding budget.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SUC3</td>
<td>We were satisfied with the project performance regarding expected functionality.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SUC4</td>
<td>We were satisfied with the project regarding expected quality.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SUC5</td>
<td>We were satisfied with the overall outcomes from our offshoring arrangement.</td>
<td></td>
</tr>
</tbody>
</table>

*Table 1.* Measurement instrument for research model.

Proceedings ECIS 2009
Together with these indicators we will also gather data regarding whether it was a recent or old project, the offshoring country, the language in which the project was carried out, whether it was a captive or outsourcing offshoring project, project duration in months, project volume in man months and currency, respondents’ personal years of offshoring expertise, career position now, role on the project at the time it was conducted, whether the respondent resided on- or offshore, and to what industry the company belongs. This additional data will be used for better understanding and will serve as control variables.

3.2 Analysis method

We will transform our research model into a structural equation model and test it using PLS analysis. PLS is especially suitable in research areas where theory-building is still in its early stage. Additionally, it works with non-normal distributed data as well as with small sample sizes and it allows for incorporating formative and reflective construct measurement. (Gefen, Straub and Boudreau 2000, Herrmann, Huber and Kressmann 2006)

4 NEXT STEPS

After data collection, we will analyse the returned data and test our research model. In doing so, we will follow the generally accepted analysis principles when using PLS. The quality of our reflective indicators will be evaluated regarding content validity, convergent validity, and discriminant validity (e.g., Gefen et al. 2000, Huber et al. 2007). The specification quality, which is important for formative indicators, is ensured by a pre-study where we conducted interviews with 47 German offshore experts at different companies. Further tests for quality of the formative indicators will focus on its predictive quality, reliability, discriminant validity, and occurrence of multi-collinearity (e.g., Diamantopoulos and Winklhofer 2001, Huber et al. 2007). We will test for non-response bias by comparing the data of early returned questionnaires with later returned ones. Finally, we will actually test our structural model and its hypotheses and analyse for the effect of control variables. The final result’s interpretation will also reflect on the generalizability of findings to other countries.

References


MANAGING AN IT CARVE OUT AT A MULTI-NATIONAL ENTERPRISE
A TEACHING CASE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0725.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Teaching Case</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Teaching Case, Mergers and Acquisitions, IT Project Management, IT/IS management</td>
</tr>
</tbody>
</table>
MANAGING AN IT CARVE-OUT AT A MULTI-NATIONAL ENTERPRISE
TEACHING CASE DESCRIPTION

Fähling, Jens, Technische Universität München, Chair for Information Systems,
Boltzmannstr. 3, 85748 Garching b. München, Germany, faehling@in.tum.de
Leimeister, Jan Marco, Universität Kassel, Chair for Information Systems,
Nora-Platiel-Str. 4, 34127 Kassel, Germany, leimeister@uni-kassel.de
Yetton, Philip, University of New South Wales, Australian School of Business, Building,
Kensington, NSW 3052, Australia, phily@agsm.edu.au
Krcmar, Helmut, Technische Universität München, Chair for Information Systems,
Boltzmannstr. 3, 85748 Garching b. München, Germany, krcmar@in.tum.de

Abstract

Mergers, acquisitions and divestments, including the carve-outs of business units or parts of them, are standard strategies used by multi-divisional organizations to adjust their business portfolios. Carve-out projects are critically dependent on their management of IT. Systems, which have been integrated in order to deliver seamless and efficient IT operations, must now be pulled apart under demanding time and compliance constraints. In 2007, Delta IT Consulting (DIC), one of France’s biggest IT-service provider, sold one of its three service provider divisions, IT Product Services (IPS). This division employed about 3,500 employees in 20 countries and previously generated 0.7 billion of DIC’s 3.8 billion euro revenues. DIC itself is a division of Delta Corporation - a French high-tech company and leading player in a wide array of businesses, industries and countries around the world. This teaching case challenges the reader to analyse and manage the IT carve-out as a critical component within the divestment project. The case includes insights into strategic and organizational challenges of planning and managing an IT carve-out project.

Keywords: teaching case, IT carve-out, IT project management, IT management.

[The names, dates and numbers have been changed for business in confidence reasons.]
1. INTRODUCTION

On December 4th, 2007, Delta IT Consulting (DIC) entered into a contract to sell its IT Product Services (IPS) division as part of “Organize4Profit”, its corporate organization-wide restructuring program. This strategic business unit (SBU), which represented 0.7 billion of the IT services unit’s 3.8 billion euro revenues, was one of three SBUs in DIC. A new CEO, Thomas Hutton, had been appointed in October with a mandate from Jean Claude, the Chairman of Delta Corporation, to restructure DIC, consolidating its sites in France from 48 to 21. The media coverage was intense.

IT was one of ten work streams in the carve-out project. All members of the IT carve-out team (IT-cut) were appointed at short notice. Nobody in the team had any experience of working on a carve-out project. However, everyone agreed that the following months would be extremely hard and time-consuming. Team members had to sign non-disclosure agreements, which were not part of their regular contracts. These were not only to guarantee confidentiality for the negotiations but also to ensure compliance with anti-trust regulations. Some team members wondered if they would have a job when the carve-out was signed off.

“We have to hurry,” said Alain Croix, Chief Information Officer at Delta IT Consulting (DIC) in his first communication with IT-cut, “Time is short. The sale closes on April 1 next year. It is only three and a half months to the hand-over.”

This case study examines the strategic and organizational challenges of planning and managing an IT carve-out project. Acquisitions and divestments are critical strategic decisions in global organizations such as Delta Group, with major political and reputation consequences, and significant impacts on their bottom lines. For example, the cost of an IT carve-out is frequently significant relative to the other sale costs. More importantly, IT carve-outs must meet deadlines and functionality imposed by the sale contract or incur substantial penalties. Therefore, the success of a divestment, such as the one made here by DIC, was critically dependent on IT-cut’s performance.

1.1 Business background/ context

Delta Group is a major global organization, operating in 160 countries. In 2007, it employed 372,000 people worldwide, generating revenues of more than 71 billion euro, with a net income of over 2.4 billion euro.

DIC was part of the Digital Processing and Transmitting business area, offering a broad portfolio of multi-vendor IT solutions and services for the private and public sector. With 10,000 customers, DIC was one of the world’s leading providers of IT services, generating revenues in 2007 of 3.8 billion euro. Its portfolio contained a range of services from consulting and system integration to the management and operation of IT infrastructure and entire business processes.

DIC was divided into three business divisions. The IT Solution Services division covered the first phase in the IT lifecycle. This included consulting services and the definition of customer solutions, including SAP, system integration, and IT and process consulting. The IT Operations Services division covered the second phase in the IT lifecycle. This included both the outsourcing of IT intensive business processes, including HR and financial services, and the management of data centres, desktop environments, LAN/WAN and call centres. The IT Product Services division (IPS) covered the third phase in the IT lifecycle, specifically maintenance and infrastructure services. IPS delivered platform-

---

1 In 2007, DIC was one of five core ‘business areas’ in Delta Group, a global corporation in electronics, engineering, energy and healthcare.

2 In April 2007, Delta embarked on a major restructuring project Organise4Profit.
independent IT infrastructure services and solutions that addressed customers’ high availability, consolidation, migration and lifecycle management requirements, supporting a range of third-party platforms.

![ DIC Business Portfolio](image)

**Figure 1: DIC Business Portfolio**

In 2007, DIC was expecting to lose 109m euro, down from a profit the previous year of 40m euro. In April 2007 Delta Group had embarked on a major restructuring project, Organise4Profit. As part of that initiative, Delta Group was looking for a solution for this challenged SBU.

2. **THE CARVE-OUT**

Starting in September 2007, there had been rumors in the media about major changes at Delta, including Delta IT Consulting (DIC), to help the Chairman achieve his target of 5%-6% profit for all SBUs by mid-2009.

Off the record, Delta managers had admitted that DIC had too many contracts that were not profitable. This had led to major write-downs. The French business press reported early in October that all of DIC would be sold.

In November, the press speculated that Siemens had made an offer but this was denied. There was further speculation about negotiations with other companies from the USA and Germany, including Gedas, the IT service provider at Volkswagen, and International Computer Services (ICS), a 50-50 joint venture between Delta and International Computer Industries.\(^3\)

Finally, on December 4, it was announced that Delta was selling the IPS business to ICS. IPS by itself was one of the largest IT service company worldwide, employing 2,300 people in 17 countries and with a global network of 60,000 service partners and system integrators delivering services in 140 countries. Selling to ICS was logical because IPS’s customers were also customers of ICS. Selling IPS to a competitor would have been a serious threat to ICS, with the new owner acquiring both major service contracts and access to valuable insights into ICS technology.

The importance of the IPS acquisition was highlighted in the ICS Annual Report 2008, page 77:

> “The Services Division of ICS commenced business under ICS management on April 1, 2008. IPS and ICS are a perfect match. This is clearly reflected in the smooth integration. Ninety four percent of all customer contracts and 86 percent of former IPS employees transferred to the new Services Division on April 1. This is a major strategic

\(^3\) ICS delivers innovative, high quality IT products, infrastructure solutions and services to corporations, SMEs, and private users in all key markets across Europe, the Middle East and Africa. ICS’s offerings range from the smallest handhelds, through personal computing systems, to fully integrated enterprise-class infrastructure solutions. ICS complements this offering with the expertise of leading technology, software and service partners.
acquisition for us, building on our vision of a full-service, customer-driven organisation. The extended service portfolio enables us to compete more effectively against the other major players in this sector. The Services Division will operate as a separate subsidiary.”

On December 20, 2007, Delta and ICS signed a contract for ICS to buy IPS.

3. THE IT CHALLENGE

Confidential negotiations between Delta and ICS had begun in early October 2007, covering strategic aspects, including customer transfer, market expansion and business portfolio development. The Project Manager of the carve-out project was not part of the negotiation. Indeed, nobody from the carve-out team, let alone any member of IT-cut, had attended any of those negotiations. As a consequence, the contract covered very few aspects of the IT handover.

The overall carve-out project structure consisted of 10 work streams, of which IT was one. For example, the Human Resources work stream was responsible for all personnel issues, including negotiations about which employees would move to ICS. The legal work stream was responsible for conforming to the relevant regulations in the affected countries. Another work stream was responsible for the transfer of real estate. With many unique and unfamiliar tasks, the project team also hired a large number of consultants to support the work streams.

The IT work stream’s tasks were divided into IT infrastructure and IT applications. IT infrastructure covered all network issues, including rules for ICS to access the DIC’s IT systems during the transition phase. Telephone, fax, email, file services, print, desktop environments, back office systems and security were also part of the IT infrastructure domain. In addition, Internet and intranet systems had to move to ICS. Finally, all services that supported the data centers were migrated.

Within Delta, some local IT applications were under the control of regional business units, while core systems were centralized. Central applications were mandated for all countries, while local applications were approved only for a business unit from a specific country. The IT system for corporate group control, reporting and financials was an example of a central application. While these centrally controlled applications were immediately salient and visible to IT-cut, local applications were not well documented and were hard to track.

Because of the difficulty in identifying the relevant local applications, ensuring their timely migration to ICS was a major challenge. Each IT application had to be classified as a Delta-wide, DIC-wide or IPS-specific application. A pre-requisite for the classification project was an agreed process for the assessment of each application.

For example, consider the spare parts ordering system and the maintenance management system. While the former was used for the order and delivery of spare parts, it also triggered the billing process. The latter scheduled maintenance calls. Both systems were IPS specific. One advantage of IPS applications was their independence from IT Operations Services and IT Solution Services, the two other divisions in DIC. The datasets for these systems belonged specifically to IPS and there was no need to identify, separate and carve-out the relevant data. In contrast, the separation of shared DIC or Delta-wide applications was challenging.
Table 1: Six types of IT applications at Delta

The scheduling and billing systems for mobile technicians was an example of a DIC-wide application. This system was used by all DIC mobile technicians regardless of whether they worked for IPS or another SBU in DIC. In this case, IPS-relevant data had to be identified and carved out. Unlike the simple, direct transfer of an IPS-wide application to ICS, the transfer of a DIC-wide application was complex. The first decision to be made was whether to transfer the applications and data to ICS, or to separate the data and integrate it into an existing, similar system at ICS. A secondary decision was whether the historical data was important to ICS or whether ICS could be given access to the data by Delta via a Transition Service Agreement (TSA) for a specified period.

Delta-wide applications included, for example, mailing, groupware systems, and business warehouse and business intelligence solutions. Mailing and groupware systems were based on Microsoft technology, while financial and controlling systems, and the business warehouse and intelligence solutions were based on SAP technology. Mailing data for the affected IPS employees could have been carved out easily because each employee had an encapsulated mailbox. However, any employee in IPS was able to store documents and post comments to the Delta groupware platform. Therefore, it was difficult to identify and separate all relevant IPS employee documents and comments in order to both transfer them to ICS’s groupware system and delete them from Delta’ groupware system.

This was also the case for other Delta systems, including business warehouse and business intelligence solutions. Delta had invested heavily to harmonize the reporting structure across all business areas and regional business units. To do that, Delta had developed a standardized architecture for the business warehouse landscape in all regional business units. With integrated templates for the regional business units, Delta was able to consolidate quickly and easily across its warehouses around the world. Not surprisingly, these templates were not useful to ICS because it had implemented a different reporting system. The contract partners had to develop solutions to break such Delta-wide applications apart and to transfer them to ICS’s existing IT landscape.

Fortunately, both Delta and ICS used Microsoft technology for mailing, office and collaboration, and SAP for their financial and administrative backbones. This meant that there was a relatively high degree of compatibility between IPS and ICS systems. Unfortunately, that was the exception, and major differences between the IT systems of IPS and ICS were found everywhere. The problems were always in the detail. For example, there were differences in the desktop standard client. Although both partners used Microsoft Windows, the installed desktop standard client of IPS could not be used in the ICS environment. As a consequence, all desktop PCs transferred from IPS to ICS had to have the ICS desktop standard client installed.

IT-cut had to decide how to handle each application separately, with particular attention to the shared DIC-wide and Delta-wide applications.
Marcel Dupuis, leader of IT-cut, identified three critical IT challenges for IT-cut.

“The first is that some IT systems at IPS are hosted on the Delta IT platform and are highly integrated with other Delta IT systems. So, we have to work out how to identify the relevant IT systems and how to cut them out of our overall IT platform. But remember, while some IT systems are IPS-specific and are managed by IPS, they may also be used by other divisions.”

The second was:

“The complexity of the carve-out project itself. IT is only one work stream, although an important one, of this carve-out project. As one of the first steps, we will define the interfaces between our IT work stream and the other work streams (including Human Resources, Customer Contracts and Legal Issues) to ensure a good information flow.”

The third was the global coverage:

“I also anticipate changes in the number of countries to be affected by the carve-out, due to legal or strategic issues. IT is not organized centrally. Regional departments have certain authorities and power over IT decisions. This existence of local infrastructure and applications significantly increases the complexity of the IT carve-out project”

4. CARVE-OUT PROCESS

Figure 1 documents the steps to be followed in a typical IT carve-out project. The overall process is divided into four phases. These are separated by specific milestones. The first phase, “Pre-Signing”, clarifies and agrees all aspects concerning the contract. This includes the dates for the Closing milestone, the Transition phase and the Cutting milestone. In addition, the contract partners agree upon the distribution of costs for the overall carve-out project. This phase ends with the milestone “Signing” of the contract between the buyer and seller, finalizing the legal framework for the carve-out.

In France, without specific approval from the anti-trust agency, the seller’s project integration teams are not allowed to talk to the buyer about the carve-out. In this case, that approval was not expected to be given before the end of January 2008. Therefore, IT-cut had to plan and begin the project without communicating with ICS.

As part of the contract, it was agreed that ICS would have access to the IPS’s IT systems for six more months after Signing. Delta/DIC would provide the IT support and access for ICS during that period.

The second phase, “Pre-Closing”, establishes a working project management structure and creates the project management office. The IT carve-out team began its work stream with the analysis of networks and IT basic services. The relevant IT applications were identified, evaluated and prioritized for the IT carve-out.

The second phase ends with the “Closing” milestone. By then, all relevant IT systems had to be separated from DIC’s IT systems and exclusively under the control of ICS, guaranteeing its independence from DIC. This included separating the IT infrastructure, including email accounts, networks and telecommunication services.
The third phase, “Transition”, commences the day after Closing. On “Day One”, ICS required access to historical financial data and to have the capacity to manage its business within its own IT systems. Recognizing time constraints, and the time needed to carve out and integrate all relevant IT systems into the ICS’s IT infrastructure, DIC created a copy of the relevant IT systems, deleting all the data that was not included in the contract. During the remainder of the transition phase, DIC gave ICS access to DIC’s IT systems. This required the development of strict rules covering the access to the DIC’s network by ICS. This is terminated at the “Cutting” milestone. After the “Cutting” milestone, – fourth phase named “Post-cutting” – while ICS still has access to DIC’s network, ICS had established its own network and the network connection at DIC was physically decommissioned.

5. GETTING GOING

With no experience of an IT carve-out, the members of IT-cut have many questions for Marcel Dupuis. One of the team members asks: “Why do we have to do the IT carve-out in two steps? Why must we first duplicate all systems to fulfill the contract in time and then carve out all affected systems of our IT landscape afterwards in order to transfer them to ICS? Can’t we carve out the relevant IT systems until Closing?” Marcel stops him there and replies: “No, the task is too complex. At this time, we have no idea how many local IT systems are running in the affected countries. And the regional companies in those countries have strong power bases at Delta Head Office. It will not be easy to get all the relevant information quickly. We have to ensure that the regional companies are working with and not against us.”

Paul Alexander, a specialist in IT contracting, asks: “And how can we warrant this for ICS?” “With the help of Transition Services Agreements (TSA),” replies Alain. “TSAs govern access by business partners to our network. Therefore, we need strict rules to manage the carve-out firewall.”

“But, we have corporate guidelines for approving IT systems. According to those guidelines, we should be able to make a good guess about the affected IT systems in each country,” comments Carol Williams, who is specialist from organizational affairs. “You are right,” says Alain. “But guidelines are guidelines and not laws. Unfortunately, we are not able to monitor and control all regional business units and their IT portfolio. A certain freedom of choice is necessary to run effective regional business units. In addition, the guidelines are frequently only applied to Delta-wide global applications, for example, the e-mail and groupware standards. In the best case, a regional business unit has its own documentation system, but many of the regional business units have no documentation. Therefore, the identification of all affected local applications is one of our biggest challenges. We need to appoint a dedicated person to coordinate across the regional business units. This will synchronize our carve-out activities and approaches.”
“When we have identified all affected IT applications, would it be possible for us to simply separate those systems out of the existing IPS IT landscape?” asked Paul. “That is difficult, because the IT systems are tightly integrated with and connected to each other. But we could try this approach for IPS-wide applications,” answers Alain.

Carol offers another solution: “Could we copy and duplicate all affected systems and give them to ICS?” All the IT-cut members pause for a moment and think over this suggestion. Finally, Ernesto, a specialist on IT infrastructure, answers: “That approach sounds interesting. Unfortunately, it would be extremely expensive for us. And, in the end, it is likely that ICS would not want to have all of our IT systems, but only want to integrate the data into their existing systems.” Alain comments: “Additionally, almost every IT system is connected to another one. This means, that data is distributed all over the IT landscape and interlinked. In the worst case, we would have to duplicate our whole IT landscape, which is not realistic“.

“Ok, but then we could go another way. We could build up all relevant IT systems for IPS from scratch and only give over the datasets,” Paul suggests. Ernesto counters: “Then, what would we do with the existing contracts and depreciation schedules? What will happen to the old IPS systems that we do not need anymore?”

A lot of work is waiting for Alain Dupuis and his team in the next few months. It is clear, that this is not a simple task. The IT carve-out will end on April 1, which is only three and a half months away. The members of IT-cut are beginning to realize that an IT carve-out is more than just carve out IT. It is time to start.

6. **CASE QUESTIONS AND ISSUES**

1. Position the IT workstream within the overall project organization of the carve-out. What are the critical roles and competences required within the IT workstream?
2. Identify the main tasks for the IT carve-out and combine them within a project plan. Use milestones to structure the plan. Begin with the basic carve-out process model and extend as needed.
3. How can IT-cut establish transparency for the existing infrastructure and applications? Develop a data collection template.
4. How should IT-cut prioritize applications/systems for transfer? Develop a prioritization logic based on your data collection template developed in response to issue 3.
5. Develop a basic reporting tool and a management dashboard for the work stream project managers.
# A CULTURAL ANALYSIS OF BUSINESS PROCESS MANAGEMENT GOVERNANCE IN INDIAN ORGANISATIONS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0379.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Process Management, Governance, Culture fit / differences / heritage/ intelligence / issues / theory / values, Offshoring / Outsourcing</td>
</tr>
</tbody>
</table>
A CULTURAL ANALYSIS OF BUSINESS PROCESS
MANAGEMENT GOVERNANCE IN INDIAN ORGANISATIONS

Jayaganesh, Malini, Department of Information Systems, University of Melbourne, Victoria
3010, Australia, m.jayaganesh@pgrad.unimelb.edu.au
Shanks, Graeme, Department of Information Systems, University of Melbourne, Victoria
3010, Australia, gshanks@unimelb.edu.au

Abstract

Business process management (BPM) is a key issue for organisations, particularly in a global
business environment. In this paper we synthesise a framework for BPM governance and then report
two case studies that explore the influence of national culture on BPM governance in India. One case
study involves a global outsourcing services company with Indian origins and the other an Indian
manufacturing company that has recently established an overseas presence. The two case studies
provide a deep understanding of how culture influences BPM governance differently within each
organization and how BPM governance practices can be established to mitigate any negative
influences of national culture in a global context.

Keywords: Business Process Management, BPM Governance, Culture, India.

Acknowledgement

This study is supported by a grant from the Australian Research Council with SAP Research Australia
as the industry partner and is part of a project that compares business process management in
Australia and India.
1 INTRODUCTION

Business Process Management (BPM) is a priority for organisations around the globe as it enables sustained competitive advantage (Hung 2006). Recent studies indicate a growing awareness of and interest in process improvement and management among organisations (Harmon and Wolf 2008, Gartner 2007). It is felt that this trend together with the enabling role of information technology in BPM renders this phenomenon of particular interest to researchers of the Information Systems (IS) discipline. BPM is defined as a “holistic organisational management practice, which requires top management understanding and involvement, process-aware information systems, well-defined accountability and a culture receptive to business processes. It is based on a process architecture, which captures the interrelationships between the key business processes and the enabling support processes and their alignment with the strategies, goals and policies of an organisation” (Rosemann and de Bruin 2005). Business Process Re-engineering (BPR) can be considered as one of the foundations of BPM and in the context of process improvement, the primary difference between BPM and BPR is that the former emphasises ongoing improvement while the latter has become synonymous with radical one-off change (Armistead and Machin 1997).

Governance is a critical success factor for any BPM initiative (de Bruin 2007) as it is concerned with delineating accountability within an organisation (Weill and Ross 2004), managing the level and distribution of risk (Zingales 1997) and providing direction for the achievement of competitive advantage (Patel 2002). For purposes of this study, BPM governance is defined as the establishment of process-related responsibility and accountability mechanisms for the purpose of encouraging desirable behaviour in BPM. The majority of existing studies tend to be fragmented with focus on specific mechanisms of BPM governance such as process ownership (Booz Allen Hamilton 2003), process modelling (Bandara et al 2006) and performance measurement (Aberdeen 2006). There are only a limited number of studies (e.g. de Bruin 2007) which explore the interactions between these various mechanisms as a whole. Furthermore, contextual factors such as national culture which could potentially exercise an influence on BPM practice have also been relatively unexplored. Reports on the influence of culture on BPR (Agarwal and Haleem 2003, Martinsons and Hempel 1998) and on corporate governance (Buck and Shahrim 2005, Haniffa and Cooke 2002, Turnbull 1997) lead us to postulate that culture could have an impact on BPM governance.

India has emerged as a major player in the global economy in recent years due to its rapidly increasing share of the IT offshoring market. Liberalisation of its fiscal and trade policies in recent times has also led to increased presence of multinational organisations in the country and a furore of activity in many of its export sectors. Investment in IT solutions to manage and support business activity is also on the rise (Tarafdar and Vaidya 2006). In the context of implementation of one such IT solution in Indian organisations, viz. Enterprise Resource Planning (ERP) systems, process-related issues have been identified (Tarafdar and Roy 2003). Therefore it is felt that India provides an ideal setting for the study of national culture influences on BPM.

This paper explores the influence of national culture on BPM governance in India. The study builds on existing theory in both culture and BPM and includes in-depth case studies of BPM in two organisations in India. The remainder of this paper is organised as follows. Following the introductory section in which the motivation for the study has been outlined, an extensive review of literature concerning national culture and BPM governance is presented. Then the case study research method is explained. The fourth section of the paper provides a description of the two case study organisations and an overview of their BPM practice. This is followed by an in-depth analysis of BPM practices in the case-studies through the lens of national culture. The discussion section then presents the main findings and implications for research and practice, and the paper is concluded with limitations of the study and directions for future work.
2 LITERATURE REVIEW

This section presents the theoretical underpinnings of the study. The concepts of BPM governance and culture are defined and operationalised.

2.1 BPM Governance

The term governance is synonymous with the exercise of authority, direction and control (Zingales 1997) and is intended to ensure an organisation’s efficiency and effectiveness (Zingales 1997, Simonsson et al 2008). BPM governance has been identified as one of the factors for measuring BPM maturity in organisations and has been defined as the establishment of relevant and transparent accountability, decision making and reward processes to guide actions (Rosemann et al 2007). While this definition emphasises the exercise of authority and control, it does not adequately explain the intent or direction of these activities. A widely-employed definition of I.T. governance is that it specifies the decision rights and accountability framework to encourage desirable behaviour in using I.T. (Weill and Ross 2004). A common feature of this perspective with that of Korac-Kakabadse and Kakabadse (2001) is the emphasis on the focus or objective of the governance activity. Therefore, for purposes of this study, BPM governance is defined as the establishment of process-related responsibility and accountability mechanisms with the objective of encouraging desirable behaviour in BPM.

The governance component of the BPM maturity model includes five capability areas which include process management decision making, process roles and responsibilities, process metrics and performance linkage, process management standards and process management controls (Rosemann et al 2007). However it is felt that some of these capability areas could be considered collectively as there they are closely related and there is some overlap in definition. Therefore, for purposes of this study, a high-level BPM governance framework consisting of four distinct components has been synthesised. This framework differs from the governance component of the BPM maturity model in several ways. Firstly, process roles and responsibilities is perceived to include roles and responsibilities in process management decision-making as well and therefore a component titled “process-related responsibilities and accountability” which incorporates both aspects has been used. Similarly process metrics, performance linkage and process management controls have been collectively considered as BPM Quality Assurance. Furthermore, while the the BPM Maturity model includes process architecture as a capability area within strategic alignment, it does not include considerations to process definition as part of BPM governance. However, it is felt that the clear and consistent definitions of process and prioritisation of these in terms of criticality to operations should be considered as part of BPM governance as it is intrinsically linked to definition of process roles, standards and quality assurance. Therefore process definition has been included as a component of BPM governance in this framework. The components of the BPM governance framework as used in this study are described as follows:

- **Process-related Responsibilities and Accountability** (Ewusi-Mensah 1997, Fitzgerald and Carroll 2005, Kirchner 2005, Rosemann et al 2007, Weill and Ross 2004) – Process owners and committees for key processes with duties and responsibilities need to be clearly specified. This includes specification of authority for providing inputs to and making process-related decisions and the specification of precise reporting structures. It also includes the rules for enforcement of accountability, definition of the corresponding penalty and reward systems as well as provision of guidelines for escalation of issues.
- **Process Definition** (Rosemann et al 2007) – Processes need to be prioritised in terms of criticality to operations. At the very least highly critical processes need to clearly defined, documented and made accessible for use.

Proceedings ECIS 2009
• **Process Standards** (Fitzgerald and Carroll 2005, Rosemann *et al* 2007) – Guidelines for conducting all BPM activity such as process modelling, communication, issue resolution and monitoring need to be established.

• **BPM Quality Assurance** (Broadbent 2002, Fitzgerald and Carroll 2005, Rosemann *et al* 2007) – Periodic reviews to manage the quality and currency of processes as well as the efficacy of the organisation’s BPM practice need to be established. Metrics, which are in keeping with the organisation’s strategic goals, must be established for each process. These may be either quantitative or qualitative in nature but need be defined in such a manner as to enable ease of measurement. Performance must be monitored on a regular basis to ensure alignment with the overall BPM and organisational goals.

2.2 Culture

Culture is defined as the collective programming of the mind that separates the members of one group or category of people from another (Hofstede 1980). At the core of culture are values which are defined as broad tendencies to prefer certain states of affairs over others (Hofstede 1980). These tendencies are manifested in rituals, heroes and symbols which are collectively represented by practices or behaviours. Although these are visible to outsiders, their cultural meaning is not necessarily so and is determined by the way in which the practices are interpreted by other members of the same group. While behaviours can be modified with the introduction of incentives, values are relatively more stable over time. There are various levels of collective mental programming such as national, regional/ethnic, gender, generation, social class, industry/professional, organisational and departmental. This focus of this study on BPM governance in Indian organisations is at the national level and henceforth the term “culture” refers to national culture unless otherwise stated. The study of national cultural differences and resultant implications for management has been dominated by the characterisation of culture along a variety of preordained attitudinal dimensions or predispositions to action. For the purposes of this study, the three most influential theories of national culture are synthesised and the resultant five dimensions deemed to be most appropriate for the study of BPM governance are presented in Table 1. These include the works of Hall (1976), Hofstede (1980) and the GLOBE project (House *et al*. 2004).

<table>
<thead>
<tr>
<th>Dimension of Culture</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individualism - Collectivism</td>
<td>Hofstede (1980), (House <em>et al</em>. 2004),</td>
</tr>
<tr>
<td>Performance Orientation</td>
<td>(House <em>et al</em>. 2004)</td>
</tr>
<tr>
<td>High Context – Low Context</td>
<td>(Hall 1976)</td>
</tr>
</tbody>
</table>

*Table 1 Dimensions of Culture*

Hofstede (1980) administered a questionnaire to and analysed the responses of 72,215 employees of IBM across 40 countries. From this analysis he developed four dimensions of culture viz. power distance, individualism, masculinity and uncertainty avoidance. Later, a fifth dimension, long term orientation, was added. Each observed country was scored relatively on an index ranging from 0 to 100 for each dimension. This work has dominated cultural studies in the information systems discipline (Myers and Tan 2002).

The Global Leadership and Organisational Behaviour Effectiveness (GLOBE) research project (House *et al*. 2004) focuses on the relationships between national/societal culture, organisational culture and leadership. It involved 150 co-investigators who collected data from approximately 9000 managers in 500 different organizations in 3 different industry sectors in 61 countries. Nine dimensions have been identified and responses to questions based on these dimensions were sought from employees in three industry sectors across sixty two nations. The dimensions identified by the GLOBE project include...
uncertainty avoidance, power distance, societal collectivism, in-group collectivism, gender egalitarianism, assertiveness, future orientation, performance orientation, humane orientation.

Hall (1976) differentiated between cultures on the basis of context for communication. His contention is that the elements which combine to produce a given meaning – events and context – are in different proportions depending on culture. Therefore he compared cultures on a scale from high to low context. High context cultures rely on extensive social networks in the everyday lives and therefore do not require as much explicit background information for communication as low context cultures.

The five dimensions of culture which have been deemed as appropriate for use as a lens to examine BPM governance are explained as follows:

- **Power Distance** is the extent to which members of institutions and organizations within a country expect and accept that power is distributed unequally.
- **Individualism** refers to loose ties between individuals as opposed to strong ties with a given social network.
- **Uncertainty Avoidance** is the extent to which the members of a culture perceive ambiguous or unknown situations as threats as opposed to opportunities.
- **Performance Orientation** refers to the extent to which a society rewards innovation, quality and performance improvement.
- **Context** refers to the degree of explicit background information that is required for effective communication.

3 RESEARCH METHOD

The primary objective of this exploratory study is to understand how and why culture might influence BPM governance in Indian organisations. Therefore case study was selected for conducting the research as it been identified as an appropriate method for conducting “how and why” enquiries into real world phenomena (Yin 2003). As a first step, an extensive literature review was conducted to derive comprehensive definitions for both culture and BPM governance. Following this an interview protocol was developed to enable a systematic enquiry. The protocol consisted of both closed and open-ended questions – the former to ensure that responses are appropriately equivalent for purposes of cross-case comparison and the latter to provoke thought and dialogue in keeping with the exploratory nature of this study.

Two case studies were undertaken, one in a global outsourcing services company with Indian origins (Case Study A) and the other an Indian manufacturing company that has recently established an overseas presence (Case study B). The case study sites were selected on the basis of matched dimensions (including turnover, number of employees, position in industry etc.) which enable cross-case analysis. The unit of analysis in each case study was the BPM initiative including people, activities and documentation. Data collection included seven interviews at the first case study site and nine at the second. Interviewees included representatives from senior and middle management as well as staff at the operational level for purposes of completeness in understanding the BPM governance practice at each organisation as well as triangulation. The duration of each interview was between sixty and ninety minutes. This was further supplemented by a greater number of shorter interviews to follow-up on particular aspects of earlier interviews as well as observation of activity over several days. Extensive notes were kept of each interaction and interviews were audio-recorded wherever permission was granted. Appropriate documents, if any, were also scrutinised.

Following the data gathering phase, the audio recordings and notes were transcribed. Initial analysis of the text was done using constructs from the BPM framework described in section 2.1. The findings are presented in section 4.2 and 4.4. Further analysis within each of these was carried out through coding based on the dimensions of culture identified in section 2.2 and is presented in section 5.
4 CASE STUDY DESCRIPTIONS

In this section a description of the two case-study organisations is presented together with an explanation of the motivations for their BPM initiatives as well as a description of their BPM governance using the framework described in section 2.1.

4.1 Case Study A

This organization provides software and business process outsourcing services. Established in the early 1980s, it trades as a public limited company and is placed in a leadership position within its industry. From its very inception the organization has focused on developing its primary client base outside India and over the years offices have been established in several countries for purposes of front-end operations such as marketing and client liaison. As a result the organization has been a key player in the phenomenal growth of the I.T. offshoring market in India. Apart from its meteoric growth, the organization is also renowned for its relentless emphasis on quality. It has adopted a number of popular methodologies for this purpose such as Six Sigma and Balanced Scorecard and has been benchmarked to a number of global quality standards such as Carnegie Mellon University Software Engineering Institute’s Capability Maturity Model, People Capability Maturity Model, Malcolm Baldridge framework, ISO9001 and EFQM to name a few. Apart from ensuring the enforcement of highest quality standards from an internal management perspective, these initiatives also serve as a powerful signal of the organisation’s competency to the external market.

The organization is structured into business units which are defined along the parameters of domain specific knowledge and skill development. The organization has a matrix structure as there is movement of personnel, information and knowledge between the business units. There is also considerable exchange of personnel, information and knowledge across the company’s global operations. In the early 2000s the lack of integration between the systems being used was identified as the primary cause for the absence of insight into initiatives across the organization, particularly in the deployment of resources that enabled/supported them. Various tactical and operational bottlenecks were also ascribed to the same cause. The senior management of the organization recognized the limitations of the existing way of working and its repercussions on the business and an enterprise-wide BPM project for integrating the rich knowledge base of the organization was initiated. A multi-disciplinary team dedicated to the capture of process-related information and its management was established. A process repository was created and made available to employees via the intranet portal.

4.2 BPM Governance at Case Study A

Process-related Responsibilities and Accountability - As soon as senior management took the decision to implement BPM as part of the overall business strategy, a dedicated BPM team was established as the preliminary step. The team consists of 6 members with clearly specified areas of responsibility which included the initial setting up of the BPM project and its ongoing management. The first task for this team was the appointment of a process owner for every critical process. The process owner was responsible for defining the process, monitoring its ongoing performance and ensuring its alignment with strategy. The main challenge was to identify appropriate process owners. In many cases, this was straight-forward and was guided by the operational role of the person concerned. However for processes which span across business units, there were no volunteers and consequently the BPM team exercised their authority to nominate a person for each role as they deemed fit. Each process owner was then advised of the scope and nature of their responsibility including their role in the decision-making process. The reporting structures within each process area and issue resolution hierarchy was clearly outlined and communicated to all the concerned persons. Extensive documentation relating to process ownership definition and organisational reporting structures was prepared by the BPM team.
Process Definition - Each process in the organisation was documented and stored in a repository. The process documentation (PD) contains comprehensive information about the process including the description and sequence of tasks, the personnel involved and the data requirements. It also contains process maps that were drawn up jointly by the process owner and BPM team. In some situations there was a need to work backwards from the operating IT systems/solutions to identify the underlying business processes. Any change to a process is documented by the process owner on an ongoing basis. All of this information is available to employees from a BPM system via the organisation’s intranet portal. Access is automatically restricted to the area of business for each employee to prevent misuse of information. Every employee is expected to use the BPM system to record everyday activity. The system has built in validation checks and authorisation requests are automatically raised. The industry in which this organisation operates exhibits high attrition rates and therefore, having detailed documentation of all processes ensures that vital knowledge is not lost with departing personnel and new employees can settle into their roles quickly.

Process Standards - Standards were established for all necessary aspects of BPM activity. This included standards for process modelling and definition to ensure common understanding and co-ordination of BPM activities throughout the organisation. All employees received the necessary training to ensure that the standards are understood and enforced. All information related to standards, including any changes, is made available to employees via the online BPM system.

BPM Quality Assurance – The establishment of standards paved the way for quality assurance activity. Appropriate metrics were defined for each process as well as standard reward and penalty structures for process-related performance. Regular review cycles were established to ensure efficiency and effectiveness of the organisation’s activities. Furthermore the BPM activity itself comes under periodic review to ensure continuing efficacy of the practice and to identify ongoing opportunities for improvement. Following the review, adherence to desirable behaviour is rewarded while non-compliance is penalised.

Case Study B

This organization operates in the manufacturing industry and is a market leader in a diverse range of consumer durables. It commenced manufacturing operations in the 1955 and prides itself on a range of innovative products with a number of patent and design registrations. The organisation has 7 manufacturing units suited for scalable operations. These are supported by 20 branch offices and an extensive distribution network to reach 400,000 retail outlets across the country. The company also operates 25 exclusive branded retail outlets. The company has achieved its leadership position in the domestic market primarily due to its ability to provide high quality products at affordable prices. All products meet global standards such as ISO9001 for both design and quality. The management believes that the company owes its survival to a fundamentally progressive philosophy which, while emphasising the organisation’s essentially Indian identity, is also receptive to flexibility and change in the way of working to keep pace with changing times.

In the early 2000s the organisation underwent a shift in focus. Having established itself as the significant player in the domestic market, the company started to explore export opportunities and after due consideration a subsidiary was established in the USA. At the same time the hitherto underlying value discipline of product innovation was perceived to be insufficient for the strategic plans outlined for the future. An emphasis on operational excellence was introduced to complement brand and design capabilities. An ERP system was identified as the appropriate I.T. solution to support operational efficiency. The introduction of the ERP system resulted in the erosion of functional silos and the organisation was restructured as a process-centric enterprise.

4.3  BPM Governance at Case Study B

Process-related Responsibilities and Accountability – Process-related decisions in this organisation tend to focus on IT rather than on business. Process-related roles were not explicitly established. This
is because, unlike at Case Study A, an explicit BPM strategy was never clearly established at Case Study B by senior management. Process management was assumed to be a part of the ERP system implementation project and fell under the overall responsibility of the IT department. Consequently, persons with operational responsibility for the various functions assumed charge of the related processes. This has resulted in considerable overlap in the case of cross-functional processes resulting in collective rather than individual responsibility and accountability. Wherever consensus is not achieved, familiarity with the ERP technology tends to dictate which person has the last word in decision-making. Reporting and issue resolution structures tend to follow a similar pattern.

Process Definition – The average duration of employment of middle managers at this organisation is twelve years and attrition does not appear to be common enough for concern. During much of the past decade the organisation experienced phenomenal growth in the domestic market and for a period of three years in the late 1990s and early 2000s, an average rate of growth of 30% was achieved. This has been attributed firstly, to the expertise of the various functional managers and secondly, to the flexibility and responsiveness of the organisation’s business processes. While these characteristics have served the organisation well during the growth phase, it has also encouraged a tendency to rely on individual expertise rather than on well-defined processes. This is also accompanied by a tendency to abstain from locking-in process definitions for fear of making them rigid and less-responsive to change. The quality assurance (QA) process and the payroll process were the only two areas within the organisation that had any kind of process-related documentation. The QA process documents were developed as part of the regulatory requirements for certification of product and design quality. The payroll process documents were prepared by the payroll manager on the basis of his own initiative rather than a directive from the organisation and rarely used to guide day-to-day operations.

Process Standards – There is also a marked absence of explicitly defined standards. In terms of process documentation, those relating to the QA process are comprehensive and meet regulatory requirements. Graphical representations of the design and manufacturing process are detailed. However, the lack of clearly established standards for documentation is evident in the payroll process documents. The process descriptions are textual rather than graphical and minimalistic in nature. Many of the terms used are ambiguous and comprehensible only to the person who created them. However, the investigation revealed that absence of explicit standards and documentation in the organisation has not hampered internal operations and dealings with local clients. Instead these interactions are guided by consciously established social relationships and networks.

BPM Quality Assurance – A structured quality assurance activity is not discernable and metrics to judge process performance are not explicitly defined. In the case of the payroll process documentation, which lists 8 core processes, the performance measure for every one has been expressed as “timely action”. There is no accompanying explanation as to what constitutes timely action. However this does not mean that activities are not monitored. Once again, the organisation relies on the expertise of its managers to track process performance and to identify shortcomings. Reviews are conducted on an ad hoc basis and in an informal manner. Furthermore, a fierce loyalty to the organisation and its leadership is evident and verbal recognition of achievement is considered sufficient reward for adherence to desirable behaviour. Non-compliance is not punished outright but gently discouraged.

5 CASE STUDY ANALYSIS

The BPM governance practices of the two organisations which were described in the previous section are now examined using the dimensions of culture as a lens.

Power Distance is described as the extent to which members of institutions and organizations within a country expect and accept that power is distributed unequally. India exhibits a tendency for high power distance and the associated behaviours include emotional relationships between subordinates and superiors and the perception of the boss as a benevolent autocrat or parent (Hofstede 1980, House et al 2004). At Case Study A, processes are adhered to mainly because of the strict enforcement of
rules and the accompanying rewards and penalties. This reflects a somewhat impersonal style of management. However at Case Study B, a strong emotional relation with one’s superiors and the respect for authority and expertise is reflected in the accompanying desire to stand favourably in their opinion which in turn motivates compliance to organisational processes. Rather than enforcing penalties for acts of non-compliance, offenders are reprimanded by their supervisors and gently shepparded back into line much like in a parent-child relationship. It was also perceived that communication between superiors and sub-ordinates sometimes suffers due to the high power distance. At Case Study A, supervisors can look up the BPM system to obtain an accurate picture of the state of operations. However, at Case Study B, supervisors request face-to-face meetings with subordinates in order to obtain an understanding of the current situation. This communication mechanism is highly dependant upon the established relationship between the people involved and effective communication is at times overshadowed by respectful reticence and fear of loss of face.

**Individualism** refers to the strength of ties within a given social network. India exhibits a tendency towards low individualism or, in other words, collectivism (Hofstede 1980, House et al 2004). At the workplace the primary visible behaviour of collectivist societies is that personal relationships prevail over task rather than the other way around and occupational mobility is lower. At Case Study A, while relationships within teams are nurtured, such structures are short-lived as employees are frequently moved across projects. Furthermore an employee’s relationship with the organisation as a whole is largely impersonal and, as mentioned earlier in this paper, attrition rates are high. At Case Study B, on the other hand, employees identify themselves strongly with the organisation and perceive the fate of the organisation as being parallel to their own. This acts as a compelling motivator for performance. Stronger bonds are forged with members of the same functional unit and this provides the foundation for all work-related activity. Most employees tend to remain in the organisation once employed and the length of period of relationship further enhances its stickiness.

**Uncertainty Avoidance** refers to the extent to which the members of a culture perceive ambiguous or unknown situations as threats as opposed to opportunities. Members of low uncertainty avoidance cultures such as India exhibit a tendency for tolerance of ambiguity and chaos and are generally better at invention than at implementation (Hofstede 1980, House et al 2004). At the first case study site, a marked intolerance for ambiguity is noticed. Roles, standards, metrics etc. are clearly specified and systematically enforced. However at Case Study B, ambiguity prevails over nearly every aspect of BPM and management of related activities is through individual expertise on an ad hoc basis.

**Performance Orientation** refers to the extent to which a society rewards innovation, quality and performance improvement (House et al 2004). India exhibits a tendency towards low performance orientation and associated behaviours include high value placed on relationships and loyalty and less emphasis on training and development, results and rewards. At first glance, these characteristics appear to correspond with the features of Case Study B and not with those of Case Study A. However, a closer study reveals that results are valued at Case Study B and this is, in fact, a reason for emphasising outcomes rather than processes. Furthermore rewards are also valued, albeit the nature of the reward is not promotion or a financial incentive (as at Case Study A) but recognition and a sense of achievement. Therefore it is concluded that the influence of performance orientation on BPM governance has not been established in either case study.

**Context** refers to the degree of explicit background information that is required for effective communication (Hall 1976). India exhibits high context tendencies which mean that members of this society rely on subtly implicit and highly contextual information to enable communication. To a person who is unfamiliar with the associated contextual details, the communication may seem ambiguous and ineffective. At Case Study A, all process-related communication is explicit and for the most part, occurs via the online BPM system. At Case Study B context rich verbal communication occurs between the peer-groups of employees constantly. Therefore the absence of explicit mechanisms does not appear to affect the quality of communication between employees.
6 DISCUSSION

A number of broad trends have emerged through the analysis of the two case studies. Firstly, this study provides empirical evidence for and insights into the influence of culture on BPM governance. The analysis of Case Study B indicates that the Indian cultural characteristics of high power distance, collectivism, low uncertainty avoidance and high context have resulted in highly informal BPM governance practices which might appear ambiguous and chaotic to persons who are unfamiliar with the cultural context. The findings related to power distance and collectivism corroborate those of a cross-cultural study on the influence of culture on BPR projects in India and the USA (Agarwal and Haleem 2004). This study established that individualism and low power distance, which are characteristics of American culture, facilitate the implementation of BPR projects whereas collectivism and high power distance exert a negative influence in the Indian context. However, the same study also established that low uncertainty avoidance facilitates BPR. This is in contrast to the empirical evidence of this study which establishes that, in fact, low uncertainty avoidance results in a more casual approach to BPM governance with less emphasis on establishment of process roles, definition and standards.

On the other hand, the analysis of Case Study A demonstrates that an organisation which operates in the same national context as CASE Study A and therefore susceptible to the same cultural influences can successfully use a highly formal and rigorous BPM governance practice to mitigate the negative impact of such influences. Furthermore, the experience of this company suggests that there is a relationship between an organisation’s motivation for BPM and its susceptibility to cultural influence. According to the definition employed in this study, the primary objective of BPM governance is to encourage desirable behaviour. At Case Study A, we find that from its very inception the emphasis has been on establishing and maintaining a client base outside India. In other words, the desirable behaviour can be described as adopting work practices that enable and support business dealings with international clients. The BPM governance practice has been consciously designed to support interactions with non-Indian persons. Therefore an examination of the various aspects of BPM governance revealed an absence of Indian cultural influence. This has proved to be a success factor for the organisation in establishing itself in the I.T. offshoring market. In contrast, until recently, the desirable behaviour at Case Study B has been the adoption of work practices that enable and support business dealings with Indian clients. The rather ambiguous and informal BPM governance that has been practiced hitherto appears to have supported this objective. This corroborates findings from a study which demonstrates that exposure to global markets has driven software firms in India to have adopt western-style corporate governance practices (Khanna and Palepu 2004).

The contributions of this study for research are two-fold. Firstly, the development of a BPM governance framework based on literature review has proven useful for the study of such practices in organisations. As mentioned earlier in the paper, existing research on BPM governance has tended to be fragmented and focussed on specific mechanisms such as process ownership (Booz Allen Hamilton 2003), process modelling (Bandara et al 2006) and performance measurement (Aberdeen 2006). The BPM governance framework employed in this study provides a holistic perspective of the topic. It has also facilitated systematic and thorough data collection and analysis. Secondly, the majority of cultural studies in IS rely on the cultural dimensions of Hofstede (1980) alone. By including dimensions from other two other sources, this study has demonstrated a richer means of analysing the impact of culture on organisational practice. Although the impact of performance orientation on BPM governance has not been established in this study, the inclusion of this dimension from the GLOBE project (House et al 2004) has enabled the investigation into the case study organisations’ approach to outcomes and rewards. The inclusion of communication context (Hall 1976) as a cultural dimension has provided a clear explanation for Case Study A’s enthusiastic adoption of an impersonal technology-based BPM system for process-related communication while Case Study B employs communication methods which appear to be highly context-dependant and therefore chaotic and ambiguous to outsiders.
For practitioners, the emphasis of the study on the relationship between motivation for BPM and susceptibility to cultural influence is enlightening. The focus on offshoring business at Case Study A has resulted in a systematic and well-designed BPM governance practice while the largely informal practice at Case Study B has proven to be sufficient for its focus on the domestic market. However, as mentioned earlier in the paper, the desirable behaviour for the latter organisation is undergoing change as it is now actively seeking to expand its export capabilities. As the example of Case Study A has demonstrated it is possible that interactions with international clients would require more explicit communication and more formalised process management thus rendering the existing informal ways of working a bottleneck rather than an enabler. Therefore one may conclude that Indian organisations seeking to establish international business should consider adopting formal BPM governance practice in a similar fashion to Case Study A in order to reduce/eliminate the negative influence of Indian culture on their operations and to facilitate open and unambiguous communication between all parties.

7 CONCLUSION

This paper makes several key contributions to the understanding of cultural influence on BPM governance. It builds on existing research on BPM governance and culture and also serves as an exemplar for conducting culture-focused case studies on BPM governance. The findings of the study will be of use to organisations in India which are en route to expanding export capabilities by alerting them to the importance of designing their BPM governance to mitigate the possible negative impact of cultural influences.

A limitation of this study is that it focuses on a single cultural context. It is felt that the influence of culture on BPM governance can be better understood by comparing practices in two or more culture contexts. This limitation of the current study helps to set the agenda for future research. An in-depth study of several organisations in India and Australia is currently underway and it is anticipated that the cross-national comparisons will explain the impact of culture on BPM governance in a comprehensive manner. Furthermore the differences between a domestic-market focussed organisation and a global one also motivates the inclusion of both locally owned/operated organisations as well as subsidiaries of multinational companies in both countries.

References


DEVELOPING A FRAMEWORK FOR IT GOVERNANCE IN THE POST-MERGER INTEGRATION PHASE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0632.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Mergers and Acquisitions, IT governance, Design Science, Organisational Change</td>
</tr>
</tbody>
</table>
DEVELOPING A FRAMEWORK FOR IT GOVERNANCE IN THE POST-MERGER INTEGRATION PHASE

Becker, Jörg, University of Münster, European Research Center for Information Systems, Leonardo-Campus 3, 48149 Münster, Germany, becker@ercis.uni-muenster.de

Pöppelbuß, Jens, University of Münster, European Research Center for Information Systems, Leonardo-Campus 3, 48149 Münster, Germany, jens.poepelbuss@ercis.uni-muenster.de

Stolze, Carl, University of Münster, European Research Center for Information Systems, Leonardo-Campus 3, 48149 Münster, Germany, carl.stolze@ercis.uni-muenster.de

Asgarian, Cyrus, Deloitte Consulting GmbH, CIO Advisory Services, Franklinstraße 46-48, 60486 Frankfurt am Main, Germany, casgarian@deloitte.de

Abstract

As today’s Information Technology (IT) penetrates the business of almost any company, the IT integration of two companies is regarded as one of the most challenging tasks in mergers and acquisitions (M&A). This integration mainly takes place during the post-merger integration (PMI) phase of an M&A endeavor. In this phase, well-defined IT governance is critical to success since it specifies the decision rights and an accountability framework to encourage a desirable course of the IT integration. It harmonizes IT-related integration issues with desired behaviors and business objectives.

In this paper, we provide a framework for IT Governance in the PMI phase (ITGoPMI). The ITGoPMI framework can serve as a blueprint for the specification of organizational units and their decision fields when the IT functions of two companies become one. We are developing the framework according to the principles of design science research in an iterative manner. Preliminary versions of the framework were evaluated through expert interviews. They approved that the here presented version of our ITGoPMI framework already has the potential to serve as a beneficial guideline for post-merger IT integration in practice. Nevertheless, we will continue our research process with further evaluation and improvement of the framework.

Keywords: Mergers and Acquisitions, IT Governance, Design Science, Organizational Change
1 MOTIVATION

Over the last few decades, mergers and acquisitions (M&A) have established themselves as important options for companies striving for growth or increasing profitability (Alaranta & Henningsson 2008). As today’s Information Technology (IT) penetrates the core business of almost any company, the IT integration is regarded as one of the most challenging tasks in these M&A endeavors (Schewe & Lohre & Böhmer 2007, Carr 2003). Surveys underline that this integration aspect is often seen as a cause for problems or unfulfilled expectations (PwC 2004, ComputerWeekly 2001, Deloitte Consulting 2008).

Consequently, IT governance can be considered as critical to the combination of IT functions that mainly takes place during the post-merger integration (PMI) phase of an M&A endeavor. Not only in the PMI context, IT governance specifies the decision rights and an accountability framework to encourage a desirable course of the IT integration (Weill & Ross 2004). It harmonizes IT-related integration issues with desired behaviors and business objectives. Appropriate structures and procedures need to be established in order to achieve success of the overall M&A endeavor.

Although the importance of this topic is striking, the issue of IT integration in M&A has only received little consideration in academic literature (Henningsson & Svensson & Vallén 2007, Mehta & Hirschheim 2004, Alaranta & Henningsson 2008). From our point of view, especially the role of IT governance in the PMI phase of M&A has been regarded insufficiently.

Addressing this insufficiency, the objective of this paper is to provide a framework for defining IT governance for the course of the PMI phase. It can serve as a guideline to cope with the issues linked to the post-merger IT integration by suggesting IT-related organizational units and adequate allocation of decision rights.

Having provided the motivation for the development of a framework, the next section will present foundations about IT governance and M&A that lay out the basis for our work. Thereafter, we will argue that the development of such a framework can be approached according to design science which aims at designing innovative artifacts that are built and evaluated in an iterative manner (Hevner & March & Park & Ram 2004). We will present our research method by illustrating the iteration steps that have been completed so far. Subsequently, we will explain the current version of our Framework for IT governance in the PMI phase (ITGoPMI) which is the result of the third iteration of our research process. Finally, we will briefly discuss implications for further research.

2 THEORETICAL BACKGROUND

2.1 IT Governance

Beyond doubt, the term IT governance refers to the steering of the use of information technology of all kinds within a company. However, as several authors have pointed out, there is no consistent understanding of the term neither in research nor in practice (Johannsen & Goeken 2007, Rüter & Göldner & Schröder 2006, Simonsson & Johnson 2006). For instance, Weill and Ross (2004) define IT governance as “specifying the decision rights and accountability framework to encourage desirable behavior in using IT”. Desirable is the behavior that is aligned with the general targets, strategy, values and norms of a company. Therefore, the design of IT governance is always specific to a given company. The scope of IT governance are not single decisions themselves but the determination which decisions need to be made, who can contribute to the decision-making processes and who is eventually eligible to make the decision. In this sense, every company has an IT governance, but only an explicitly designed one is able to align IT effectively and efficiently to the goals of the company. In order to achieve this, IT governance needs to be in line and linked with the general corporate governance. Information and information technology are seen next to employees, finance, physical
Not only in the previous definition, but also according to Luftman (2003), Van Grembergen (2002) and the IT Governance Institute (2003) IT governance refers to the leadership of the use of IT. Furthermore, there is consensus that the ultimate goal of IT governance is IT business alignment (Gentle 2004). Steering and monitoring of structures form the key tasks of IT governance.

From our point of view, a strict differentiation between two possible conceptions of IT governance is necessary. On the one hand, IT governance can be regarded as the definition of structures and procedures of decision-making, and the allocation of decision rights to people and units. On the other hand, some definitions also include the actual decision-making. Adopting the definition of IT governance by Ross and Weill in this paper, we agree with the first conception.

2.2 Post-Merger Integration of IT

Mergers and acquisitions can be classified entirely different based on either legal terms or practical power differences between the two fusing companies (Alaranta & Henningsson 2007, Henningsson & Carlsson 2007). A discussion of these differences is beyond the scope of this paper. In fact, both mergers and acquisitions typically require the IT of two companies to be combined to a certain degree.

Integrating former independent IT functions is one of the major challenges of an M&A process, even seen more critical than resistance of employees in general. Nevertheless, this challenge is not always considered as it should be in respect to its importance (McKiernan & Merali 1995, Wirtz & Wecker 2006, Deloitte Consulting 2008). For the series of actions, which need to be proceeded to form the combined IT, the term *IT integration process* has been established. However, it is not necessarily an integration in the truest sense of the words. As a matter of fact, the IT integration process covers all decision areas of IT governance, not just the technical core parts but also its soft aspects like organization and strategy. Therefore, we assume that the integration of IT also covers the integration of information systems (IS). In this sense, we see IT as the broader term that encompasses not only technical facilities but people, organizational structures, procedures, tasks and processes as well.

The IT integration process can be regarded as a transformation caused by the overall M&A process (Farhoomand 2005, Schewe & Lohre & Genius 2007) and takes place during the *post-merger integration* (PMI) phase (see Figure 1). Together with the two phases ‘preparation’ and ‘transaction’ PMI forms the M&A process as a whole (Meckl 2006, Henningsson & Carlsson 2007).

![Diagram of IT integration process](image)

**Figure 1. Transformation of IT during the post-merger integration**

According to Wijnhoven & Spil & Stegwee & Fa (2006), there are four prototypic strategies which refer to different levels of integration. For instance, a superior player usually assimilates the inferior one, whereas there are also mergers of coequal organizations that have to proceed differently. The
The most rigorous strategy is to replace any existing IT (see outer left box in Figure 2). The second strategy is the adoption of existing IT instead of creating an entirely new one. Standardization as the third option can be described as taking parts out of both present ones and recombine them into the new IT. This is often done in a “best-of-breed” fashion. Synchronization finally describes a strategy in which the new IT is basically a consolidated interface for the still existing pre-merger ones which can stay unchanged. Independent from the chosen strategy, the PMI of IT requires in any case adequately defined structures for decision-making and monitoring as well as accountability to enable and encourage an efficient management of the desired changes (Stolzenberg & Heberle 2006, Ringlstetter & Kaiser & Schuster 2006).

Figure 2. Transformation of IT during the post-merger integration (see Wijnhoven et al. 2006)

Because of the explicit or implicit existence of IT governance in the two merging firms, IT governance during the PMI cannot be entirely detached from the existing IT governance. In fact, their structures need to be investigated during the initial stage of the PMI. Existing boards, committees and levels of management need to be carefully integrated (Wijnhoven et al. 2006). Although the situation is different for every single M&A process, it is not unlikely that the existing IT governance incorporates elements of standardized frameworks, especially ITIL or CoBIT (Fink & Ploder 2008, IT Governance Institute 2008).

2.3 Facilitating the Definition of IT Governance in the PMI Phase by a Framework

A framework is a conceptual model that describes the structures of the organization in scope on an abstract level by a selected organizational paradigm (Meise 2001). The purpose of a framework is to provide an overview of relevant elements and their interdependencies. As the arrangement of elements in a framework can be chosen freely, the visualization of the overall has priority over the detailed depiction of single elements. Every element itself can be represented by a model of its own again and thereby a hierarchy of models within a framework can be formed (vom Brocke 2003, Meise 2001).

The development of a framework often aims at creating a reference model that includes abstract solutions and generic schemes for the design of organizations and information systems (Meise 2001). Those can be based on either theoretical considerations or practical experiences. The resulting framework then describes either an ideal and theoretically sound status (“best practice”) or a systematic compressed image of the perceived reality of the modeler (“common practice”). Both types hold recommendations to help the practitioners and scientists alike in their work (vom Brocke 2003).

Whilst the main characteristic of a framework is the provision of an overview over a certain field, we especially aim at creating a model with a normative character. It is expected to give a guideline for the IT integration within the M&A process. As Weill and Ross (2004) emphasized, IT governance means to define which decisions about IT should be made by whom and in what manner in order to provide accountability and control. Finally, as discussed in the previous section, existing IT governance approaches need to be respected when designing a new one for the post-merger company. A general vision of the change process needs to be determined which eventually has to be put into practice. Altogether, these assumptions form the basic requirements for our framework for IT governance during the post-merger integration phase (see Table 1).
Framework

- Provide an overview about organizational units and decision fields
- Provide reference for IT integration endeavors

IT governance

- Define decisions to be made
- Clarify who is allowed to make these decisions
- Specify the way decision-making is done as well as accountability and control

Post-merger integration phase

- Respect that elements of existing IT governance are the starting basis for the new one
- Help to decide about the vision for the IT transformation
- Support steering and monitoring of the IT integration

Table 1. Requirements for a framework for the IT governance during the PMI phase

3 RESEARCH METHOD

3.1 Developing the Framework According to Design Science Research

We regard the development of a framework as being subject to design science research which is motivated by the desire to improve the environment by means of innovative artifacts in terms of constructs, models, methods, and instantiations (Hevner et al. 2004). However, contrary to design science research as depicted by Hevner et al. (2004) that focuses the IT artifact per se, we also include the non-technological context, i.e. people and organization. We aim at developing a framework that provides practical knowledge that is applicable for the management of IT during M&A endeavors (Carlsson & Henningsson & Hratinski & Keller 2008). This practical knowledge can also be referred to as being a design theory (Carlsson et al. 2008, Gregor 2006, Hevner et al. 2004).

![Figure 3. Framework Development Process](image)

According to Hevner et al. (2004) and March & Smith (1995), design science research proceeds as a cycle of the two phases ‘build’ and ‘evaluate’. Similarly, literature on design theory development gives ‘propose/refine design theory’ and ‘test design theory’ as two key research activities (Carlsson et al. 2008). Thus, first within an iteration loop, artifacts or design theories are built addressing problems in an innovative way. Then, they are evaluated or tested against the utility they provide in solving these problems, and they are examined whether requirements and restrictions are met (Hevner et al. 2004). If an artifact fails this evaluation, the heuristic process is resumed beginning again with the next build phase aiming at refinement of the previous result. Our research process, which is currently in its third iteration, reflects these basic principles (see figure 3) and is described in detail in the following sections.

Proceedings ECIS 2009
3.2 First Draft of the Framework

Initially, we designed a first and – in retrospect – comparatively simple and incomplete draft of the ITGoPMI framework. We selected the shape of a house according to Meise (2001) as the basic design of our framework. This reference design allocates the core processes in the center of the framework which are surrounded by steering and support processes.

The first draft was evaluated by discussing it with an academic expert in the field of PMI and innovation management. He approved the systematic design of the framework. However, he questioned the explicit consideration of an integration of IT cultures due to concerns (especially German) workers’ representatives possibly have. We answered to these concerns by abandoning the accentuation on IT cultures from our framework. Henceforth, we regarded this rather as one of the human resources topics in M&A (Picot 2005). Our overall objective was sharpened, namely to design a framework that provides a guideline for the specification of decision rights as well as an accountability framework to enforce desirable outcomes during the PMI phase.

3.3 Extending the Knowledge Base

Design science research refers to a broad knowledge base of theories, models or methods that provide the basis for the design process (Hevner 2007). Similarly, design theories are expected to ground in previous research which has to be reviewed systematically (Carlsson et al. 2008). Due to the identified shortcomings of our first draft, we recognized the need to extend our knowledge base with further existing artifacts, experiences, and expertise that define the state-of-the art.

First, we analyzed frameworks like COBIT and ITIL since organizations frequently adopt those when explicitly designing their IT governance (IT Governance Institute 2008, Rüter & Göldner & Schröder 2006, Fink & Ploder 2008). For instance, decisions regarding the ‘Service Strategy’ aspect of ITIL V3 (Cartlidge & Hanna & Rudd & Macfarlane & Windebank & Rance 2007) largely cover the IT decision fields presented by Weill and Ross (2004).

Second, we conducted a literature review in order to identify possibly existing theories and artifacts regarding IT governance in the PMI phase. To analyze the state-of-the-art of research in this field, we examined high quality publications listed as ‘A’ or ‘B’ in the well-established rankings VHB JOURQUAL 2 (Hennig-Thurau & Schrader 2008) and WKWI (2008). We searched the online databases of these journals and conferences for the terms „post-merger“, „PMI“, „information technology“, „governance“ and „IT integration“. To guarantee up-to-date results, we limited our literature review to papers published from July 2005 to June 2008.

This literature review reinforced our proposition that currently there is no comprehensive framework for our scope, since almost any of the 25 publications we identified referred to a quite specific sub-aspect. Nevertheless, most of these aspects were relevant to at least some degree and thus had to be considered in the further development of our framework. For instance, Vaast and Levina (2006) advise against over-codification of procedures, and as already presented in section 2.2, Wijnhoven et al. (2006) provide an excellent overview about possible integration strategies.

3.4 Enhancing the Framework

We refined the ITGoPMI framework driven by the identified shortcomings of the first draft and our extended knowledge base. We kept the basic design of the framework, but added certain organizational units that address key subprograms and decision fields relevant during the PMI phase. Referring to the structuring found in the frameworks ITIL (Cartlidge et al. 2007) and COBIT (IT Governance Institute 2007) as well as in further publications (Rüter et al. 2006, Weill & Ross 2004, Deloitte Netherlands 2008), we derived seven sub-aspects of governing IT in the PMI phase. These comprise: IT Strategy, IT Governance, IT Organization, IT Processes and Services, Information
Systems, IT Personnel and Change Management. Each subprogram is managed by an integration team that is lead by a team leader who is responsible for his/her particular aspect.

The subprograms are steered by superior organizational units: the IT Integration Strategy Committee which holds the overall responsibility for the IT integration in the M&A process and makes fundamental decisions, and the IT Integration Steering Committee that is responsible for the coordination with the departments of the demand organization. The IT Integration Office serves as an interface between steering level and the integration teams. Located below the subprograms, the units IT Governance Assurance Team and IT Integration Assessment Team monitor the compliance with the defined IT governance and the progress of the integration endeavor. Additionally, the Demand Organization (in the organizational view) with its levels represents those parts of both companies outside the actual IT organization that demand and consume IT services of any kind (see Figure 4). It should be noted that these organizational units are also affected by the M&A transformation and therefore may be constantly changing.

Figure 4. Enhanced Version of the Framework

3.5 Expert Interviews

Proceeding with the next evaluation phase in our design process, the ITGoPMI framework was evaluated through five (semi-structured) expert interviews. Four of these experts were professional IT consultants; the fifth person was the academic expert we had already interviewed in the previous evaluation phase. All interviewees were experienced with M&A endeavors (see Table 2). We especially asked IT consultants for their view because they are typically involved in M&A endeavors in order to facilitate the IT integration with their expert knowledge and tools (e.g. a normative framework). Furthermore, we argue that consultants, as being not directly affected by the outcomes of the M&A process itself, are less biased in their view on how the IT governance should be shaped during the PMI.
Table 2. Sample of Experts

<table>
<thead>
<tr>
<th>Interviewee</th>
<th>Consultant A</th>
<th>Consultant B</th>
<th>Consultant C</th>
<th>Consultant D</th>
<th>Academic Expert</th>
</tr>
</thead>
<tbody>
<tr>
<td>Involvement in M&amp;A Processes</td>
<td>Yes, but more demergers</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Number</td>
<td>3-5</td>
<td>More than 5</td>
<td>1</td>
<td>1</td>
<td>More than 5</td>
</tr>
<tr>
<td>Role</td>
<td>External consultant</td>
<td>External consultant and affected employee</td>
<td>External consultant</td>
<td>External consultant</td>
<td>External assessor</td>
</tr>
</tbody>
</table>

The practical value of the framework as well as theoretical premises and the relevance of included elements were discussed. Parts of our interview outline were inspired by a past PMI survey conducted by Bauch (2004). Again, the interviewees appreciated the structuring of the framework and they honored its level of detail (the experts received a description of the framework that was of considerably more detail than the previous section 3.4). However, the heterogeneity of their perceptions – even of basic terms like ‘IT governance’, ‘change management’ and ‘framework’ – became apparent through the interviews. This also left some degree of uncertainty of how a framework for IT governance in the PMI phase precisely should look like. Nevertheless, some common criticisms led the further development of the framework. For instance, a distinct IT Integration Strategy Committee was regarded as superfluous which we consequently dropped in the next build phase. The result of this phase represents the current version of the framework being subject to the next chapter.

4 FRAMEWORK FOR IT GOVERNANCE IN THE PMI PHASE

As a result of the suggestions made by the experts during the interviews we developed the ITGoPMI framework further to its subsequently presented form. The house as the basic design of the framework has proven to be sufficient to structure the matter in a consistent and comprehensive way and thus has been kept throughout the development process. The structure is still steering as the roof, subprograms to structure the integration work itself as the body and monitoring as the basement of the framework. However, the organizational units in each part have been revised to reflect insights gained through the feedback and suggestions of the interviewed experts.

The steering level at the roof of the ITGoPMI framework has been consolidated to two units: The IT Post-Merger Integration Office and the IT Post-Merger Integration Steering Board. The first one acts as a project management office by coordinating the work of all integration teams and keeping them on track from the operational perspective. The IT Post-Merger Integration Steering Board is responsible for making strategic and far reaching decisions about the transformation of IT in general. It decides especially about the desired outcome of the overall transformation process and which strategic integration approach should be followed (see section 2.2). As this has profound impact on the way IT can support the demand organization, the demand-sided senior management level needs to be represented in the IT Post-Merger Integration Steering Board. As an M&A process is usually about more than IT, both organizational units in the framework’s roof (see Figure 5) have to keep close ties with the steering committee and project office of the entire post-merger integration process.

The steering level is supported by IT Post-Merger Integration Steering Committees for specific IT PMI decisions which require the involvement of the department level of the demand organization. For instance, these can refer to the consolidation of CRM systems, which the marketing departments heavily rely on. Nevertheless, the exact number and topics of those committees needs to be individually determined for any M&A process with its specific requirements and situation. Hereby also the likely transformation of the demand organization needs to be taken into account.

On the operational level, Business Liaison Managers form the interface between the subprogram integration teams and the departments of the demand organization. Hence, these positions possess a political dimension as especially cooperation and communication needs to be facilitated between the
different stakeholders. In this context, according to our interviewed experts, it is advisable to involve external professionals as neutral instances for this role. At least one Business Liaison Manager is needed for each subprogram. A larger number seems to be appropriate for situations with highly complex work in the subprograms or many departments involved.

**Figure 5. ITGoPMI Framework**

As shown in Figure 5, a dedicated Integration Team is responsible for each subprogram. Each team should comprise members from the two merging companies in order to avoid political disputes and facilitate knowledge sharing. All integration teams must report at least to the steering level. Out of the wide range of thinkable subprograms for the post-merger integration of IT itself, seven subprograms have been identified as being repeatedly used and/or suggested by the interviewed experts. These seven are the following: IT Strategy, IT Organization & IT Governance, IT Processes, IT Services, IT Applications, IT Systems and IT Facilities & Networks.

The subprogram **IT Strategy** aims at developing an integrated and sustainable IT strategy for the newly-merged company. Hence, the general parameters for the future developments of IT are defined in this subprogram. Moreover, this must also include the definition of an exact strategy for the transformation of IT during the PMI phase as this is highly related with the to-be future IT strategy as well as the general logic behind the M&A process.

To make the new IT strategy come alive after the merger, appropriate organizational structure and governance are needed. Due to the highly interrelated nature of IT organization and IT governance both are jointly worked on in the subprogram **IT Organization & IT Governance**. One of the key tasks in this subprogram is to determine which mechanisms and structures exist already in both companies and how they can be used or integrated into the structures during the PMI as well as afterwards.

Whilst the subprogram **IT Processes** defines – as the internal perspective of structuring – how IT operates, IT services describe what the outcomes of IT use are. These outcomes can be either delivered to the – from the point of view of the IT function – external demand organization or being used...
internally as inputs of another service. Therefore the subprogram IT Services determines what the IT is supposed to deliver to whom after the completion of the PMI.

The remaining three subprograms address the technical transformation itself during the PMI phase. The subprogram IT Applications is about the software layer that runs the business-relevant programs and their interfaces within and across the newly-formed company. The subprogram IT Systems deals with desktop hardware and system software that enables the applications to run. Last but not least, the subprogram IT Facilities & Networks is covering the topological aspect of IT by determining for example where data centers should be located and how the networks should be operated.

In the basement of the ITGoPMI framework, two teams are monitoring the IT integration. First, the IT Integration Assessment Team is concerned with tracking the progress of the transformation. Second, the IT Governance Assurance Team checks the compliance with the governance structures valid for the course of the IT integration. In order to cope with conflicting interests, both teams should consist not only of members of the involved firms but of external professionals as well. Especially from a corporate governance point of view, it is important that steering and monitoring do not influence each other. Consequently, it can make sense to involve external consultants in these two teams.

To breathe life into the organizational units, especially in steering and monitoring, we developed a set of typical processes they need to execute. Maybe the foremost obvious is that the integration office of the general M&A process needs to make sure that the IT Post-Merger Integration Steering Board is formed. Then, this board appoints the IT Post-Merger Integration Office, the integration teams at the core and all other needed organizational units. These are the most responsible actions as the choice of people with their talents, skills and motivations determines the performance that can be achieved for the whole transformation. Moreover, as the IT Post-Merger Integration Steering Board is responsible for the outcome, it needs to receive reports of the subprogram integration teams and the monitoring teams. These disclose the achieved progress of the integration and the compliance to the PMI-specific IT governance structures. Finally, the IT Post-Merger Integration Office manages the day-to-day coordination between the integration teams and drives their work efforts if needed.

Altogether these organizational units form a comprehensive structure to assign decision rights and define accountability during the PMI of IT. However, it merely represents a heuristic proposition and guideline. Adjustments may be necessary for individual M&A.

5 IMPLICATIONS FOR FURTHER RESEARCH

In this paper, we presented a framework for IT governance in the PMI phase which is the preliminary result of a still ongoing research process. Building on the well-established definition of IT governance according to Weill and Ross, the ITGoPMI framework provides a blueprint for the specification of decision rights and accountabilities. Precisely, it outlines organizational units and their decision fields for the course of the IT integration. It aims at providing a means to ensure desirable outcomes of IT use, even during the fairly instable time when two companies and their IT functions merge.

Nevertheless, there might be limits in applicability for M&A processes in which small and medium-sized enterprises join forces. We see the need for further applicability checks to explore which structural elements of our framework can be regarded as universal and which are only needed in large-scale M&A endeavors. There may even be cases in which no integration of IT takes place at all and consequently such a framework guiding the IT integration process is not needed either.

As we have seen, there are different strategies for the post-merger IT integration. Hence, there is no single best way of how to integrate two former independent IT functions. Further research should analyze if varying strategies also require specific adjustments to the framework.

We learnt that developing a comprehensive framework for the IT governance in the PMI phase is a challenging task. So far, our empirical research was limited to interviews with German experts. Even in this relatively small sample of interviewees there were divergent perceptions of basic terms like ‘IT
governance’ and ‘framework’. This opened up controversial discussions about what IT governance actually is, especially during the post-merger integration phase. In spite of the heterogeneity of their perceptions, we have been assured by all experts we interviewed that our ITGoPMI framework already has the potential to serve as a beneficial guideline for post-merger IT integration in practice. Obviously, this opinion needs to be validated by empirical research on a larger scale.

As the output of design science research must be returned to the environment for evaluation (Hevner 2007), we will continue our research process by interviewing IT managers with M&A experience as well as by trying to apply the framework in practical M&A situations. This will possibly lead to further improvements. In the long term, we will strive for establishing it as a reference model for IT governance in the PMI phase.

References


ARCHITECTURES OF PARTICIPATION AND EMERGENT FORMS OF INTER-ORGANISATIONAL ACTIVITY: A PRELIMINARY ANALYSIS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0579.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Open Innovation, Collaboration, interorganizational business processes, Architecture</td>
</tr>
</tbody>
</table>
ARCHITECTURES OF PARTICIPATION AND EMERGENT FORMS OF INTER-ORGANISATIONAL ACTIVITY: A PRELIMINARY ANALYSIS

Feller, Joseph, University College Cork, College Road, Cork, Ireland, jfeller@afis.ucc.ie
Finnegan, Patrick, University of New South Wales, Sydney, Australia, p.finnegan@unsw.edu.au
Nilsson, Olof, Mid Sweden University, 851 70 Sundsvall, Sweden, olle.nilsson@miun.se

Abstract

This research-in-progress paper explores ways in which inter-organisational networks can fulfil both collective (network) and self-interest (member) goals by implementing architectures of participation that govern and support participant interaction. We draw on recent studies to derive a conceptual framework consisting of technological, legal, economic and social mechanisms in inter-organisational networks, and using initial findings from three case studies of networks in Sweden and Denmark, we illustrate how these architectures operate. Our analysis shows the relative importance of these mechanisms in facilitating individual and collective value creation and we conclude by presenting our ongoing research plans.

Keywords: Inter-organisational networks, Collaboration, Open Innovation, Architectures of Participation
1 INTRODUCTION

Over recent decades organisations have been radically transformed from single location hierarchical structures with well-defined boundaries to more flexibly organised and dispersed configurations. Indeed, arguments for flatter structures (Levitt and Whisler 1958, Snow et al 1992) and flexible forms (Drucker 1988, Starkey et al 1991, Bahrami 1992, Benkler 2002, 2006) have mounted. Advanced information and communications technology as well as an increasingly competitive environment have been cited as the main contributors to this change (Miles and Snow 1986, Peters 1992). Collaboration frequently occurs on an inter-organisational basis, as partnerships with suppliers and customers are critical in many sectors (Cash 1985, Christianse et al 2004, Tapscott and Williams 2006). Consequently, the boundary between organisations and their partners is becoming even less distinct with interdependencies between them being more important (Gulati and Kletter 2005, Premkumar et al 2005).

Although much inter-organisational activity has been production based, the exchange of specialised knowledge and skills is increasingly important (Bartlett and Ghoshal 1989, Sonnenberg 1992, Chesborough 2003). The image that emerges from these developments is one of an extended or virtual organisation where external boundaries are obscured by inter-organisational dependencies characterised by substantial communication and collaboration. The sharing of information, ideas and expertise becomes central to the success of such extended enterprises (Bartlett and Ghoshal 1989), and there is a realisation that external forces, rather than internal organisational, technological or managerial variables, are the keys to explaining organisational success (Joynt 1991).

Despite the importance of inter-organisational co-operation in relation to servicing consumer needs for products and services (Subramani 2004, Okamura and Vonortas 2006), organisations have been slow to harness the same type of external cooperation in relation to innovation (Lane and Probert 2007). Nevertheless, innovation is the result of combining different knowledge sets (Nonaka et al 2003, Tidd et al 2005), and such knowledge is frequently to be found outside the organisation (Chesbrough 2003, De Wit et al 2007). However, with the exception of notable examples of collective invention (cf. Allen 1983, von Hippel 1987), organisations have been slow to engage in open innovation (cf. Chesbrough 2003). In addition to worries about the quality and suitability of external ideas, organisations have resisted co-operative approaches to innovation due to perceived competitive necessities and issues relating to organisational control (Chesbrough 2004).

This paper presents research-in-progress on inter-organisational networks that aim to facilitate process, product and service innovation rather than just support inter-organisational transactions. The study seeks to establish ways in which inter-organisational networks can fulfil both collective (network) and self-interest (member) goals by implementing architectures of participation that govern and support participant interaction using technological, legal, economic and social mechanisms. Drawing on initial findings from three networks in Sweden and Denmark (a Danish business association, a group of Swedish public authorities, and a Swedish network with private, public and academic partners), we illustrate how these architectures operate and conclude by presenting our ongoing research plans.

2 CONCEPTUAL GROUNDING

The term open innovation has been used to describe the shift from a paradigm in which firms exclusively sought to create, improve, and exploit products/services within the boundaries of the firm, to one which “places external ideas and external paths to market on the same level of importance as that reserved for internal ideas and paths to market” (Chesbrough 2006, p. 1). In practice, the open innovation concept can be implemented in many ways, e.g. creating new/improved products/services by building on external ideas and innovations, creating new revenue streams by allowing external parties to exploit internal ideas and innovations, and forming alliances with complementary partners to
support ongoing innovation/exploitation processes (Gassmann and Enkel 2004). However, new products are not the only tangible manifestation of open innovation as services and process transformation can also be developed in this manner (Morgan and Finnegan 2008). Such activities require participating firms – whether they act as consumers, producers or partners – to engage with external parties; they therefore depend on the establishment and management of effective inter-organisational networks (Vanhaeverbeke and Cloodt 2006).

Participants in inter-organisational networks believe that collaboration will result in adaptive efficiency; “the ability to change rapidly and at the same time provide customized services or products, and at low cost” (Alter and Hage 1993). An inter-organisational network is a social action system as it exhibits the fundamental principles of any organized form of collective behaviour. These include the aim to achieve both collective (network) and self-interest (member) goals, interdependent processes utilized by network members, and the ability of the cooperative entity to act as a unit with a separate identity from its individual members (Van de Ven 1976).

While these activities can be supported by traditional inter-organisational governance structures like hierarchies, markets and brokerages (Feller et al 2008a), many examples of open innovation instead leverage what Benkler (2002, 2006) has described as peer production: a model for organizing production that does not rely on markets, hierarchies, property and contracts. Specifically, the collaborative creation of software by development communities (open source software) has been used as a defining example of the peer production model (Benkler 2002, 2006), and the engagement of firms with such communities and the products they create has been identified as a key exemplar of open innovation (West and Gallagher 2006).

Any understanding of open innovation processes must, therefore, not only include inter-organisational interactions embedded in traditional governance structures, but also those embedded in peer production contexts. In such contexts, the tension between the collective goals of groups and the individual goals of profit-seeking participants can be problematic. In the current work, we address this tension through the concept of an architecture of participation, which has emerged in the literature with two distinct meanings.

Firstly, an architecture of participation can be understood simply as the “various technologies and activities designed to facilitate and promote participation, communication and the active co-construction of meanings and knowledge” (Attwell and Elferink 2007); in other words, they are collections of mechanisms that allow the members of a community or network to interact. Secondly, and more subtly, the label has been used to describe systems that help transform individual activities into communal resources. Bricklin (2001) observed that the technical characteristics of peer-to-peer music sharing systems like Napster could potentially transform the “tragedy” of the commons into the “cornucopia” of the commons, where adding value to the system is an automatic result of using the system. Building on Bricklin’s insight, O’Reilly (2005) has applied the concept to open source software development, where he argues that such architectures “may actually be more central to the success of open source than the more frequently cited appeal to volunteerism. The architecture of Linux, the Internet, and the World Wide Web are such that users pursuing their own ‘selfish’ interests build collective value as an automatic by-product” (O’Reilly 2005, p. 476). Within the open source context, many different types of architectures have been identified; including technological architectures (e.g. collaboration and communication platforms, software development kits and application programming interfaces, etc.), legal architectures (e.g. software licences), economic architectures (e.g. direct and indirect incentives and rewards for participation) and social architectures (e.g. shared cultural values, reputation building through participation, etc.)(Feller et al 2008b).

Thus, the current work seeks to identify the technological, legal, economic and social architectures of participation used within the innovation networks studied, and to understand the roles played by these architectures in both enabling interaction (a la Attwell and Elferink 2007) and transforming individual activities into communal goods (a la Bricklin 2001 and O’Reilly 2005), as illustrated in Figure 1.
3 RESEARCH CONTEXT AND METHOD

The objective of this study is to explore architectures of participation in inter-organisational networks. Given the exploratory nature of this research, and the need to obtain rich data in a complex inter-organisational context, a case study approach, with embedded units of analysis, was considered appropriate. ‘A case study examines a phenomenon in its natural setting, employing multiple data collection methods to gather information from a few entities. The boundaries of the phenomenon are not clearly evident at the outset of the research and no experimental control or manipulation is used’ (Benbasat et al. 1987). Cases are most appropriate when the objective involves studying contemporary events, without the need to control variables or subject behaviour (Yin 2003). Our method is consistent with the case study approach of Benbasat et al. (1987) and Yin (2003) in that we study the phenomenon in its natural setting, employing multiple data collection methods to gather information from a few entities, without employing experimental control or manipulation. We follow in the tradition of Eisenhardt (1989) and Madill et al. (2000) by seeking to reveal pre-existing, relatively stable and objectively extant phenomena and the relationships among them.

Data gathering activities to date have focused on three networks with different characteristics and aims. The first is a Danish business association of private companies active in the open source software market. The second is a group of Swedish public authorities that cooperate in procurement processes. The third is a network including partners from the Swedish public sector, private sector and academia that focus on co-creation of public e-services.

The researchers first conducted an archival search of public domain material on the networks and their participants, including web resources and articles in the public press. Based on this preliminary analysis a case study protocol (cf. Yin 2003) was prepared in order to ensure the consistency of data gathered. Interviews with key informants were conducted during 2008 (see Table 1). The interviews, which followed an interview guide (cf. Patton 1980), were of 30-60 minutes duration and conducted in Swedish and Danish both in person and by telephone.

The interviews were transcribed and translated (by one of the authors), and follow-ups were made by e-mail and telephone to clarify and refine issues that emerged during the transcription/translation process. The interview transcripts were supplemented with 15 official documents provided to the researchers by the interviewees. The documents included policy statements, bylaws and project reports.
Content analysis was carried out on both the interview and document data sets. A coding system was derived using the conceptual framework illustrated in Figure 1, and a two-phase coding process was employed (c.f. Miles and Huberman 1994). During the first-level coding phase, each segment of the interview/documentation data was summarized and labelled. This was followed by a pattern coding process in which the segments of data were organized, analyzed and synthesized within the themes/concepts embedded in the conceptual framework. While the emphasis of the first-level coding phase was on description, the pattern coding process focused on explanation.

### Table 1. Key Informants Interviewed

<table>
<thead>
<tr>
<th>Network</th>
<th>Interviewees</th>
</tr>
</thead>
</table>
| OSL                                               | 1. Chairman  
|                                                   | 2. Board member  
|                                                   | 3. Board member |
| Environmentally Procurement in County Västernorrland | 1. Project Leader |
| The Sundsvall Group                                 | 1. Researcher  
|                                                   | 2. CIO |

**Table 1. Key Informants Interviewed**

### 4 FINDINGS AND ANALYSIS

#### 4.1 Network, dynamics, composition and goals

The initial analysis of the three networks involved characterising them based on their dynamics, member composition, and goals (both individual and collective). Table 2 summarises these characterisations, which are discussed in detail below.

<table>
<thead>
<tr>
<th>Dynamic</th>
<th>OSL</th>
<th>Procurement</th>
<th>Sundsvall Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Association</td>
<td>Cooperation</td>
<td>Collaboration</td>
</tr>
<tr>
<td>Composition</td>
<td>Homogenous (private companies)</td>
<td>Homogenous (public authorities)</td>
<td>Heterogeneous (private/public/academia)</td>
</tr>
<tr>
<td>Collective Goal</td>
<td>Build collective brand and influence</td>
<td>Implement public policy and stimulate “green” manufacturing</td>
<td>Product development</td>
</tr>
<tr>
<td>Individual Goal</td>
<td>Develop individual business opportunities</td>
<td>Lower costs</td>
<td>Various</td>
</tr>
</tbody>
</table>

**Table 2. Network, dynamics, composition and goals**

The Danish Open Source Business Association [*Open Source Leverandører i Danmark*], or OSL, was founded in 2003 by five companies active in the open source solutions business sector. The members of OSL are a relatively homogeneous group of private sector organisations, all Danish-based companies that deliver solutions and services based on open standards or open source software. At the time of writing, OSL comprises 30 ‘full member’ companies and an additional 31 ‘supporting member’ companies. Most of the member companies are small (5-20 staff), but there are also a few multinationals with Danish headquarters (e.g. IBM, Sun and Oracle). The key dynamic evident in OSL is one of *association*, in which the network acts to provide a unified voice for its member companies. The member firms are motivated to participate in the network by a desire to develop individual business opportunities; however, the network acts as a lobby group for all of its members, seeking to enhance awareness of open source software in the Danish market and influence Danish IT policies. Specifically, the goal of OSL is “to promote a genuine market-based choice between different types of
software development and licensing in order to make the choice of open source software based on quality, price, usability and suitability ... The association will work to secure a genuine choice of IT-architecture in the public as well as the private sector” (OSL 2008). Thus, the network’s effectiveness as a brand building and policy influencing mechanism is enhanced by the active participation of individual companies, who in turn benefit from these activities.

‘Environmentally Procurement in County Västernorrland’ is an initiative, established in 2006, in which eight Swedish public authorities (the seven municipalities and the County Council in the county of Västernorrland) and the Association of Local Authorities in Västernorrland, have established a network for joint procurement activities. Other partners include the Swedish Road Administration (‘Region Mitt’), the Västernorrland Administrative Board, the Swedish Agency for Economic and Regional Growth and the Swedish Environmental Management Council. The network is thus a homogenous grouping of public authorities, in which the key dynamic is cooperation in the form of joint procurement activities. Individual authorities are motivated by a desire to lower costs by leveraging the negotiating position that comes from aggregating demand across the county. The network, in turn, seeks to implement public policy favouring “green” products (like low-energy lightbulbs and hybrid automobiles), which it can do through the joint purchasing power of its collective members.

‘The Sundsvall Group’ [no official name exists] is a network, established in 2005, to support collaborative software development activities amongst its members, who include Swedish public authorities (Municipality of Sundsvall), private firms (Logica), non-profit associations (Åkroken Science Park) and academic institutions (the CITIZYS Research Group at Mid Sweden University). The network was originally established with the aim of carrying out a single joint project (the development of a municipal e-service in the municipality of Sundsvall, ECHOES (Everyday Communication Home School)) but has evolved to support other projects. Because this collaborative network is heterogeneous in composition, the individual goals of the members vary considerably from member to member; e.g., the municipality delivers e-services at a lower cost, the academic partner enjoys access to research opportunities in an applied setting, the private company gets a new business opportunity, etc. Collectively, the network seeks to develop e-services that improve the quality of life for individual citizens through easier access to public information and services.

4.2 Architectures of participation

We found that the different kinds of networks have different sets of architecture of participation. We refer to the components that form the architecture of participation as the technical, legal, economic and social architectures.

<table>
<thead>
<tr>
<th></th>
<th>Association</th>
<th>Cooperation</th>
<th>Collaboration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technical architectures</strong></td>
<td>Tools for communication</td>
<td>Tools for communication</td>
<td>Tools for communication</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tools for coordination</td>
<td>Tools for coordination</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tools for co-development</td>
</tr>
<tr>
<td><strong>Legal architectures</strong></td>
<td>Identity</td>
<td>Identity</td>
<td>Project Identity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Process</td>
<td>Ownership</td>
</tr>
<tr>
<td><strong>Economic architectures</strong></td>
<td>Fees</td>
<td>Fees</td>
<td>Projects</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Projects</td>
<td></td>
</tr>
<tr>
<td><strong>Social architectures</strong></td>
<td>Shared views</td>
<td>Shared views</td>
<td>Shared views</td>
</tr>
<tr>
<td></td>
<td>Trust</td>
<td>Trust</td>
<td>Trust</td>
</tr>
<tr>
<td></td>
<td>Reputation</td>
<td>Reputation</td>
<td>Reputation</td>
</tr>
<tr>
<td></td>
<td>Networking</td>
<td>Networking</td>
<td>Networking</td>
</tr>
<tr>
<td></td>
<td>Collective sanctions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Architectures of participation
4.2.1 Technological architectures

All three networks were analysed in order to identify the technological architectures that enabled interaction between members and/or acted to transform individual activities into communal resources. Across all three networks, technologies such as email, websites and telecommunications were used to communicate within the network. In the procurement and software development networks, these same technologies were also used to explicitly coordinate activities as well as for communication. In the procurement network, such co-ordination activities were limited to specific procurement activities, while in the software development network longer term coordination took place. Finally, within the Sundsvall group, additional technologies supported the distributed collaborative development of software. These included version control tools, issue tracking systems, etc. Additionally, the software products created by the network were themselves architecturally designed to support future participation; i.e. the way in which ECHOES was developed meant that the code base could be largely reused in future projects. It is worth noting that with the exception of this final point regarding code reuse, the roles played by technological architectures in all networks appeared to be limited to enabling interaction, not transformation.

4.2.2 Legal architectures

Within OSL, where the primary collective goal was the creation of a joint voice for marketing and lobbying, the dominant legal architectures focused on the management of the association’s identity and of member behaviour through formal bylaws. For example, all potential members must demonstrate that their commercial activities are in-line with OSL’s intentions (i.e. the promotion of open source solutions). Furthermore, to become a full member, the applicant must be a company and be able to prove that during the last financial year they employed the equivalent of more that one full time employee. Smaller companies, and even individuals, are also entitled to membership but as supporting members without a vote in network level decisions.

Within the procurement (cooperative) network, legal architectures exist not only to manage network identity but also to ensure that the members are able to engage in procurement activities following agreed upon procedures. A complex system of agreements governs the joint procurement processes, including, national regulatory and legislative frameworks, local policy decisions made by public authorities, and purchase-specific agreements between the network members.

Finally, within the collaborative network, there is less evidence of legal architectures governing behaviour or identity. Instead, network identity is fluid (re-established with each new project), “rules” for behaviour are informal or absent, and the primary emphasis is on legal mechanisms to prevent conflict over the ownership of the collaboratively created software product (e.g. licensing decisions and sub-contracts with service providers, etc.).

In contrast with the technological architectures previously discussed, the legal architectures evident in the networks more directly support the transformation of individual efforts into communal resources. For example, in OSL the bylaws ensure that the membership act and speak with a cohesive voice; in the procurement group they are able to act as a single purchasing agent; and in the Sundsvall group there are clearly defined rights of collective and individual ownership.

4.2.3 Economic architectures

In OSL, annual fees provide a financial mechanism to support the day-to-day administrative activities of the network and also act as a signal of a firm’s commitment to the collective goals of the network. Annual fees are also used in the procurement network to offset administrative costs (primary the salary of the overall project leader). Additionally, financial agreements are made between a subset of the network membership for each specific purchasing event. Within the collaborative software development network, financial agreements are limited to specific aspects of projects (e.g. the
management of a specific project budget or sub-contracting of a particular activity). As with the technological architectures discussed previously, the economic architectures evident appeared to be focused more on enabling interaction than transformation.

### 4.2.4 Social architectures

Within OSL the formal bylaws governing member behaviour are complemented by several informal social mechanisms that both enable interaction and help ensure that the network as a whole benefits from individual activities. For example, the members of the association share a common goal (the expansion of the market for open source products and related services) and share the belief that they can all benefit from this expanded market without needing to directly compete with each other. There is also a shared ethos that the members should focus on winning customers from the proprietary software companies, not from each other. Although the association does not explicitly seek to facilitate cooperation between the member companies, many of the companies have formed relationships through the association, which can lead to exchanges of knowledge, contacts, etc. Interpersonal trust and knowledge about other actors were seen to play an important role in such interactions. Rumours about ‘bad’ jobs or ‘bad’ behaviour spread quickly, which effects decisions regarding potential cooperation or sourcing of competencies, but could also lead to collective sanctions e.g. exclusion.

Within the cooperative network the members also share a collective vision and a common goal of effectively implementing environmental procurement policies. Personal relationships and networking between the participants play an important role in building trust and sharing knowledge. Periodic meetings and workshops reinforce such relationships. As with the OSL, informants highlighted the importance of reputation in choosing procurement partners/leaders.

Finally, in the collaborative network, social architectures were seen to play an important role, particularly in the absence of written rules, bylaws, etc. Specifically, personal knowledge, trust, and a common view on the future of public e-services act as the uniting ‘glue’. There are no regular meetings; instead, new ideas and projects emerge from frequent contacts and discussions. Through the personal networks of the ‘core’ project members, new members are invited to participate in projects when there is a need for external competencies or expertise.

As with the legal architectures evident in the networks, the social architectures play a key role in the creation of communal resources. In all three networks, a shared worldview and an implicit acknowledgement of acceptable behaviour enable members to pursue individual goals while creating/preserving communal resources.

## 5 CONCLUSIONS

This paper has explored two connotations of *architectures of participation* in the association, cooperation and collaboration networks studied (see Figures 2, 3 and 4); as mechanisms for enabling interaction and as mechanisms for transforming individual action into collective resources. Both connotations are critical to the successful implementation of open innovation strategies.

First, open innovation activities are, by definition, extra-organisational and require mechanisms that facilitate interaction between participants. This study has given us insight into the wide variety of mechanisms employed by open innovation networks to meet this need; the use of technological, legal, economic and social architectures were all visible. The demonstrated use of a variety of mechanisms has implications both for practitioners and future researchers. For practitioners, this study suggests the need to move beyond a narrow technological viewpoint, and to consider all four types of architectures in managing network activities. For future researchers, including ourselves, there is a need to develop a better understanding of the interaction and interdependencies between the various mechanisms.

Second, as noted at the outset of the paper, issues of trust, control, and the potential “tragedy of the commons” can all play a role in dissuading organisations from participating in open innovation.
Therefore, mechanisms for transforming individual action into collective resources (i.e. enabling all participants to benefit from the collaboration) are critical. The study showed that although all four types of architectures have a function in enabling interaction, only the legal and social architectures have a strong role in transforming individual action into collective resources. It is noteworthy that informal/social mechanisms and formal/legal mechanisms both play a key role in the networks studied. For practitioners, this suggests the need to take into account both formal and informal mechanisms in making governance decisions and in dealing with issues of appropriation and sharing. For researchers, the findings signal the need to reconsider the formal/legal-centric view of inter-organisational networks dominant in extant literature, and contribute to the emerging characterisation of networks based on the interplay between formal and informal mechanisms (c.f. Feller et al 2008).

Figure 2. Architectures of Participation (Association)

![Figure 2. Architectures of Participation (Association)](image)

Figure 3. Architectures of Participation (Cooperation)

![Figure 3. Architectures of Participation (Cooperation)](image)
The analysis of technological, legal, economic and social architectures of participation described above provide us with an initial set of models for understanding the interplay between individual and collective value creation. By treating the three networks as abstract types (Association, Cooperation and Collaboration), we hope to stimulate related research focused on networks with similar goals/characteristics in other contexts. The work reported is research-in-progress, and our own plans for future research activity include (1) the identification of additional open innovation network types, (2) further data gathering in the existing networks (both through interviews with a wider range of stakeholders and the administration of a survey to network participants), and (3) focusing on additional embedded units of analysis (projects). In doing so, we aim to uncover a more exhaustive inventory of the various architectures of participation in use, and also develop a more sophisticated set of models for understanding how these architectures enable interaction and support transformation.

**Acknowledgements**

This work has been financially supported by the Irish Research Council for the Humanities and Social Sciences (IRCHSS) through the *Open Code, Content and Commerce Business Models* project (http://www.o3cbusinessmodels.org).

**References**


THE DEVELOPMENT OF AN EVALUATION MODEL OF E-COMMERCE WEBSITES FOR THE TAIWANESE AIRLINE INDUSTRY

<table>
<thead>
<tr>
<th>Journal</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ECIS2009-0058.R1</td>
</tr>
<tr>
<td>Submission Type</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword</td>
<td>Decision making, E-commerce (B2B / B2C / B2G / G2C), E-Marketing, Web Site Analysis</td>
</tr>
</tbody>
</table>
THE DEVELOPMENT OF AN EVALUATION MODEL OF E-COMMERCE WEBSITES FOR THE TAIWANESE AIRLINE INDUSTRY

Tsai, Wen-Hsien, National Central University, No.300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan, whtsai@mgt.ncu.edu.tw
Leu, Jun-Der, National Central University, No.300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan, leujunder@mgt.ncu.edu.tw
Chou, Wen-Chin, National Central University, No.300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan, 954401001@cc.ncu.edu.tw; Yu Da College of Business, No.168, Hsuehfu Rd., Tanwen Village, Chaochiao Township, Miaoli County 36143, Taiwan, wcchou@ydu.edu.tw

Abstract

The airline industry employs advanced e-commerce technologies to retain frequent customers and attract new passengers, but it is likely that not all airlines have clear knowledge about how many gaps should be filled between the status quo and an ideal website. This study proposes an integrated model for evaluating airlines’ websites in terms of the perspectives of “marketing mix 4Ps” and “website quality”. In order to verify the practicality and usefulness of this model, an empirical study of the Taiwanese airline industry is offered to illustrate the application of the proposed model. The model first applies the DEMATEL method to cope with the interdependencies between criteria. It then converts the criteria’s cause and effect relations into a visual structural map where the ANP method can help compute the weight of criteria. Finally, it uses the modified VIKOR method to rank e-commerce websites of the five Taiwanese airlines. Overall, the results show that the Taiwanese airlines still have a great deal of room to improve their websites. This proposed model not only provides helpful information for airlines to understand their websites’ quality level, but also contributes to industrial applications in terms of providing some worthwhile recommendations for building an ideal website.

Keywords: Decision Making, E-Commerce, E-Marketing, Web Site Analysis
1 INTRODUCTION

The Internet and information technology (IT) have become common practice of the airline industry and helped sharpen airlines’ competitive edge through operation efficiency improvement (Hanke & Teo 2003). One of the most valid solutions for enhancing business values and attracting more customers is to sell low-fair air travel tickets and facilitate boarding processes such as e-ticketing and online check-in through an airline’s own website (Wei & Ozok 2005). Currently, many airlines are utilizing dedicated websites to market their products to potential customers. Some airlines also offer discounts to customers who purchase their tickets online (Hanke & Teo 2003). With the ongoing trends of the Internet marketing, this raises the critical issue of how airlines can effectively measure their websites’ characteristics. Due to the multidimensional characteristics of website design in this issue, it must be appropriately addressed by multiple criteria decision-making (MCDM) methods.

Nowadays, there is no a standardized model for evaluating e-commerce websites of airlines, and existing ranking methods do not offer enough insights for airlines’ proprietors to determine whether their websites meet ideal levels in terms of the perspectives of “marketing mix 4Ps” and “website quality”. Therefore, an effective model is proposed that combines the Decision Making Trial and Evaluation Laboratory (DEMATEL) method, the Analytic Network Process (ANP) method, and the modified VIKOR (VlseKriterijumska Optimizacija i Kompromisno Resenje in Serbian, or Multi-criteria Optimization and Compromise Solution) method for an inter-website comparison of airlines. First, the DEMATEL method is applied to deal with the interdependence between evaluation criteria and also to convert the criteria’s cause and effect relations into a visual structural map. Next, the ANP method is employed to determine the relative importance of evaluation criteria. The ANP, a relatively new MCDM method, can systematically deal with network-like decision problems. Finally, the modified VIKOR method is used to evaluate and rank the websites of airlines. The VIKOR, a compromise ranking method, had been introduced by Opricovic (1998). It established an aggregating function based on the measure of closeness to the ideal solution. This ranking index of the VIKOR method is an aggregation of all criteria, the relative importance of criteria, and a balance between total and individual satisfaction (Opricovic & Tzeng 2004). Therefore, this integrated model is useful for an airline company to identify the performance level of its website in the airline industry, and then to take necessary actions to improve shortcoming for enhancing competitive advantages.

The purpose of this paper is to build an evaluation model of e-commerce websites for airlines. By suggesting comprehensive measurements based on the perspectives of “e-marketing” and “e-quality”, the proposed model can provide helpful information for airlines’ proprietors to understand the unimproved gaps of their websites. Moreover, this study contributes to industrial applications in terms of providing some worthwhile recommendations for building an ideal website. In order to verify the practicality and usefulness of this model, an empirical study of the Taiwanese airline industry is offered to illustrate the application of the proposed model.

2 THE CONCEPTS OF E-MARKETING AND E-QUALITY

2.1 The perspective of e-marketing

The Internet is especially valuable for coping with the intangible nature of service and transforming “marketing mix” variables to exploit the informational and transactional potential of the Internet (Baloglu & Pekcan 2006). The “marketing mix” concept was introduced in the 1950s, and the mix of different means of competitions was soon regrouped to 4Ps (i.e. product, price, place, and promotion) by McCarthy (Grönroos 1997). These 4Ps are useful to highlight some unique aspects of e-marketing (Kalyanam & McIntyre 2002).

To update the marketing mix for the Internet, Kalyanam and McIntyre (2002) proposed the
e-marketing mix that included new elements such as personalization, privacy policy, and website design. Dutta and Biren (2001) used the 4Ps model comprising customer relationships and technologies of interactivity and connectivity as a tested framework (called the Marketspace Model) for evaluating business transformation on the Internet. Blum and Fallon (2002) examined 53 Welsh visitor attraction websites according to product, price, promotion, place, customer relations, and technical aspects. In addition, Law and Leung (2000) evaluated the efficiency of airlines’ online reservation services according to online reservation services, provision of extra benefits, reservation time, and the availability of additional services/facilities.

2.2 The perspective of e-quality

The term “e-quality” is important for assessing e-commerce websites. Aladwani and Palvia (2002) considered Web quality to be a complex thing and multidimensional measurement in nature. Madu and Madu (2002) noted that the dimensions of e-quality may be different from the traditional practice of quality. DeLone and McLean’s updated information systems (IS) success model (2003) consists of three quality factors: information quality, system quality and service quality. Parasuraman et al. (1985) conceptualized service quality as the relative perceptual gaps between exceptions and perceptions of performance levels and then developed the gap model as an approach to measure service quality. According to this gap model, the SERVQUAL instrument, which includes the five dimensions of tangibles, reliability, responsiveness, assurance, and empathy, was developed by Parasuraman et al. (1988). Subsequently, Zeithaml et al. (2000) developed e-SERVQUAL to measure electronic service quality as an updated measure of traditional SERVQUAL model in the Web setting. This multi-item scale includes the seven dimensions of efficiency, reliability, fulfillment, privacy, responsiveness, compensation, and contact.

Website evaluation has been widely studied in the literature. Lee and Kozar (2006) evaluated online electronics and online travel websites by adopting DeLone and McLean’s IS success model and applying the Analytic Hierarchy Process (AHP) method. AHP can only obtain the relative weights of alternatives; it cannot compute gaps between the status quo and an ideal point. Büyüközkan et al. (2007) used the Fuzzy VIKOR method to evaluate 21 e-learning websites according to seven criteria. They ignored interrelationships between these seven criteria when determining their weights. Büyüközkan and Ruan (2007) used fuzzy AHP and fuzzy TOPSIS to rank 13 Turkish government websites according to six e-service quality dimensions. TOPSIS, however, has some limitations. According to Wang et al. (2007), TOPSIS’ closeness coefficient values do not reflect the superiority or inferiority of alternatives and therefore cannot be used for ranking purposes. Baloglu and Pekcan (2006) applied content analysis to analyze the websites of hotels in Turkey in terms of site design and marketing characteristics using a measurement variable of yes-or-no (one-or-zero). The shortcoming of binary variables is that they cannot express the performance on each criterion. Wan (2002) also applied content analysis to evaluate the websites of 30 tourist hotels and 39 tour wholesalers using a 5-point rating scale. He took into account the performance on each criterion, but ignored the relative importance of various criteria. Shchiglik and Barnes (2004) used a 5-point scale and important ratings to evaluate website quality according to four dimensions (i.e. site quality, information quality, interaction quality, and airline-specific quality) for the airline industry in New Zealand. They took into account the global weighted average performance, but neglected to specifically consider the worst one.

As mentioned above, previous studies have failed to provide a comprehensive and systematic approach that quantitatively measures a website’s overall performance, and their research methodologies must be improved. Therefore, this study proposes an effective model for assessing e-commerce websites in terms of combined “marketing mix 4Ps” and “website quality” perspectives. This model overcomes the drawbacks of prior studies and offers enough insights for practitioners to accurately measure the current level of their websites according to critical criteria that determine their competitive advantages.
3 AN INTEGRATED EVALUATION MODEL

The assessment process of e-commerce websites required the construction of an evaluation model. To effectively evaluate e-commerce websites of airlines, an integrated MCDM model is proposed. The proposed model includes six phases: (1) defining the goal; (2) establishing the analytic structure; (3) using the DEMATEL method; (4) using the ANP method; (5) applying the modified VIKOR method; (6) finding the ranking of airlines’ websites. Accordingly, an overview of the integrated evaluation model for e-commerce websites is shown in Figure 1. The details of each phase are described as follows.

- (1) Defining the decision goal
- (2) Establishing the analytic structure
- (3) Using the DEMATEL to analyze interdependent relationship between criteria
- (4) Using the ANP to calculate the weight of each criterion
- (5) Applying the modified VIKOR to calculate the index of each alternative
- (6) Ranking alternatives and finding a set of compromise solutions

Figure 1. The integrated evaluation model for e-commerce websites.

The first phase defines the decision goal. The goal is to evaluate and rank e-commerce websites for the airline industry. In phase 2, the analytic structure, according to related literature and expert interviews, is used to assess e-commerce websites. The four groups are defined and identified in this structure, including: goals, perspectives, criteria, and alternatives. Such combination shapes an evaluation mechanism for airlines’ websites, as shown in Figure 2. As seen in this figure, this MCDM problem is considered in terms of two perspectives. There is no interrelationship between these two perspectives. On the perspective of “e-marketing,” the criteria help airlines create more sales such as advertising, price negotiation, online order and payment, and online community features. They are: “high value-added product (P1),” “product customization (P2),” “price negotiation (P3),” “low price (P4),” “discount (P5),” “advertising (P6),” “communication (P7),” and “transaction function (P8).” These eight criteria are then divided into four clusters: product, price, promotion, and place. On the perspective of “e-quality,” the criteria support the website’s designer to create a high quality website such as a personalized user interface and ease of navigation. They are: “reliability (P9),” “responsiveness (P10),” “credibility (P11),” “currency (P12),” “relevance (P13),” “personalization (P14),” “navigability (P15),” “security (P16).” These eight criteria are then divided into three clusters: service quality, information quality, and system quality. In addition, the “alternatives” comprises the five airlines’ websites.

Phase 3 applies the DEMATEL method prior to the ANP method in order to improve the procedure for dealing with the interrelationships between criteria. DEMATEL quantifies complex relationships between criteria and converts them into a visible structural map. It has been successfully applied in a wide range of situations such as e-learning programs evaluation (Tseng et al. 2007), safety measurement (Liou et al. 2007), and social responsibility programs selection (Tsi & Hsu 2008) problems. A series of steps are implemented to complete the analysis of this phase (for a detailed exposition of these steps, please refer to the above mentioned studies). Finally, the strength of
interdependence between criteria and an influence-relation-map are determined from the results of the DEMATEL.

![Diagram](image-url)

**Figure 2.** The analytic structure for website evaluation.

Phase 4 applies the ANP method to calculate the relative importance of each criterion. This method is a comprehensive decision-making technique that has the capability to comprise all the interdependent criteria in arriving at a decision (Jharkharia & Shankar 2007). The AHP suffers from the limitation of only being applicable to straightforward hierarchical structures while the ANP can be applicable to the complexities of many real world problems (Chang et al. 2007). In reality, however, such evaluation criteria are seldom independent (Liou et al. 2007). Therefore, instead of adopting the commonly used AHP method for solving these types of problems, we apply an ANP-based model for calculating the weight of each criterion. For the detailed steps of the ANP method the reader is referred to Saaty (2001).

Phase 5 applies the modified VIKOR method to compute gaps between the status quo and an ideal website. The VIKOR method provides measurements of determining the aggregate distance \(R_j\) to the ideal point. This method considers two distance measurements, \(S_j\) and \(Q_j\), based on an aggregating function \(D_p\)-metric in the compromising programming method. \(D_p\) represents the distance of an alternative from the ideal solution for different levels of \(p\). \(S_j\) is the \(j^{th}\) alternative with respect to all criteria as calculated by the average distance from the ideal solution. \(Q_j\) is the \(j^{th}\) alternative with respect to a specific criterion as calculated by the maximum distance from the ideal solution. The traditional \(R_j\) does not indicate how many performance variations exist between the status quo and the ideal point; therefore, following Ou Yang et al. (2008), we use the modified \(R_{j}^{mod}\) index to overcome this drawback. Eq. (1) is used instead of the traditional \(Q_j\). Then, the modified
\[ R_{ij}^{mod} \] index is listed in Eq. (2).

\[ Q_{ij}^{mod} = \max_i \left\{ \left( \frac{x_i - x_j}{x_i} \right) \right\} \text{ for } j = 1, 2, \ldots, m \] (1)

\[ R_{ij}^{mod} = v S_{ij} + (1 - v) Q_{ij} \] (2)

Finally, phase 6 obtains the ranking of e-commerce websites and a set of compromise solutions according to the modified VIKOR index.

4 AN EMPIRICAL STUDY OF THE AIRLINE INDUSTRY

To describe the proposed model clearly, an empirical study of the Taiwanese airline industry was conducted in order to demonstrate the efficacy of this model for assessing and ranking e-commerce websites. There are five domestic airlines in Taiwan: China Airlines, Eva Air, Uni Air, Mandarin Airlines, and TransAsia Airways. Their e-commerce websites are indicated as A1 (China Airlines), A2 (Eva Air), A3 (Uni Air), A4 (Mandarin Airlines), and A5 (TransAsia Airways), respectively. We invited 32 experts to express their opinions in September 2008. Nineteen of them were from various industries and the remaining ones were from academics and research institutes. These experts are professionals who have been designing e-commerce websites or studying on various projects about the Internet marketing for a certain time; therefore, their answers to the questionnaires can appropriately reflect the status quo of e-commerce websites of the five airlines.

4.1 Calculation of the weights of evaluation criteria

First, the experts examined the status quo of the five airlines’ websites beforehand and then used the cut-off value method with the 9-point scale to screen for proper criteria. According to the survey results, 16 criteria were constructed as very important to the evaluation of airlines’ websites, with a mean value exceeding 6.0 (i.e. cut-off value) for each criterion. Second, a zero–four DEMATEL scale was used for determining the relationship structure among the 16 criteria. Once the relationships between criteria had been measured by the experts, an initial direct-relation matrix could then be obtained. On the basis of the initial direct-relation matrix, a normalized direct-relation matrix was obtained by using the DEMATEL formula. Next, a total-relation matrix (including D, R, D+R and D–R) was identified (Table 1). Finally, a threshold value was chosen as 0.66 to filter the minor effects in the element of the total-relation matrix. After deciding on the threshold value, the influence-relation-map could be obtained by mapping a dataset of (D+R, D–R) (Fig 3).

<table>
<thead>
<tr>
<th>( P_1 )</th>
<th>( P_2 )</th>
<th>( P_3 )</th>
<th>( P_4 )</th>
<th>( P_5 )</th>
<th>( P_6 )</th>
<th>( P_7 )</th>
<th>( P_8 )</th>
<th>( P_9 )</th>
<th>( P_{10} )</th>
<th>( P_{11} )</th>
<th>( P_{12} )</th>
<th>( P_{13} )</th>
<th>( P_{14} )</th>
<th>( P_{15} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.999</td>
<td>0.696</td>
<td>0.735</td>
<td>0.729</td>
<td>0.718</td>
<td>0.670</td>
<td>0.676</td>
<td>0.741</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.638</td>
</tr>
<tr>
<td>0.741</td>
<td>0.544</td>
<td>0.752</td>
<td>0.728</td>
<td>0.722</td>
<td>0.676</td>
<td>0.672</td>
<td>0.736</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.638</td>
</tr>
<tr>
<td>0.669</td>
<td>0.013</td>
<td>0.619</td>
<td>0.769</td>
<td>0.722</td>
<td>0.640</td>
<td>0.641</td>
<td>0.743</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.581</td>
</tr>
<tr>
<td>0.657</td>
<td>0.981</td>
<td>0.713</td>
<td>0.988</td>
<td>0.726</td>
<td>0.539</td>
<td>0.591</td>
<td>0.691</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.579</td>
</tr>
<tr>
<td>0.638</td>
<td>0.967</td>
<td>0.687</td>
<td>0.723</td>
<td>0.579</td>
<td>0.637</td>
<td>0.567</td>
<td>0.688</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.560</td>
</tr>
<tr>
<td>0.684</td>
<td>0.616</td>
<td>0.709</td>
<td>0.740</td>
<td>0.748</td>
<td>0.550</td>
<td>0.647</td>
<td>0.711</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.514</td>
</tr>
<tr>
<td>0.750</td>
<td>0.982</td>
<td>0.803</td>
<td>0.770</td>
<td>0.765</td>
<td>0.685</td>
<td>0.582</td>
<td>0.763</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.507</td>
</tr>
<tr>
<td>0.706</td>
<td>0.831</td>
<td>0.746</td>
<td>0.728</td>
<td>0.717</td>
<td>0.632</td>
<td>0.653</td>
<td>0.612</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.507</td>
</tr>
<tr>
<td>0.899</td>
<td>0.099</td>
<td>0.099</td>
<td>0.099</td>
<td>0.099</td>
<td>0.099</td>
<td>0.099</td>
<td>0.099</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.507</td>
</tr>
<tr>
<td>0.750</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.507</td>
</tr>
<tr>
<td>0.750</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.507</td>
</tr>
<tr>
<td>0.750</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.507</td>
</tr>
<tr>
<td>0.750</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.507</td>
</tr>
<tr>
<td>0.750</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.507</td>
</tr>
<tr>
<td>0.750</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.699</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.507</td>
</tr>
</tbody>
</table>

Table 1. The total-relation matrix of the 16 criteria.
As shown in Figure 3, communication ($P_7$), with the highest value of D–R, was called the master dispatcher. Similarly, product customization ($P_2$) obtained the second highest value of D–R and became the second powerful criterion. Navigability ($P_{15}$), with lowest value of D–R, was the master receiver.

Subsequently, the ANP method was employed to determine the relative importance of each criterion. The experts first responded to the questionnaire through a series of pair-wise comparisons with Saaty’s one-nine scale, comparing the relative importance of one element over another. A scale of one to nine ranges from equal importance to extreme importance (Saaty 2001). Super Decisions 1.6.0 software was applied to aid in the calculations. This software is an easy-to-use tool for constructing network-like decision models with the ability of calculating the weight of each element. Due to space limitations, the pair-wise comparisons are not shown. Second, consistency ratios of all the pair-wise comparison matrices were calculated. The consistency measure is very useful for identifying possible errors in judgments. When consistency ratios are less than 0.1, the judgments are considered to be reliable. In this study, consistency ratios were less than the acceptable threshold value and the eigenvectors displayed were appropriate to enter into the unweighted supermatrix. Third, the experts conducted the pair-wise comparisons on the clusters. Then, the unweighted supermatrix was transformed to be column-stochastic as shown by $M^W$ after completing the pair-wise comparisons of the clusters. Finally, the weighted supermatrix was raised to limiting powers of $M^L$ to capture all interactions and obtained a steady-state outcome. The limit supermatrix is shown in Table 2. The results of the limit supermatrix yielded ($P_1$, $P_2$, $P_3$, $P_4$, $P_5$, $P_6$, $P_7$, $P_8$, $P_9$, $P_{10}$, $P_{11}$, $P_{12}$, $P_{13}$, $P_{14}$, $P_{15}$, $P_{16}$) = (0.084, 0.024, 0.115, 0.129, 0.130, 0.011, 0.014, 0.128, 0.088, 0.049, 0.081, 0.042, 0.026, 0.015, 0.028, 0.036). Ranked by the weights, the top five evaluation criteria in order of importance were: discount ($P_5$), low price ($P_4$), transaction function ($P_8$), price negotiation ($P_3$) and reliability ($P_9$), respectively. This implies that airlines’ proprietors need to pay more attention to these five criteria for achieving sustainable competitive advantages in the e-marketplace. Under the perspective of e-marketing,

Figure 3. The influence-relation-map the 16 criteria.
discount ($P_5$) had the highest weight of 0.130. This criterion is a key factor for successful e-marketing. Apparently, the use of online discount is essential requirement for effective promotion.

<table>
<thead>
<tr>
<th>Goal</th>
<th>Perspectives</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>EW</td>
<td>EM</td>
<td>EQ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. The limit supermatrix.

4.2 Performance measurement of the websites

After finishing a series of pair-wise comparisons, the experts were asked to provide linguistic values for evaluating the five airlines’ websites. In this study, linguistic values were used to design the evaluation questionnaire. These performance values, which were very good, good, median, poor, and very poor, were transformed by scaling them into the numbers 100, 75, 50, 25, 0, respectively. The results of evaluation are shown in Table 3. This table shows that all five websites were not utilize the Internet to its full potential and performed poorly in terms of price negotiation ($P_3$), low price ($P_4$), responsiveness ($P_6$), and communication ($P_7$). In addition, a coordinate diagram (Figure 4) was acquired by mapping the dataset of (e-marketing score, e-quality score), where the horizontal axis was “e-marketing” score, and the vertical axis was “e-quality” score. Figure 4 shows that $A_2$ had the best overall performance and was the closest to the ideal point ($A'$). Nevertheless, $A_2$ was the farthest from the ideal point, both “e-marketing” and “e-quality” performed poorly. These two perspectives needed the most improvement.

Next, the experts determined that a score of 100 represented the positive-ideal solution ($x^+_i$) and a score of zero represented the negative-ideal solution ($x^-_i$). The performance score was calculated using normalized formula (i.e. $x_i^+ - x_j^-$) to obtain performance variance rates between the status quo and the ideal point. Finally, $S$, $Q^mod_j$, and $R^mod_j$ were computed by selecting $\nu=0.5$ (Table 4). Given these results, we observed that $A_2$ did not have an acceptable advantage; in other words, $R^{[2]} = 0.001 \leq DR = 0.021$. On the other hand, we observed that $A_2$ was stable within the decision-making process; in other words, it was the highest ranked in $S^1$. Therefore, we proposed $A_1$ and $A_2$ as a set of compromise solutions. Looking at Table 4, the values $R^mod_j$ are ($A_1$, $A_2$, $A_3$, $A_4$, $A_5$)=(0.436, 0.435, 0.485, 0.518, 0.494), in which the ranking is: $A_2 > A_1 > A_3 > A_5 > A_4$, where $A > B$ indicates that A is preferred to B.
The results show that the comprehensive variance rate of A2 was 0.435. This website still needed improvement in order to achieve its ideal point or desired level. On the other hand, A4 was the farthest from the ideal solution, as its $R^\text{mod}_j (0.518)$ was larger than all others. A particular attention should be given to price negotiation (P3) and low price (P4) because these two criteria had higher weights and lower performance than others. Price negotiation (P3) was in most urgent need of improvement for airlines’ proprietors. Consideration should also be given to the possibility of developing an “online price negotiation” area for special customers such as group traveler. In addition, we suggested that the airlines should try to improve their pricing mechanisms.
4.3 Discussions

The value of the weight $v$ had a central role in the ranking of alternatives. According to the VIKOR, when $v$ was larger (>0.5), the value $R_j$ will prefer majority rule. A scenarios analysis can be undertaken by setting $v$ systematically to some values between 0 and 1. To give a valuable insight for the change of evaluators’ preferences (or concerns), we proposed another six scenarios for evaluating e-commerce websites; six sets of different weights were associated with six scenarios. Table 5 presents six sets weight values and the ranking of the VIKOR. The results of this table indicated that alternative $A_1$ was the best ranked, with an acceptable advantage, for scenario 1. From scenario 2 to scenario 6, the compromise sets were obtained {A$_1$, A$_2$}, the first ranked alternative had no advantage to be a single solution.

It was interesting, however, to note that the value $S_j$ of $A_2$ was the smallest in Table 4 but the value $R^\text{mod}_j$ of $A_2$ was the second smallest in scenario 2 and scenario 3. In other words, $A_2$ had the smallest average variance rate, but it only obtained the second ranking order in the VIKOR method. Moreover, $A_2$ had the smallest variance rate and the best ranking order in scenario 4, scenario 5 and scenario 6. On the other hand, when $v$ was smaller than 0.5, $A_1$ had the smallest variance rate and the best ranking order. $A_2$ had the largest variance rate and the worst ranking order in all scenarios. According to the results of these scenario analyses, we could find the benefits of the VIKOR method since it additionally considers the value $Q^\text{mod}_j$ which is the farthest from the positive-ideal solution (i.e. the largest variance rate) in the VIKOR index. On the contrary, the shortcoming of traditional performance evaluation is that it only takes into account weighted average scores and neglects to specifically consider the worst one. Thus, the decision-makers can select a suitable weight ($v$) in terms of their preferences (or concerns) to make an appropriate decision.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
Websites & $S_j$ & \multicolumn{2}{|c|}{$Q^\text{mod}_j$} & \multicolumn{2}{|c|}{$R^\text{mod}_j$} \\
& Value & Ranking & Value & Ranking & Value & Ranking \\
\hline
$A_1$ & 0.324 & 2 & 0.547 & 1 & 0.436 & 2 \\
$A_2$ & 0.308 & 1 & 0.563 & 2 & 0.435 & 1 \\
$A_3$ & 0.408 & 4 & 0.563 & 2 & 0.485 & 3 \\
$A_4$ & 0.450 & 5 & 0.586 & 3 & 0.518 & 5 \\
$A_5$ & 0.403 & 3 & 0.586 & 3 & 0.494 & 4 \\
\hline
\end{tabular}

*Table 4. Ranking of the five websites by VIKOR ($v=0.5$).*

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
Websites & Scenario 1 & Scenario 2 & Scenario 3 & Scenario 4 & Scenario 5 & Scenario 6 \\
& $R^\text{mod}_j$ ($v=0$) & $R^\text{mod}_j$ ($v=0.2$) & $R^\text{mod}_j$ ($v=0.4$) & $R^\text{mod}_j$ ($v=0.6$) & $R^\text{mod}_j$ ($v=0.8$) & $R^\text{mod}_j$ ($v=1$) \\
& Value & Ranking & Value & Ranking & Value & Ranking & Value & Ranking & Value & Ranking & Value & Ranking \\
\hline
$A_1$ & 0.547 & 1 & 0.502 & 1 & 0.458 & 1 & 0.413 & 2 & 0.369 & 2 & 0.324 & 2 \\
$A_2$ & 0.563 & 2 & 0.512 & 2 & 0.461 & 2 & 0.410 & 1 & 0.359 & 1 & 0.308 & 1 \\
$A_3$ & 0.563 & 2 & 0.533 & 3 & 0.501 & 3 & 0.470 & 3 & 0.439 & 3 & 0.408 & 4 \\
$A_4$ & 0.586 & 3 & 0.559 & 5 & 0.531 & 5 & 0.504 & 5 & 0.477 & 5 & 0.450 & 5 \\
$A_5$ & 0.586 & 3 & 0.549 & 4 & 0.513 & 4 & 0.476 & 4 & 0.440 & 4 & 0.403 & 3 \\
\hline
Compromise solutions & $A_1$, $A_2$ & $A_1$, $A_2$ & $A_1$, $A_2$ & $A_1$, $A_2$ & $A_1$, $A_2$ & $A_1$, $A_2$ \\
\hline
\end{tabular}

*Table 5. Ranking of the five websites in six scenarios.*
5 CONCLUSIONS

In view of the pressure from the current fierce competition and high oil costs, airlines gradually turn to the Internet marketing methods to increase competitive advantage. Airlines are concerned about how to build an ideal e-commerce website by comprehensively considering the role of a website on technology, service and marketing constructs.

This study proposed an integrated model combining the DEMATEL method, the ANP method and the modified VIKOR method for evaluating e-commerce websites of airlines and provided some directions to airlines on how to improve overall e-commerce performance. First, the DEMATEL were used to construct the interrelationships between criteria. Second, the ANP were employed to calculate the relative importance of each criterion. The weights obtained through ANP were then combined with the modified VIKOR method to compute performance variance rates between the status quo and the ideal point for ranking the five Taiwanese airlines’ websites. The results showed that $A_1$ and $A_2$ were the two best websites and that $A_4$ was the worst one. Moreover, a scenario analysis based on the change of experts’ preferences was also provided. The objective of scenario analyses was to provide an insight to decision-makers when a parameter was changing. Finally, the overall results showed that the Taiwanese airlines did not utilize the Internet to its full potential and still had a great deal of room to improve their websites. All five websites performed poorly in terms of price negotiation ($P_3$), low price ($P_4$), responsiveness ($P_{10}$), and communication ($P_7$). To apply the Internet marketing effectively, these airlines should put efforts to add an “online price negotiation” feature to site and adjust pricing strategies. Airlines can use their websites to gather information about customer needs, buying patterns and preferences. This information can be a worthwhile input to the development of high value-added products and services. Furthermore, although this paper is related to the airline industry in Taiwan, the same concept can be applied and extended to other industries to handle any evaluation problem with interdependent criteria. However, this study also has some limitations. First, survey data were collected from a limited number of visits to each website at a specific time. Due to the highly dynamic nature of these websites, similar studies at different times are quite likely to show different results. Second, this study transformed linguistic values into five crisp numerical values for measuring a website’s performance, but this might not be true in the real world. Concerning future research, it would be beneficial to extend this study to a fuzzy environment. In addition, the benchmarking analysis provides a means to enhance the websites’ competitive advantages by learning from the best practices of the industry. Researchers could also extend this study to the benchmarking analysis.

Acknowledgement

The authors would like to thank the National Science Council of Taiwan for financially supporting this research under Grant NSC97-2410-H-008-029.

References


# ASSESSING THE BUSINESS VALUE OF ELECTRONIC ORDER-TO-PAYMENT CYCLE

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0177.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Case Study, IS metrics, IT performance management, IT investment</td>
</tr>
</tbody>
</table>

---

Proceedings ECIS 2009
ASSESSING THE BUSINESS VALUE OF ELECTRONIC ORDER-TO-PAYMENT CYCLE

Lempinen, Heikki, Helsinki School of Economics, Runeberginkatu 14-16, 00100 Helsinki, Finland, heikki.lempinen@hse.fi
Penttinen, Esko, Helsinki School of Economics, Runeberginkatu 14-16, 00100 Helsinki, Finland, esko.penttinen@hse.fi

Abstract

In this paper, we build an evaluation tool for assessing the business impacts of an electronic order-to-payment cycle. Based on a literature review and expert interviews, we formulate a three-stage model which includes performance indicators for electronic order, electronic invoice, and electronic payment processes. In addition, we pinpoint the inter-process linkages. We test the proposed evaluation tool in a business context and find that the impacts of automating the order-to-payment cycle relate closely to cost avoidance. However, a strong emphasis on asset utilization can be observed as well – better use of IT could enhance utilization of existing human resources and capital, affecting company profitability.

Keywords: Performance indicators, evaluation tool, e-order, e-invoice, e-payment.
1 INTRODUCTION

How can we utilize information technology (IT) to increase productivity? What IT innovations enable us to perform tasks in smarter ways than we used to? One important source of productivity growth lies in improving the processes of the financial administration. Using IT in financial administration has been recognised as one of the most important sources of profitability growth in Europe (European Commission 2007; EU 2006). As an example, EACT has estimated that by moving from paper-based invoicing to electronic invoicing, companies could save 243 billion euros in processing costs alone (European Commission 2007). Therefore, increasing productivity by using information technology (IT) in the financial administration is high on the EU agenda (EU 2006).

The extant literature provides a plethora of metrics and measures related to supply chain performance and financial administration. What is missing, though, in the current literature, is a holistic analysis on the potential productivity gains that could be derived from a full-scale digitalization of the order-to-payment cycle. Therefore, the objective of this study is to pinpoint the potential productivity gains from adopting an electronic order-to-payment cycle. To do this, based on an extensive review of the existing literature and expert interviews, we divide the order-to-payment cycle into three distinct processes (ordering, invoicing, and payment) and build an evaluation tool to assess the potential productivity gains. In addition, we report the findings of a multiple case study, testing the proposed evaluation tool in a business context. The proposed evaluation tool is developed from the point of view of the buying organization.

The structure of the study is as follows. In the second section, we discuss the literature on business value of information technology. In the third section, we develop the evaluation tool. In the fourth section, we test the proposed evaluation tool in the context of a Finnish design company. In the remaining sections, we provide some anecdotal evidence from two other case studies, draw conclusions, and suggest avenues for further research.

2 BUSINESS VALUE OF INFORMATION TECHNOLOGY

Effective measurement of information system success has been a serious concern for both managers and scholars - attempts to quantify the benefits of IT have often resulted in inconclusive or inconsistent results (Byrd et al. 2006). Research in IT business value examines the organizational performance impacts of information technology, such as productivity and profitability improvement, cost and inventory reduction, competitive advantage and other performance measures (Melville et al. 2004). There is a vast amount of studies regarding IT business value and almost as many proposed instruments to evaluate it; quantitative financial measures, information value measures, service quality tools (SERVQUAL), and multi-dimensional analysis have been used, among others (see Cronk & Fitzgerald 1999 for review).

2.1 Productivity paradox and beyond

In spite of the great promise of IT driving the biggest technological revolution men have known, there has been heated debate in the IS literature for the past decades about whether IT usage actually pays off (Brynjolfsson 1993). Labelled the productivity paradox of information technology, Brynjolfsson (1993) explains that although computing power in the (U.S.) economy has increased by more than two orders since 1970s, productivity seems to have stagnated. Particularly in the 80s, many studies claimed that the overall IT productivity impacts are neutral or even negative (e.g. Salerno 1985). Robert Solow, winner of the Nobel Price in economics 1987 stated that “we see the computer age everywhere except in the productivity statistics” (New York Times Book Review 1987). Yet, the well established view today is that IT-intensive firms are more productive (e.g. Dedrick et al. 2003, Aral et al. 2006).
According to Brynjolfsson & Hitt (1998), the critical question for IT managers is not “Does IT pay off?” but “how can we best use computers?” Brynjolfsson & Hitt (1998) refer to the business value of IT as its ability to contribute to productivity growth. Productivity growth comes from using resources more efficiently, or as Brynjolfsson & Hitt (1998) express it “productivity growth comes from working smarter”. This is the promise of computers; this is what IT should be able to offer. According to Tallon et al. (2000), some insights into IT payoff can be attained by firm-level research on the “productivity paradox”, principally in the form of returns on IT investment (IT productivity on the company level can be measured by comparing some IT factor to an organizational performance measure, e.g. annual IT expenditure vs. pre-tax profit). They argue, however, that too little attention has been given to other IT impacts such as improved inventory management, greater product variety and customer service. To be able to fully understand the benefits, they claim, additional indicators should be considered.

Although the discussion so far has provided few practical tools for assessing the business value of IT regarding this study, some valuable points have emerged. First of all, it is crucial to see that there is a variety of tangible and intangible variables to consider – not all IT investments are made simply to cut cost. Second, a process management and design view should be highlighted when building the evaluation tool. Gonzales-Benito (2007) found that IT investments have positive effect on operational purchasing performance since the use of IT allow companies to adopt certain purchasing practices and facilitates greater strategic integration of the purchasing function. Finally, it has become clear that the majority of the literature reviewed so far aim to measure direct impacts of IT on company level performance or even on a wider scope. Now the question remains how to measure business impacts of an electronic order-to-pay process; one specific business activity inside a company?

According to Silvius (2006), there are two distinct approaches to be found in IT business value literature; the variance approach investigating what the relationship between IT investment and organizational performance is, while the process approach tries to find out how this relationship works. Most research presented so far try to explain the productivity paradox by investigating direct company level impacts of IT and thus falls under the “variance approach” category. A growing body of IT business value research, however, prefer the process approach, suggesting a multi-dimensional impact structure in the organizational hierarchy.

2.2 Process-oriented approach for evaluating IT business value derived from the order to payment cycle

To gain insight on activity-specific impacts inside the process, it has to be split open into smaller phases and try to identify attainable benefits in each phase. This will, however, result only in a group of operative-level performance indicators which hardly give much insight on business impacts of automating the entire chain. On the other hand, the overall company, industry or macro-economic level analysis of IT productivity impacts are too general and cannot provide information that is accurate enough for process monitoring purposes. Thus, there should be some instrumentation in the middle to combine the lower-level and the top-level business impact indicators.

According to Barua et al. (1995), the growing concern of scholars is that IT effects on the enterprise level performance can be identified only through a web of intermediate level contributions - there is some evidence that IT impacts exist and that they can be detected by lower-level analysis in the organization. The lower level impacts, in turn, are expected to affect higher level performance. To evaluate these, Barua et al. (1995) propose a process-oriented methodology, which involves a two stage analysis of intermediate and higher level output variables. They strive to open up the black box of IT usage, detect and measure IT impacts where they occur - their main thesis is that economic contributions of IT can be measured at the operational level, where IT systems are implemented. The higher order impacts can be then traced through a chain of relationships within the organizational hierarchy.
An important side note to make here is that actually, the functions; ordering, invoicing and payment are better yet sub-processes of the order-to-pay activity. Tallon et al. (2000) argue that IT creates value for the company via individual business processes, or inter-process linkages, or both – the greater the impact of IT on processes and inter-process linkages, the greater the contribution of IT to firm performance. This encourages investigating IT impacts of individual sub-processes e-ordering, e-invoicing and e-payment but also pay great attention to IT impacts on inter-process linkages – in other words investigate the impact of electronic integration in an order-to-pay process.

We divide IT impacts to three levels; application-specific impacts (e-ordering, e-invoicing, and e-payment), integrated process level impacts (order-to-pay) and finally the company level business impacts. First of all, the use of IT affects sub-processes directly, resulting in some impacts which are specific to each function. Second, IT usage impacts the entire order-to-pay process via inter-process linkages thus enabling full chain optimization. Finally, linkages between the process and the company level reveal the contribution to company level business value of IT. In short, automating an order-to-payment process yields application-specific impacts and through some yet unidentifiable intermediate variables contributes to overall IT business value for the company.

### Figure 1. Conceptual framework for the study.

#### 2.3 Uncovering the hierarchy: identifying the chains of IT impacts

To better identify the chains of different stage indicators for IT impact evaluation, some studies have borrowed a very well known multi-stage framework from the strategy literature. The Balanced Scorecard is a tool that translates a company’s mission and strategy to a comprehensive set of organizational performance measures (Kaplan & Norton 1996). The Balanced Scorecard communicates a holistic model that links individual efforts and accomplishments to business unit objectives through a consistent series of objectives and measures. The idea is that each individual person could see how their efforts contribute to achieving broad organizational goals, thus offering means for organizational learning and improvement. To be able to track these linkages, however, the model needs to be more than a collection of critical indicators – the scorecard incorporates a complex set of cause-and-effect relationships among outcome measures and the related performance drivers. These relationships can be then traced through sequences of “if-then” statements.

Respectively, efforts to increase IT intensity in a specific function could be seen to affect company level financial measures through a similar chain. Following this idea, Epstein & Rejc (2005) introduce the IT Balanced Scorecard which is a tool for tackling the problem of properly evaluating IT payoffs. The authors argue that an IT performance measurement and management system must focus determining the key drivers of IT success and the causal relationships among them and develop numerous performance measures to track IT performance.
3 METHODOLOGY

Since the purpose of this study is to find out how IT creates business value in the order-to-payment cycle context, the first logical step was to conduct an extensive review on the general subject of business value of IT. It soon became clear that there are two quite different approaches to evaluate the impacts of IT, the variance approach studying what the impact is and on the process approach studying how this value is created. The process approach was chosen because of its suitability regarding our research question i.e. to answer the question “how”. A group of process-oriented models were reviewed next. Due to the specificity of the context, we soon realised that there is not a possibility to choose only one model and use it as such, but rather combine elements of a couple of them and create a modified evaluation tool.

As the second part of the literature review, we collected benefits regarding e-ordering, e-invoicing and e-payment to gain in-depth knowledge of order-to-payment cycle automation potential. Also, interviews with field specialists were conducted to get comments and new ideas along the way. The expert interviews were particularly useful in determining the value created by IT impacts on the inter-process linkages – it is something that until now was implicitly understood by business people but not explicitly stated in the books.

To illustrate and test how the proposed evaluation tool works, we chose to use the case study methodology to make an in-depth analysis in one case company. The case study methodology has distinct advantage when a “how” or “why” question is being asked about a contemporary set of events, over which the investigator has little or no control (Yin, 1994). The case study methodology was especially suitable for our research as we investigate how companies use information technology in their financial administration and what performance indicators they perceive as important in the process. Furthermore, the strength of the case study approach is that it enables the capture of “reality” in considerably greater detail and the analysis of a considerably greater number of variables than is possible with many other approaches (Galliers, 1991). This was especially important in our work as the objective of this research was to find the underlying performance indicators to IT investments in financial administration, and not just the company policy statements.

More specifically, to develop the evaluation tool, we conducted 10 industry expert interviews (leading specialists in companies providing e-payment, e-invoice, and e-procurement services, the Federation of Finnish banks, the Finnish Information Society Development Centre etc.) and academic professionals (professors at Finnish business universities). For the illustrative case part, at Marimekko, we interviewed six people in total each from different functions that were relevant regarding our study. We covered extensively the entire order-to-payment cycle by interviewing people from buying, invoice handling, treasury, IT system development, and supply chain management. In addition, we conducted supplementary case studies in two large Finnish companies. These case studies included altogether eight interviews.

4 DEVELOPMENT OF THE EVALUATION TOOL

In developing the evaluation tool to analyze the order-to-payment cycle, we adopt the process view discussed above and divide the order-to-payment cycle into three distinct processes: e-ordering, e-invoicing, and e-payment.

4.1 E-ordering impacts and performance indicators

The classic argument for adopting e-procurement systems is that it creates substantial cost savings (e.g. Bakos 1997). To find out how these savings can be achieved, Johnson & Klassen (2005) discuss three different dimension of e-procurement in their article; e-sourcing, e-coordination and e-communities. Many e-procurement studies address e-sourcing issues e.g. the emergence of electronic
(reverse) auctions and other transparent e-marketplace structures which have, e.g., allowed companies to negotiate better prices for purchases and reduce search costs. E-communities, on the other hand, refer to different e-procurement systems platform structures. Proprietary platform procurement systems are point to point links between buyers and suppliers (often EDI-systems), open platform procurement systems are open Internet-based systems, and hybrid platforms have elements of both.

Johnson & Klassen (2005) note that reported benefits of e-coordination are diverse and varied, including cost savings from process improvements (less slack and reduced rework due to errors), price reductions, greater visibility of orders and enhanced inventory turnover and accuracy, among other things. Also, it has been reported that fewer Request for Proposals (RFPs) elicit no bids when sent electronically. Ordering automation can lead to purchase price reductions due multiple reasons. First, since e-ordering allows for a more coordinated and centralized buying process, orders can be bundled to be able to negotiate volume discounts. Second, e-ordering systems provide a transparent bidding platform for a growing number of suppliers, thus increasing price competition. Reduced probability of human mistakes in ordering and storing due to the use of IT allows improved inventory accuracy. Enhanced inventory turnover, in turn, is largely due to reductions in average inventory levels.

Inventory level reductions come from increased inventory accuracy and enhanced transparency of incoming raw material and replenishment deliveries. Other effectiveness and efficiency gains provided by e-ordering systems include lower transaction costs mainly due to less manual work in the process (Presutti 2003) and shorter order cycle times meaning that users receive requested goods or services faster (e.g. Johnson & Klassen 2005, Presutti 2003, Reunis et al. 2001). The following figure (Figure 2) depicts the performance indicators for the e-ordering process.

![E-ordering performance indicators diagram](image)

Figure 2. E-ordering performance indicators.

4.2 E-invoicing impacts and performance indicators

The benefits from the electronic invoicing process of invoicing stem primarily from the decreased processing costs and numbers of errors, reduction of materials use, and decreasing handling times. Penttinen & Hyytiäinen (2008) argue that adopting e-invoicing has clear benefits; transition from
paper bills to e-invoicing brings considerable financial savings. It has been estimated that an incoming paper invoice incurs cost of 30-50 euros to the receiving company. By moving to electronic invoicing, these costs can be cut considerably mainly due to reduction in handling costs (Penttinen 2008). The following figure (Figure 3) depicts the performance indicators for e-invoicing process.

Figure 3. E-invoicing performance indicators.

4.3 E-payment impacts and performance indicators

The promise of e-payment has been that they ease payment and lower transaction cost (Southhard 2004). Cotteleer et al. (2007) listed top ten attributes based on perceived value for B2B e-payment users across multiple U.S. industries and found that transaction cost savings was considered most valuable. These cost savings come from improved processing efficiency – no more manual feeding of payment information is needed etc. Shorter payment cycles can also result in savings for the buyers in form of possible vendor discounts for early payments (Jolly 2007). The following figure depicts the performance indicators for electronic payment process.

Figure 4. E-payment performance indicators.
4.4 Links between the processes

So far, we have reviewed a group of indicators that can be more or less assigned to each function along the order-to-payment cycle. This is insufficient due to a couple of reasons. First, in real-life, the situation often is that operational level indicators are in active use but too often there is not a proper connection to business thinking. This leads to operations-driven improvement, rather than business-driven improvement – people in operations do not necessarily know the business impacts of their efforts. Also, some of the indicators used to evaluate process level performance can be included in cross-organizational service level agreements (SLAs) without actual knowledge of their economic effects. Second, it was already proposed that in addition to the individual sub-process level impacts, IT usage actually impacts the inter-process linkages as well. Therefore, we must strive to uncover the linkages between IT inputs on the grass-root level and the company level economic outputs as well as between the three functions. In developing the holistic evaluation tool (illustrated in the next section through the case example), we identified a number of links between the processes. Due to page limitations, these links will only be available as an appendix (Appendix 1).

5 CASE

To test the proposed evaluation tool, we conducted an in-depth case study at Marimekko Corporation, which is a leading Finnish textile and clothing design company that was established in 1951. The company designs, manufactures and markets high-quality clothing, interior decoration textiles, bags, and other accessories under the Marimekko brand, both in Finland and abroad (for further details see www.marimekko.fi). We conducted six face-to-face interviews with representatives from the company’s financial administration, purchasing, logistics and IT departments.

Using IT in financial administration is nothing totally new to Marimekko – they adopted e-banking systems and even electronic ordering already a few years ago. However, they have only quite recently adopted electronic invoice handling systems for incoming invoices. Approximately 16,000 – 17,000 invoices flow through their invoice management system (IM) every year. Although they have systems at place, they do not currently measure the impacts of IT usage in this context. However, they consider measuring important, particularly as basis for process development and as justification for the recent IT investments.

First, we decided to analyze the current state of processes; examining which phases exactly along the order-to-payment cycle are automated and which are still conducted manually. Further, we asked people from different functions to intuitively identify benefits that they consider important regarding process automation in the financial administration. Benefits they assigned for e-buying were mainly related to cost avoidance and time savings but also enhanced user satisfaction was reported. As for future challenges, enhanced asset utilization was considered most critical due to the fact that Marimekko currently holds substantially large inventories. Reported benefits regarding electronic invoice handling included centralization, decreasing circulation time and also reliability of electronic document exchange compared to regular mail. Perceived e-payment benefits were undisputed: reduced manual feeding leads to fewer errors, handling time reduction, and cost savings.

Finally, we asked people from different functions to comment our evaluation tool (regarding e-ordering, e-invoice handling and e-payment). It was very much a pleasure to see that during this exchange of thoughts, other important impacts were identified in addition to the ones that were first intuitively stated. The collected feedback regarding each measure was categorized (high, medium or low) according to expressed importance. Appendix 1 illustrates the evaluation tool from the point of view of Marimekko. We colour-coded the indicators so that green stands for highly important, yellow is medium, while red stands for low. After this, it was quite easy to draw “critical paths” for each sub-process. The ones marked with thick solid arrows illustrate the most critical mechanisms of value
creation. Dashed arrows show the paths that were considered slightly less critical yet worth considering.

E-payment systems in the case company are fully developed and well integrated into invoice management systems. Accounts payable sends a batch of payment proposals to the cashier and after that it takes only three clicks and approximately half a minute to approve and complete payments. Hence, there is limited potential for improvement in the payment section. That is why most of the boxes in payment are green and critical paths were not even drawn between them – there is hardly anything “critical” left to be improved in the payment side. Challenges at Marimekko relate more to the electronic integration aspects between payment and the other functions, particularly in how the order information could be better utilized in cash flow management and forecasting (illustrated by yellow boxes and the dashed path in payment).

It became clear that the yet untapped potential lies mainly on automating invoice handling and ordering even further. Critical paths in e-ordering include the above mentioned inventory level challenges and its effect on holding costs as well as asset utilization (turnover). One major challenge would be to automate inventory management even further to save buyers’ time. Additionally, reduced routine work enhanced user satisfaction. Another e-buying related issue is visualized by a dashed path: currently, buyers go manually through stock listings in order to identify purchase requirements and inevitably miss some requirements. This will of course lead to longer lead times and late deliveries to customers. If information systems would manage inventories automatically and alarm buyers when needed, human mistakes would decrease and thereby they would be able to deliver on-time to their customers. This could, through improved customer satisfaction, increase sales and thus create economic value for the company. The problem is, however, that these revenue side impacts are difficult to measure.

In invoice handling, circulation time was considered more critical than actual handling time and related costs. The largest obstacle was actually identified to be “approval delays” which refer to the time each invoice has to wait for approval at each person’s “desk”. Another related critical phase is getting purchased goods visible in stock as soon as possible. Standardization of documents was also considered highly important in order to avoid mistakes and needless rework. The same goes for system integration, where the question is how to standardize documents so that systems would understand each other leading to fewer mistakes and how to utilize the potential of information visibility in order to promote dialogue between purchasing and sales?

Part of the proposed indicators (such as inventory levels and some cost items) could be translated into relevant metrics directly; they could be derived from data already stored in company information systems. Others would require field studies, for example measuring the handling time spent on particular activities or user satisfaction. In fact, an important lesson learned from this case study is that the actual metrics have to be agreed upon according to individual characteristics of each case company.

All in all, it seems that in the case of Marimekko, impacts of automating the order-to-payment cycle relate closely to cost avoidance as intuitively stated. Yet contradictory to the intuition, a strong emphasis on asset utilization can be observed – better use of IT could enhance utilization of existing human resources and capital, affecting company profitability. They also leave the door open for revenue impacts with some reservations; the message was that these impacts are important yet difficult to observe and evaluate. However, it seems that the evaluation tool can be used in a real-life company context as an instrument for communicating, categorizing and sketching impacts and related measures. To be able to really measure and get accurate quantitative results, one would have to dig deeper to identify actual metrics that best indicate desired impacts and make efforts to collect relevant data.
6 DISCUSSION AND CONCLUSIONS

In this paper, we set out to build an evaluation tool to assess the impacts of the electronic order-to-payment cycle. We reviewed the extant literature and conducted expert interviews. Based on these, we built an evaluation tool that was divided into three sub-processes: e-ordering, e-invoicing, and e-payment. We then conducted a case study at Marimekko to illustrate how the proposed model works in real-life context.

In addition to case Marimekko, we tested the model in two supplementary case companies. Cost related effects were repeatedly emphasized which further supports the argument about impacts of automating order-to-payment cycles being cost-centric. However, the message was that once applications were implemented, challenges and post-auditing efforts related closely to better asset utilization. As a comparison, a retailing company reported that the share of e-orders in direct buying is already close to 100%, whereas current challenges relate to digitizing and centralizing indirect purchases. Concerning the e-invoice indicators, enhancing efficiency of the e-invoice “approval loop” was considered as a major challenge in all case companies.

The use of IT offers undisputed potential for process improvement in financial administration. Companies are making efforts in this area, yet progress is slow. Most companies have already well-established e-payment systems at place and electronic invoice management systems seem to be the current area of concern. Probably the largest unutilized potential lies in developing smart electronic applications for ordering. Despite of the potential, companies do not often monitor impacts of IT usage in this context. They rather implement, see what happens and never actually quantify the benefits. This might have undesirable effects – it is difficult for the companies to improve operations and further IT development investments might be impossible to justify. Measurement could thereby enhance the already great potential of process automation, yet it is rarely utilized. This contradiction acted as motivation for the whole study.

We developed an evaluation tool which strives to visualize mechanisms of value creation in the electronic order-to-payment cycle. The tool offers visibility over and across the entire business activity. As the case study showed, it works quite well at least as an instrument for cross-functional communication. Also, hierarchical categorization of metrics turned out to provide a useful skeletal structure for identifying critical cause-and-effect linkages between impacts. The purpose of building the model was to raise thoughts, promote awareness and provide a platform for productive dialogue. Based on the first company case, we can safely say that this objective was already fulfilled. Exact measurement would require extensive field studies within the organization.

References


A metrics system for the performance measurement of online distribution channels of multi-channel retailers

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0386.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>E-commerce (B2B / B2C / B2G / G2C), E-Tailing, Coordination, Empirical study</td>
</tr>
</tbody>
</table>
A METRICS SYSTEM FOR THE PERFORMANCE MEASUREMENT OF ONLINE DISTRIBUTION CHANNELS OF MULTI-CHANNEL RETAILERS

Hienerth, Claudia, Friedrich-Schiller-Universität Jena, Carl-Zeiß-Straße 3, 07743 Jena, Germany, claudia.hienerth@uni-jena.de
Meyer, David, Wirtschaftsuniversität Wien, Augasse 2–6, 1090 Wien, Austria, david.meyer@wu-wien.ac.at

Abstract

The Internet changed from a pure information media to a distribution channel, challenging companies in measuring the performance of their online distribution channels. Multi-channel retailers are particularly concerned since they need to coordinate their efforts with traditional, offline retailing activities. The paper at hand presents a corresponding metrics system derived from literature and a Delphi study, allowing for comparability with offline channels.

Keywords: Multi-channel retailing, Metrics system, E-commerce, Performance measurement.
1 INTRODUCTION

The development of the Internet from a pure information media to an increasingly important distribution channel resulted in significant intra- and inter-organizational changes for firms. Intensified by the collapse of the new economy by the turn of the millennium, stakeholders increasingly ask for performance measures tailored for online distribution channels, especially in the context of multi-channel retailers. Companies, however, are confronted with issues such as lack of experience and benchmarks, and missing or insufficient IT support. Challenges specific to multi-channel organizations include cannibalization effects and synergies as well as coordination tasks between online and offline channels (Schäffer & Weber & Freise 2002, Wall 2002, Welling & White 2006, Hienerth 2006).

There is already considerable work presenting models for e-commerce success (Hess 2001, Straub & Weber & Steinfield 2002). However, many of them are pure theoretic approaches whose applicability to real situations still needs to be evaluated. A closer look also reveals that some models either lack a precise definition of the research target, or focus on selected success factors such as web shop usability. Further, the supply chain and the corresponding business processes have changed with the increasing importance of online channels, thus limiting previous results valid for traditional organization models (Hansen 1997, Picot 1991, Gosh 1998). Hence, there is still a need for new metrics systems tailored to the requirements of specific business models (Schäffer & Weber 2001, Wall 2002).

Additionally, recent advances in IT multiplied Internet-based business models, fostering the development of new web/e-metrics such as Visits or Page Impressions, making the choice of meaningful, decision-relevant metrics even more difficult (Welling & White 2006). Also, these new metrics create a number of new technical challenges related to data collection and storage. Therefore, the relevance of these measures is still a matter of debate among researchers and practitioners alike (Schwickert & Wendt 2000, Marr & Neely 2001, Link & Schmidt 2001, Bhat & Bevans & Sengupta 2002, Palmer 2002, Nikolaeva 2005).

The paper at hand presents a metrics system specific to the performance measurement of online distribution channels of multi-channel retailers. The model in particular allows for comparability with traditional channels, supporting an integrated controlling system of a company's sales process. The model is based on existing literature on the one hand, and on the results of a Delphi study conducted on the other. The remainder of the paper is structured as follows: In Section 2, we first present design and results of the empirical study. In Section 3, these are combined with theoretical results from literature to a new model, which is then compared to metrics for stationary shops to demonstrate the comparability of both distribution channels. Section 4 concludes.

2 EMPIRICAL STUDY

2.1 Design of the Delphi study – Selection and number of participants

The two-stage Delphi study presented here was carried out using an e-mail based online questionnaire. First, we contacted 54 experts\(^1\) with a strong background in retailing and e-business, either conducting research or doing business in Austria or Germany. The choice was based on an extensive literature and internet review, as well as on recommendations from the experts themselves (Mullen 2003, Häder 2002). From these, 32 accepted to participate. In the run-up of stage 1, an extensive list with

\(^{1}\) We used Beckers (1974, p. 146) definition of expert.
performance metrics derived from the literature was provided to all participants. Five participants quit due to time restrictions, nine other claimed to be unfamiliar with the presented metrics.

In stage 1, the list of experts comprised 18 participants: six web shop managers, three researchers with background in e-commerce / retailing, four members of inter-trade organizations and certification authorities, and five executive consultants. Stage 2 involved 14 experts: four web shop managers, three researchers from the field of e-commerce / retailing, three members of inter-trade organizations and certification authorities, and four executive consultants.

2.2 Objectives of each Delphi stage

The objective of the stage 1 was to evaluate the importance of several metrics deduced from Palloks’ Model to measure the success of the online distribution channel. To this end, the experts evaluated 101 metrics within nine categories deduced from IS/e-business literature (especially from the main scholarly journals). Additionally, they were asked to supplement missing metrics and to point out metrics for measuring cannibalization and synergy effects between the web shop and the stationary shops. The first online questionnaire included 26 questions. The first five questions were intended to introduce the topic and to replace a qualitative preliminary study, which is recommended when applying a Delphi study. The remaining questions measured the importance of the listed metrics using a 5-point Likert Scale (1 = very unimportant, 2 = unimportant, 3 = neutral, 4 = important and 5 = very important).

The goal of stage 2 was to rank the most important metrics within the same categories from stage 1, defined as those with an average score of at least 4. All participants received the consolidated results from stage 1, summarized by mean and standard deviations. In total, the experts had to rank 42 metrics (within the different categories). In addition, the experts were asked to evaluate the importance of the metrics for measuring cannibalization and synergy effects between the web shop and the stationary shops, which they completed in the first stage. For this, again a 5-point Likert scale was used. Here again, we selected those with a score above 4, yielding five more metrics that were added to the system. The final model thus comprised 47 metrics in total.

2.3 Results

The 15 most important metrics according to the mean scores to evaluate the success of an online distribution channel for a multi-channel retailer are, in descending order:

1. system availability, 2. number of repeat customer, 3. product availability, 4. delivery time, 5. average server response time, 6. error rate, 7. online sales growth, 8. delivery quality, 9. return, 10. number of first buyers, 11. profit margin per customer, 12. acquisition cost per visitor, 13. acquisition cost per first buyer, 14. service level, and 15. order growth.

Clearly, from the experts’ point of view, the most relevant metrics for evaluating the success of the online sales process assess the performance of the information system (e.g., system availability) and goods handling/logistics activities (e.g., delivery time, product availability). This underlines the central importance of information systems in online selling. Also, the choice of metrics such as number of first buyers, number of the repeat buyers, and profit margin per customer shows that “classical” customer metrics from the offline-world are equally valued, particularly regarding the determination of the customer structure, customer retention and customer satisfaction. The same applies—with a grain of salt—to financial metrics such as total acquisition cost, acquisition cost per visitor, online sales, and return on sales.

We now turn to the 15 least important metrics, still according to the mean scores, and listed in ascending order:
42. online sales per daytime, 41. online sales per day of week, 40. online sales per employee, 39. day of week with the highest visit frequency, 38. time of day with highest visit frequency, 37. shop finder rate, 36. cross selling click rate, 35. product impressions, 34. focus, absence time, 33. online sales per payment method, 32. cross selling conversation rate, 31. no referrer, 30. number of clicks, and 29. one-click-rate.

Given the top metrics results, it seems surprising at first that the three most unimportant metrics (online sales per daytime, online sales per weekday, and online sales per employee) are also financial metrics. Obviously, compared to classical controlling figures, a more differentiated view is needed for online channels, taking into account existing structural differences. It is clear, for example, that input factors such as human resources will not have the same importance in online and offline channels.

Another unexpected result is the relatively minor importance attributed to web metrics (e.g., no referrer, focus, product impression, cross selling click rate, and cross selling conversation rate) for the success evaluation of the online distribution channel. Our results do, therefore, not confirm the work from other researchers (e.g., Weischedel et al. 2005) who attach strategic meaning to web metrics. This might, in fact, stem from technical issues such as collection and storage problems, or from a lack of guidance in interpreting these numbers (Novak & Hoffman 1996, Bensberg & Weiß 1999).

3 A NEW METRICS SYSTEM

3.1 Palloks’ model as a starting point

Our model is based on the hierarchical sales controlling system by Palloks (1995), consisting of 24 metrics organized in three main- and several subcategories: structural analysis (sales process, market), profitability analysis (performance of sales activities, efficiency of logistics processes, product profitability), and current state analysis. There are several reasons for choosing this model as a starting point: First, it matches our assumption that online sales activities can completely be represented and analyzed as a distribution channel on its own, a view supported by many authors (e.g., Vishwanath & Mulvin 2001, Chan & Pollard 2003, Hukemann 2004). Second, because of the model’s broad scope, it accounts for many influencing factors, opportunities, and risks influencing online sales activities, and thus supports management effectively and comprehensively (Steinfeld & Bouwmann & Adelaar 2002, Anderson & Srinivasan 2003, Wirtz & Schilke & Büttner 2003, Zentes & Schramm-Klein 2006). Third, our goal is to maintain the comparability to traditional sales channels as much as possible. Fourth, it complies with our definition of success as achieving distributional objectives (Zhuang & Lederer 2006, Wade & Nevo 2006).

3.2 Deduction and scope of the new model

The final model (see Figure 1) was deduced from Palloks’ model and the results of the empirical study. The starting model was modified to account for the process differences in online and offline retailing, as well as for the unequal use of input factors (e.g., staff and IT). The subcategory: “Efficiency of Logistics Process” was therefore split into “Handling & Logistics” and “Information System” to better account for IT-related factors, playing an essential role in online retailing. Further, the subcategory “Market” is extended to “Market & Customers”, customer-oriented figures being key especially in the start-up stage of online shops. As described above, the metrics were selected according to the Delphi ranking at stage 2, comprising the most important indicators in each category, as well as the five most important metrics for measuring cannibalization and synergy effects.

The final model comprises 47 quantitative and qualitative metrics. It can be employed by small- and medium-sized retailers as well as by web shop managers of larger multi-channel companies to measure the operational performance of their online sales activities. In addition, the model takes into account future developments of the web shop in the “Current State Analysis” category, supporting
strategic decisions for the development of the web shop and the company-wide marketing strategy. The system is generic and needs to be further adapted to the special conditions of a concrete firm implementing the model. It is limited “by definition” to multi-channel retailers, i.e. to companies offering tangible, movable goods in both stationary shops and a web shop.

3.3 Comparability with stationary shops

As mentioned above, it is challenging to directly compare online and offline channels by means of performance metrics due to significant differences of processes and input factor use. Such a comparison is, however, of vital importance to multi-channel retailers in order to assess cannibalization effects and synergies, and to optimize their overall marketing strategy (von Oelsitz 2006). In the following, we compare our metrics system to the integrated controlling model of Becker & Schütte (2004) for stationary retailers and discuss comparability issues.

We categorized the metrics into three groups according to the degree of comparability between online and offline channels, based on a comparison of their formulas and the required data sources. Table 1 lists metrics from both models that are identical or can directly be compared. Table 2 comprises performance measures that have similar semantics. Table 3 shows metrics without stationary counterpart. Note that in practice, the actual measurement of some metrics can be difficult due to technical or organizational issues. Also, most of the metrics listed are not uniquely defined, possibly causing biased values if different formulas are used for compared values.

As can be seen from the tables, most metrics (e.g., return, complaint rate) in fact allow direct comparison. Some others, more technical ones (e.g., system availability, down time) can be compared by analogy with corresponding metrics calculated for the ERP/POS-system used in the stationary shop. Only a few metrics (e.g., most frequent search term) have no matching counterpart in the stationary world. This supports the claim that, even though e-commerce has transformed the value chains, business principles have not changed after all (Hansen & Neumann 2005).

The comparison, however, raises another important issue regarding the implementation of the system: even though most metrics are identical conceptually, the required raw data in both channels often stem from different systems and data sources (e.g., log files versus database of ERP system). Differences in data quality and actuality might therefore cause additional variance in comparisons. It is therefore recommended to avoid redundant, external controlling systems, and to favour an integrated systems approach instead.

<table>
<thead>
<tr>
<th>Metric web shop</th>
<th>Metric stationary shops</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return on sales</td>
<td>stationery shops sales growth</td>
</tr>
<tr>
<td>reactivation rate</td>
<td>acquisition cost</td>
</tr>
<tr>
<td>market share</td>
<td>fulfilment cost in percentage of stationery shops sales</td>
</tr>
<tr>
<td>inventory turnover ratio</td>
<td>personnel cost in percentage of stationery sales</td>
</tr>
<tr>
<td>product availability</td>
<td>marketing costs in percentage of stationery shops sales</td>
</tr>
<tr>
<td>service level</td>
<td></td>
</tr>
<tr>
<td>delivery quality</td>
<td></td>
</tr>
<tr>
<td>liability failure rate</td>
<td></td>
</tr>
<tr>
<td>margin per product</td>
<td></td>
</tr>
<tr>
<td>complaint rate</td>
<td></td>
</tr>
<tr>
<td>margin per customer</td>
<td></td>
</tr>
<tr>
<td>market share growth</td>
<td></td>
</tr>
<tr>
<td>online sales growth</td>
<td></td>
</tr>
<tr>
<td>fulfilment costs in percentage of online sales</td>
<td>fulfilment cost in percentage of stationery shops sales</td>
</tr>
<tr>
<td>personnel cost in percentage of online sales</td>
<td>personnel cost in percentage of stationery sales</td>
</tr>
<tr>
<td>marketing costs in percentage of online sales</td>
<td>marketing cost in percentage of stationery shops sales</td>
</tr>
</tbody>
</table>
delivery time (to the customer) delivery time (to the stationary shops)
number of unique visitors (of the web shop) number of unique visitors (of the stationary shops)
number of first buyers
average order value per first buyer average purchase value per first buyer
average visit frequency
purchase / order frequency
order number development / growth purchase number development / growth
acquisition costs per web shop visitor acquisition costs per stationary shop visitor
acquisition cost per first buyer
repeat customer conversion rate
acquisition cost per repeat buyer
average orders per repeat buyer average purchases per repeat buyer
average order value per repeat customer average purchase value per repeat customer
return rate
basket to buy rate
customer overlap between the web shop and the stationary shops
turnover allocation between the web shop and stationary shops
web shop to stationary sales ratio
brand awareness

Table 1. Metrics with direct comparability to stationary shops.

<table>
<thead>
<tr>
<th>Metric web shop</th>
<th>Metric stationary shops</th>
</tr>
</thead>
<tbody>
<tr>
<td>system availability</td>
<td>system availability ERP/POS system</td>
</tr>
<tr>
<td>average response time</td>
<td>average response time ERP/POS system</td>
</tr>
<tr>
<td>error rate</td>
<td>error rate ERP/POS system</td>
</tr>
<tr>
<td>breakdown rate</td>
<td>breakdown rate ERP/POS system</td>
</tr>
<tr>
<td>breakdown intensity</td>
<td>breakdown intensity ERP/POS system</td>
</tr>
</tbody>
</table>

Table 2. Metrics with analog comparability to stationary shops.

<table>
<thead>
<tr>
<th>Metric web shop</th>
<th>Metric stationary shop</th>
</tr>
</thead>
<tbody>
<tr>
<td>most frequently used search word externally</td>
<td>-</td>
</tr>
<tr>
<td>development duration per product web page</td>
<td>-</td>
</tr>
<tr>
<td>average download time</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3. Metrics without comparability to stationary shops.

The following example should illustrate the notion of the comparability of both distribution channels and the meaning of this comparison for the management. Let us assume an average purchase value per repeat buyer of 45 Euros in the web shop and of 30 Euros in the stationary shops. This could indicate that web shop buyers are different (e.g., income, gender, ...) to the stationary shop visitors. Further, in this situation, the management should have a closer look at the market share growth. If, e.g., the web shop grows by 6 percent and stationary shops grow by 2 percent, the management should foster the online channel and accordingly allocate more resources to it. Moreover, it should investigate the consumers’ motivations for choosing the online distribution channel and adapt the corporate strategy accordingly. In any case, the different metrics employed should be monitored over time to capture important structural changes.

4 CONCLUSION

This paper presents a new metrics system suitable to measure the performance of online retailing systems. The metrics are based on literature and the results of a Delphi study. The system...
comprehends 47 quantitative as well as qualitative metrics. Despite the structural differences of online and offline retailing regarding the work flows and priorities of input factors, the system allows comparisons between web shops and stationary shops. Comparative analyses over time should reveal cannibalization and synergy effects, effectively supporting the management of both online and offline distribution channels in a coordinated way.

The presented metrics system is limited to multi-channel retailers offering movable goods to their customers. Also, the validity of the results is limited by the relatively small number of participants of the Delphi study. There is, however, no clear recommendation in the scientific literature on the minimum number of participants in Delphi studies. Also, the small number of participants (and especially their drop-off from stage 1 to stage 2) might indicate some lack of awareness among practitioners regarding the existing metrics, and the need for researchers to investigate and publish case studies where metrics are employed in practice.

Eventually, the efforts of investigating metrics systems for online and offline channels should result in the development of a comprehensive multi-channel controlling system. To this end, further investigation of existing interaction effects between the different channels is needed, especially with regard to the customers’ switching behaviour and the optimal allocation of marketing resources.
Figure 1. Metrics system for measuring the performance of online retailers.
References


# A DESIGN RESEARCH STUDY ON ENHANCING CREATIVITY – THE CASE OF DEVELOPING PRODUCT-SERVICE BUNDLES

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0544.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Creativity management, IT artifact, Knowledge Management Systems, Design research</td>
</tr>
</tbody>
</table>
A DESIGN RESEARCH STUDY ON ENHANCING CREATIVITY – THE CASE OF DEVELOPING PRODUCT-SERVICE BUNDLES

Müller-Wienbergen, Felix, European Research Center for Information Systems, University of Münster, Leonardo-Campus 3, 48149 Münster, Germany, felix.mueller-wienbergen@ercis.uni-muenster.de
Seidel, Stefan, Institute of Information Systems, University of Liechtenstein, Fürst-Franz-Josef-Strasse 21, 9490 Vaduz, Principality of Liechtenstein, stefan.seidel@hochschule.li
Müller, Oliver, European Research Center for Information Systems, University of Münster, Leonardo-Campus 3, 48149 Münster, Germany, oliver.mueller@ercis.uni-muenster.de
Knackstedt, Ralf, European Research Center for Information Systems, University of Münster, Leonardo-Campus 3, 48149 Münster, Germany, knackstedt.ralf@ercis.uni-muenster.de
Becker, Jörg, European Research Center for Information Systems, University of Münster, Leonardo-Campus 3, 48149 Münster, Germany, joerg.becker@ercis.uni-muenster.de

Abstract

As the sole production of goods is more and more becoming a commodity, many manufactures start bundling their physical products with related value-added services in order to differentiate their value propositions. The process of developing such customer-oriented combinations of products and services is an innovation process that requires the input of creative individuals. In this design-oriented paper we describe the conceptual design and prototypical implementation of an IT system that supports creative individuals performing such creative development efforts. The conceptual design builds on the provision of explicit knowledge by means of multi-perspective, hierarchical navigation structures. We argue that this concept (a) fosters information access that accords with the creative nature of a product development process and (b) furthering creative thinking through the provision of three distinct types of stimuli that actuates new entry points in a person’s cognitive network. The conceptual design of the system has its underpinnings in associative theories on creativity.

Keywords: Creativity Support, Product-Service Bundles, Information Retrieval, Navigation Hierarchies
INTRODUCTION AND RESEARCH DESIGN

Over the last decades, in most developed countries we have been witnessing a transition from a primarily goods-based to an increasingly service-based economy (Vargo & Lusch 2008). One explanation for the growing importance of services is the observation, that the sole production of physical goods is more and more becoming a commodity which can be almost equally provided by a constantly growing number of companies around the world (Rai & Sambamurthy 2006). At the same time, services are advancing as increasingly differentiated value propositions that are thought to lead to higher levels of customer satisfaction and loyalty (Howells 2003). Consequently, many traditional manufacturing companies try bundling their products with related value-added services to offer integrated customer solutions. Examples of such product-service bundles can be found in the automotive (e.g. automobile plus insurance, maintenance, trade-in, etc.) or telecommunication (e.g. mobile phone plus calling plan, messaging, music downloads, etc.) industry, but also in B2B markets like the machine tools industry (machine tool plus transport, integration, start-up, training, operating personnel, etc.).

The development of innovative product-service bundles comprises the discovery of prospective combinations of products and services. This innovation process can be characterized as a creative act. Creativity is the process of generating valuable products, services, processes, or ideas that are both novel and useful (DeGraff & Lawrence 2002, May 1959). This process results in innovations that are thought to effectively address a specific purpose in order to prosper (Amabile 1998). Against this background, product-service bundles can be characterized as innovations. A bundle’s value proposition must be new in order to stand out from competitors and it has to meet customer expectations or stimulate new demands. Accordingly, frameworks and models for product and service engineering (Scheuning & M. 1989) put creative acts like brainstorming or idea generation at the very beginning of a development process.

In order to support such creative processes, particularly two types of IT systems must be considered. First, research on so-called creativity support systems (CSS) has examined how computer systems can positively influence creativity (Massetti 1996). CSS implement different creativity techniques in order to provide guidance through the idea generating process (Malaga 2000). Second, theories on creativity identify a person’s knowledge as one major antecedent to the ability of being creative (Amabile 1998, Weisberg 1999). Knowledge management systems aid the creative process as they store organizational knowledge and support the creative individual in information retrieval (Shneiderman 2000). Yet, to our knowledge, there have been few efforts on integrating knowledge management and creativity support systems.

The purpose of this paper is to address the question of how to construct an IT system, which supports both creative thinking and information retrieval in the development process of a product-service bundle. The designed artefacts resulting from this design research process are a model and an instantiation (March & Smith 1995). We develop a model that explains how multi-perspective, hierarchical navigation structures can be applied to access digital information repositories in order to appropriately support creative persons in their effort of developing product-service bundles. Since models “that work ‘on paper’ will not necessarily work in real world contexts” (March & Smith 1995), we also present a prototype which is meant to serve as an evaluation object in a subsequent step of our research agenda demonstrating the feasibility of the proposed model. Our design rational is based on a theoretical underpinning from literature and theory. The design of the prototypical implementation incorporates results from an empirical study we conducted in the field of product-service bundle development. The focus of our research in its current state is to rather derive a sound design based on the existing body of knowledge in the IS field than to empirically validate the artefacts. However, our implementation is meant to provide the basis of consecutive evaluation steps.

The remainder of this paper is structured as follows: In the subsequent section we introduce a theoretical model that describes the conceptual basis of an IT system supporting the creative
development process of product-service bundles. To do so, we gradually develop the facets of our model from literature and theory (section 2). This is followed by the description of a prototypical implementation (section 3). In doing so, we exhibit the graphical user interface which reflects how the different aspects of the conceptual model expand into the prototype (section 3.1). Since hierarchical navigation structures play a prominent role in our approach, we subsequently apply the results of a survey on classification schema for product-service bundles to derive the navigation structures for the domain in focus (section 3.2). For the reason of presenting our design proposal in a more tangible manner, we present an example illustrating the intended value of our artifacts (section 3.3). The paper concludes with a discussion of contributions and limitations and an outlook to our future research agenda (section 4).

2 THEORETICAL MODEL

Creativity in an organizational context can be analysed on various levels. Woodman, Sawyer, and Griffin (1993) consider three levels of analysis: the individual, the group, and the organizational levels. Their model provides a comprehensive account on factors that impact creativity on these different levels and the interrelations that exist between them. The present study focuses creativity on an individual level. The theoretical model is meant to conceptualize how an IT system may aid a creative individual in solving her creative task.

Creative tasks are characterized by a high demand for flexibility and result in both deliverables and process orchestrations that are hard to predict and control in advance (Seidel & Adams & ter Hofstede & Rosemann 2007). For instance, the structure of the development process and the information needed are different whether a mobile phone is combined with a lifetime insurance against theft (e.g., gather relevant legal formalities, undertake risk assessment) compared to developing a bundle that valorises a mobile phone by providing a subscription for a music download service (e.g., contract negotiations with music labels, develop technical infrastructure). Thus, creative persons need a high level of autonomy in utilizing and controlling instruments they apply. We further argue that an approach towards creativity support must address several aspects of creative work. Our argument rests in the literature that identifies three primary components of creativity: motivation, expertise, and creative-thinking skills (Amabile 1998). While motivation is crucial to become creative at all, the way people approach problems and solutions (creative-thinking skills) and their knowledge about the domain of their work (expertise) also fundamentally contribute to the quality of a creative task’s output (Amabile 1998). Existing knowledge is a critical factor in the creative process; being creative often means “to put existing ideas together in new combinations” (Amabile 1998). For example, running shoes (product) that transfer one’s exercise performance online to a physician who analyses the data and elaborates a personal training schedule (service) constitute a highly creative product-service bundle. However, its parts are well-known. Thus, creative work comprises of both the convergent process of identifying relevant, existing ‘things’ (Weisberg 1999) and the divergent process of fusing these in novel ways (Seidel & Rosemann & Becker 2008b); “for a creative person to produce socially useful products, his or her divergent thinking must come hand in hand with convergent thinking” (Woodman et al. 1993, p. 299). IT systems may support this processes by providing its users additional expertise and furthering their creative-thinking skills.

In the remainder of this chapter we introduce an approach that aids both facets of creative work simultaneously: it provides explicit knowledge to support convergent thinking in a way that allows for the particular features of creative tasks and it provides means to foster ones creative-thinking skills in order to advance divergent thinking processes (Seidel & Müller-Wienbergen & Rosemann & Becker 2008a).

Throughout this work we rely on the assumption that knowledge positively correlates with the creative capacity of an individual. Although this is strongly supported by literature, there is literature which claims that existing knowledge can actually be a hindrance for creativity and innovation. The main aspect is often seen in the danger of biasing creative people by providing knowledge and thus limiting
their imagination (Cheung & Chau & Au 2008, Levitt & March 1988). This is the reason why we argue in favour of a combined provision of existing knowledge and means for stimulating creative thinking to prevent individuals from merely running down well-known alleys. Carlile and Rebentisch (2003) take a critical stance regarding the suitability of knowledge reuse in complex and novel situations. They argue, that if the circumstances surrounding the original knowledge development have changed, the knowledge is no longer relevant and its reuse can even be problematic. However, in creative problem solving situations the mere adoption of existing solutions never is an option. To be creative such a solution has to be novel by definition. Thus, reusing existing knowledge in creative tasks always involves its transfer into the new problem context. Carlile (2002) exposes that knowledge proves both a barrier to and a source of innovation in new product development. Its negative impact comes into effect when different communities of practice participate in an innovation process since boundaries between them hinder effective communication and cooperation. However, these boundaries neither exist within a single community of practice nor do they bear relevance for the individual level, which is the focus of the present study. Against this reasoning, we stick with Amabile (1998) by claiming that existing knowledge can fuel the creative process.

A database that provides a diverse set of information related to products and services establishes the basis of our approach. Such a database can broaden the expertise of persons involved in the development process of product-service bundles. For instance, checking the feasibility of a product-service combination often requires expert knowledge related to technical specifications of the product or legal formalities on the provision of services in individual countries. As the central concept of our model, we propose multi-perspective navigation structures as a means of accessing such digital information items. We argue that this approach to information retrieval is in favour of the convergent aspect of creative problem solving. First, hierarchical navigation structures support a creative problem solver with the explication of her information needs: Due to its creative nature, requirements for a product-service bundle are often rather vague. Thus, people involved in this creative process are often not capable of stating explicit search queries. Hierarchies can serve as an intuitive representation for the notions of abstraction and aggregation (Furnas & Zacks 1994). They can lead individuals along a stepwise refinement process in order to satisfy their information needs (Brelage 2006). Second, multiple navigation hierarchies that represent diverse perspectives on digital information provide access to the database by different users in different problem solving situations. The explicit knowledge that is considered relevant by a person as well as how this knowledge is accessed depends on both a person’s worldview and the specific situation’s context (Mey 1982, Polanyi 1975). Thus, information retrieval is about aligning the cognitive structures of a system’s users, its designers, and the information providers (Ingwersen 1992). Consequently, we argue that a single hierarchy is not sufficient to serve the often very heterogeneous (Markus & Majchrzak & Gasser 2002) group of possible information seekers (Furnas & Zacks 1994). Third, multi-perspective navigation means can provide a powerful device to expressing restrictions on the result set of an information retrieval process. As multiple hierarchical navigation structures classify the assigned items from different perspectives they may be applied to formulate multi-dimensional, set-theoretic constraints. In a system that provides the functionality of browsing the different navigation structures simultaneously, the user may refine constraints interactively. Such a dynamic approach of expressing information needs is in favour of the highly vital process of identifying relevant knowledge in creative problem solving situations. A similar approach is known from the area of business intelligence, namely the concept of online analytical processing (OLAP) (Pendse & Creeth 1995) which is applied to navigate through comprehensive sets of structured data.

In addition to providing appropriate means for the convergent aspect of creative work, we suggest to also aid the divergent facet of the development process for product-service bundles, so as to support a user’s creative-thinking skills. Associative theories explain the theoretical underpinning of our approach (Runco 2007). Mednick (1962) detected that truly creative ideas originate from remote associations within an person’s cognitive network. Likewise, Santanen and Briggs and de Vreede (2000) suggested that creativity is a function of the distance between the areas of an individual’s cognitive network that have been activated and combined to form a solution. According to Taylor
(1975), the areas of a cognitive network – mental states as he refers to them – are linked and via these connections the different mental states tend to activate each other provoking new ideas. The more remote the activated states are the more likely is the generation of a creative idea. In this context the concept of associative hierarchies becomes relevant. Such hierarchies denote the strength and organization of a person’s cognitive associations (Mednick 1962). Individuals with steep associative hierarchies tend to be less creative than people with rather flat associative hierarchies. For the former, it is harder to leave the very strong connections between their mental states and recall the more remote associations (Malaga 2000). Thus, an approach to supporting a person’s creative-thinking skills has to provide stimuli to break up inflexible associative hierarchies and to activate more remote mental states that otherwise will not be reachable via an individual’s associative connections.

Based on the above discussion, we propose an information retrieval system that is accessible via multi-perspective hierarchical navigation structures in order to provide stimuli. At this, a stimulus is not meant to provide immediate value by explicating a concrete example of re-arranging existing things and thus constituting a creative idea itself. It rather fuels the creative process of experimentation and exploration by activating a new entry point into a creative person’s cognitive network to initiate the process of going beyond the already known. The path this cognitive process takes is definitely outside the scope of an IT system. We argue that such a system is suitable to provide three primary types of stimuli (cf. Table 1). In the following we describe the different types in detail.

<table>
<thead>
<tr>
<th>Stimulus pattern</th>
<th>Stimulus source</th>
<th>Intended creativity impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-perspective stimulus</td>
<td>Navigation structure</td>
<td>Hierarchies provide a multi-level, disjoint categorization of database items. Starting from a category in focus, they give a glance at the range of possible alternatives at the current level of abstraction. These alternative options may serve as stimuli, as the user may not have considered all of them.</td>
</tr>
<tr>
<td>Inter-perspective stimulus</td>
<td>Navigation structure</td>
<td>Alternative navigation hierarchies represent alternative perspectives on the same facts. The explication of these diverse world-views may provoke new cognitive associations.</td>
</tr>
<tr>
<td>Content stimulus</td>
<td>Database content</td>
<td>Digital artefacts contained in an information repository exist in various formats addressing different senses. Various representation styles activate different creativity potentials and may in conjunction even better the recall of mental associations.</td>
</tr>
</tbody>
</table>

Table 1: Stimulus patterns

Intra-perspective stimulus

A navigation hierarchy can indicate alternatives that a creative person has not considered before. It represents a stepwise, disjunct categorization of the digital items contained in an information repository. All the navigation nodes on a selected hierarchy level that a system’s user does not choose along her navigation path to a specific item of interest may reveal alternative options, which lay outside her associative hierarchies. In becoming creative “everything is raw material.” (Coutu 2008) We refer to this stimulus pattern as intra-perspective stimulus.

Inter-perspective stimulus

“Within the mind of an individual, diversity enhances creativity.” (Amabile & Khaire 2008) The existence of multiple navigation hierarchies may motivate a system’s user to consider a different perspective onto the problem at hand. If a single repository item is accessible via discriminative navigation hierarchies, these hierarchies represent different classification schemas. Every schema may be understood as a different perspective on a repository’s content. Thus, the diverse navigation structures not only serve users with discriminative world-views as an appropriate means to access information items; the explication of these different world-views may also stimulate a user of such navigation hierarchies so as to consider different perspectives on her creative task. This way she is
empowered to identify additional entry points into her cognitive network. This stimulus pattern we call
inter-perspective stimulus.

Content stimulus

Apart from the navigation structures, the information items themselves can serve the purpose of
providing stimuli to a user. Digital artefacts contained in an information repository as sketched above
exist in various formats. Such different formats address different human senses. According to the dual
coding theory (Paivio & Lambert 1981) human memory and cognition are served by two separate
symbolic systems which are interconnected but also capable of functioning independently. One system
handles verbal information and the other system processes non-verbal information. Due to the linkage
between both systems, representations in one system can evoke associations to representations in the
other system. For instance, a picture can be named, and images may leap into someone’s mind while
reading a specific word (Malaga 2000). Experiments evidence that using lexical and pictorial stimuli
simultaneously amend one’s ability to recall an association from memory (Paivio 1983). Furthermore,
the separation between both systems implies that naming a concept and visualising the same concept
can stimulate different associations and thus provokes different ‘creative’ ideas. Moreover, different
individuals have different abilities in verbal and visual problem solving strategies (Malaga 2000).
Hence, different content formats activate different potentials in becoming creative. Against this
background, providing an ample set of contents in various formats and styles proves reasonable in
order to stimulate divergent thinking. These types of stimulus we refer to as content stimulus.

3 PRACTICAL INSTANTIATION

3.1 Graphical User Interface

Part of our research is the development of a prototypical instantiation that is ought to both proof the
theoretical model’s feasibility and serve as a tool for evaluation (Hevner & March & Park & Sudha
2004). As this research aims at designing an IT system that supports human creative performance, we
rather focus our descriptions on user interface than dwelling on internal data structures and algorithmic
issues. Thus, in the following we describe the graphical user interface and illustrate how the different
conceptual aspects of the above introduced model were implemented. The theoretically informed
model imposes several requirements on the user interface:

1) It has to present hierarchical navigation structures.
2) It must support the user in browsing several hierarchies in parallel.
3) Alternative navigation nodes have to be visible on every hierarchy level in order to provide intra-
perspective stimuli.
4) Alternative navigation hierarchies on a database item or navigation node have to be evident so as
to serve as inter-perspective stimuli.
5) Various content formats have to be pictured simultaneously catering for different user preferences.

We decided to not use classical tree structures as known from browsing file systems for the following
reasons. First, presenting a deep and broad hierarchy in a tree structure consumes much space on
screen. Especially if several hierarchies are presented simultaneously, this approach is not feasible.
Second, navigating along a tree structure by dynamically expanding and collapsing branches results in
constantly changing positions of the nodes and, thus, in a very restive navigation experience.
Therefore, we apply a navigational concept, which is to some extend leaned on the approach of
collapsible cylindrical trees as proposed by Dachselt & Ebert (2001). Child nodes in a hierarchy are
mapped onto rotating, three-dimensional cylinders while the path to the currently focused navigation
node is illustrated by means of a list above each cylinder (requirement 1) (cf. Figure 1). This approach
exhibits several advantages: The width of every hierarchy illustration is constant which results in an
easy segmentation and good utilization of the available screen space. Because hierarchies grow to the
bottom, several navigation structures may be placed side by side on one screen (requirement 2). Moreover, the appearance of the navigation structures is more static and thus easier to comprehend and less strenuous to use.

![Diagram of conceptual design of graphical user interface]

**Figure 1: Conceptual design of graphical user interface**

Although rotating cylinders intuitively display hierarchy levels of variable breadth on a screen area of constant size, they lack the ability to show all elements of a hierarchy level at the same time. Hence, they fall short of providing intra-perspective stimuli. In order to overcome this shortcoming, we implement the ability to expand a cylinder or path element to show all available alternatives at a corresponding hierarchy level (requirement 3). As a means to implement what we referred to as the inter-perspective stimulus, the interface highlights every content item or navigation node that is related to a navigation hierarchy different from the one in focus with a small badge. This badge expands to list all related navigation hierarchies, which may be chosen in order to change the navigation ‘perspective’ (requirement 4).

In order to select relevant content items from the database, the user can select multiple navigation nodes from different navigation hierarchies to work as a filter. The current filter expression is shown on the right of the navigation cylinders, while the related content items reside in the bottom area of the screen. The database items are represented by previews of their content and, thus, provide an immediate content stimulus (requirement 5). Furthermore, the system can retrieve remotely related content for every database item that is arranged around the item in focus, in order to function as another source of content stimuli. An algorithm determines the contents’ relatedness independently of the navigation structure but calls on content-based similarity metrics (Knackstedt & Kuropka & Müller & Polyvyanyy 2008).

### 3.2 Navigation Hierarchies

As indicated, we suggest hierarchical navigation structures as a device in order to stimulate creative thinking and to further information retrieval. To do so, such structures must provide means of information access that fit to the context they are applied in. In order to illustrate this, in the following we present the results of a survey on classification schema for product-service bundles that provides an empirical basis on which appropriate navigation structures for this particular domain could be developed. The underlying, highly iterative process of data collection was conducted as follows.
Quality | Quality properties (e.g. reliability, credibility) of the service | 1.59 | 0.925 |
Personnel | Qualification, experience and competences of provider personnel | 1.70 | 0.898 |
Availability | Availability in terms of time (e.g. 24 hours, 7 days a week) and location (e.g. only in Germany) | 1.73 | 1.025 |
Price | Price per unit of measurement (e.g. per use, per time period, flat) | 1.74 | 0.942 |
Function | Goal or purpose of the service | 1.76 | 0.949 |
Rights & duties | General rights and duties of the provider and consumer (e.g. confidentiality) | 1.79 | 0.990 |
References | References to past projects or customers | 1.79 | 0.944 |
Utility | Utility for the consumer (e.g. sales increase, cost cutting) | 1.85 | 1.105 |
Resources | Resources (e.g. tools, information) employed by the provider during service provisioning | 1.86 | 1.040 |
Ratings & rankings | Independent benchmarks or official certifications of the provider | 1.87 | 1.040 |
Service process | Sequence of activities executed during service provisioning | 1.92 | 1.100 |
Terms & conditions | Definition of contractual obligations (e.g. regarding payment or delivery) | 1.95 | 1.144 |
Product lifecycle | Reference to the lifecycle stage of the related physical product | 1.98 | 1.013 |
Industry sector | Industry sector for which the service is provided | 1.98 | 1.123 |
Customer input | Which resources (e.g. objects, information, personnel) has the customer to provide during the service provisioning process | 1.99 | 1.098 |
Frequency | Units and time intervals of service provisioning | 2.02 | 1.158 |
Contract duration | Duration of the service contract (e.g. one-time, subscription) | 2.17 | 1.270 |

Table 2: Empirical relevance of perspectives on product-related services

First, we conducted a broad review of the literature in order to identify perspectives for which navigation hierarchies could be identified. We concentrated our efforts on literature on conceptual models, ontologies, and standards of the services domain. The considered work covers both, a primarily business-oriented view (e.g., Baida 2006, DIN 2002) and a more technology-oriented view (e.g., Oaks & ter Hofstede & Edmond 2003, W3C 2004, W3C 2005) of the domain. We identified 37 initial perspectives (Knackstedt et al. 2008), which we consolidated into 17 central perspectives (cf. Table 2).

In a second step, over a period of six weeks in summer 2008 we conducted a telephone-aided questionnaire-based survey in the German mechanical and electrical engineering industry in order to empirically validate the relevance of the identified perspectives. Sales and marketing managers, responsible for cross-selling product-related services together with their core physical products, from 200 different companies (only one interviewee per company) participated in the study. The sample is nationally representative in terms of company size (number of employees), product technology, and industry subsectors. The participants were asked to rate the relevance (on a 5-point Likert scale from 1 (very relevant) to 5 (not relevant at all)) of the 17 consolidated perspectives when describing product-related services to their customers. All perspectives were attested in terms of relevance (cf. Table 2),
with quality (1.59), personnel (1.70), availability (1.73), price (1.74), and function (1.76) being the five most relevant.

In a third step, having validated the empirical relevance of the consolidated perspectives, we identified concrete navigation hierarchies for each perspective. In doing so, we relied on existing classification schemas (e.g. taxonomies) where possible. Table 3 lists an exemplary selection of such available classification schemas. For some perspectives several schemas could be easily identified. For example, the North American Product Classification Schema (NAPCS), the United Nations Standard Products and Services Code (UNSPSC), and the eCl@ss standard are promising candidates for a navigation hierarchy concerning the ‘function’ perspective. Likewise, there are numerous published lifecycle models that can be applied for navigating through the ‘product lifecycle’ perspective. Appropriate hierarchies for other perspectives, e.g. ‘quality’ or ‘price’, are dependent on the product or service at hand and have to be custom-made.

<table>
<thead>
<tr>
<th>Perspective</th>
<th>Publicly available classification schema</th>
<th>Excerpt from the content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality</td>
<td>SERVQUAL</td>
<td>Reliability/Assurance/Tangibles</td>
</tr>
<tr>
<td>Availability (locative)</td>
<td>ISO-Norm 3166</td>
<td>EU → DE → NW</td>
</tr>
<tr>
<td>Function</td>
<td>North American Product Classification System (NAPCS)</td>
<td>Services → Telecommunications → Messaging services</td>
</tr>
<tr>
<td>Resources</td>
<td>Classification of Resources (Seppänen &amp; Mäkinen 2007)</td>
<td>Resource → Legal → Copyrights</td>
</tr>
<tr>
<td>Service process</td>
<td>MIT Process Handbook</td>
<td>Make → Make-to-order → Manufacture and test</td>
</tr>
<tr>
<td>Terms &amp; conditions</td>
<td>INCOTERMS</td>
<td>Group C (Main Carriage Paid) → CIF (Cost, Insurance, and Freight)</td>
</tr>
<tr>
<td>Product lifecycle</td>
<td>VDI-Norm 2884</td>
<td>Before utilisation → Realisation → Procurement</td>
</tr>
<tr>
<td>Industry sector</td>
<td>North American Industry Classification System (NAICS)</td>
<td>52 Finance &amp; Insurance → 5222 Non-depository credit intermediation → 52221 Credit card issuing</td>
</tr>
<tr>
<td>Customer input</td>
<td>Classification of Resources</td>
<td>Resource → Legal → Copyrights</td>
</tr>
</tbody>
</table>

Legend: x/y: x and y are on the same hierarchy level; x → y: y is a sub-category of x

Table 3: Examples for publicly available classification schemas

In the following section we illustrate the application of these classification schemas to form the proposed multi-perspective, hierarchical navigation structures.

### 3.3 Example Case

We now demonstrate the presented creativity support approach using an example case. The portfolio manager of a global provider of enterprise communication solutions wants to develop a novel product-service bundle. The company so far only offers installation services when selling a telephone system. Browsing to installation services in an IT system realizing the presented approach, the portfolio manager notices that installation is classified as a ‘pre-use’ service in the ‘product lifecycle’ hierarchy. She realizes that there also is a ‘use’ and ‘after-use’ lifecycle stage (intra-perspective stimulus). She selects the ‘after-use’ stage and a collection of related documents appears on her screen. Since the result set contains more than thousand items, she sets a further filter by selecting ‘EU’ within the ‘availability’ hierarchy. Within the shrinked set of documents, she picks a newspaper article that is about the latest European regulations on disposal of electrical waste and the financial burdens that arise from these responsibilities. The article also elucidates novel approaches of some industries to recycle and re-market used equipment in order to generate additional revenues (content stimulus). By reading this, she comes up with the idea of developing a similar re-marketing strategy for her company. Customers should be motivated to return used telephone systems when they switch to a new
model. These will then be refurbished and sold again via a special web shop. The portfolio manager creates a service description for this new idea. When she adds the document to the system she is asked to classify it along several dimensions. One of the presented perspectives is ‘frequency’ of service provisioning which in her case is ‘one-time’ (inter-perspective stimulus). Getting this new perspective on her product-service bundle encourages her to think of possible services that are repeated on a subscription basis. From the frequency hierarchy she selects the ‘subscription’ category and discovers amongst others ‘remote diagnostics services’. The respective content screen offers a short description and diagram explaining that the service is completely virtualized and no extra hardware is involved (content stimulus). This evokes the idea in the portfolio manager’s mind that it could also be an option to completely virtualize a telephone system and merely handing IP-telephones to the customer. So, her company can substitute the error-prone and expensive telephone hardware by a software system that is centrally hosted.

4 CONCLUSION

Design research is concerned with “devising artifacts to attain goals” (Simon 1996). In this paper we presented a conceptual design and prototypical implementation of an IT system that supports creative work by both facilitating information retrieval and activating creative thinking in an integrated manner. We contribute to the IS body of knowledge by developing two artefacts: First, we developed a conceptual framework that is grounded in the existent literature and particularly draws from associative theories on creativity. The framework explains how multi-perspective navigation structures answer the aforementioned purposes. It can serve as an analytical and descriptive framework that can inform future research. Moreover, the framework provides a starting point for the development of new or the adaptation of existing information systems artefacts to support creativity. Second, we developed a prototypical implementation in order to evaluate the framework. We used the creative process of product-service bundle development to both elucidate and exemplify our design ratio.

Design research consists of the two basic processes of build and evaluate (March & Smith 1995). It must be noted that thus far, this research lacks evaluation that goes beyond a ‘proof of concept’. However, an evaluation in terms of a ‘proof by demonstration’ assessing the artefact in use is part of our future research agenda. For this purpose, a development process of a product-service bundle, as outlined in this paper, will provide the evaluation case. The developed system instantiation will be applied within this evaluation case to give evidence of the different stimuli types by analyzing the users’ interaction with the system while solving a creative task. In addition to supporting the evaluation of the theoretical framework, the IT system may become subject matter of assessment itself. Along with testing the system based on criteria such as efficiency, simplicity, ease-of-use, and elegance (March & Smith 1995), the impact on its users’ creative performance can be assessed. Firestien (1993) states that “the evaluation [of a creative product] must occur on a number of levels; not with a single factor, or a single total effective criterion score.” O’Quin and Besemer (1989) have developed a scale that allows to test whether a product is ‘creative’. It is called the Creative Product Semantic Scale (CPSS) and consists of three dimensions. These are novelty, resolution, and elaboration and synthesis. Particularly the first two dimensions (novelty and resolution) correspond to the understanding of creativity underlying this research that defines a product as being creative if it is original (novel) and if it is purposeful or appropriate (Seidel et al. 2008a).

Acknowledgements

This paper was written in the context of the research project ManKIP (Manangement of Creativity-Intensive Processes). The project is funded by the German Federal Ministry of Education and Research (BMBF) and by the European Social Fund of the European Union, promotional reference 01FM07061. We gratefully acknowledge the support of the Project Management Agency as part of the German Aerospace Center (PT-DLR).
References


From Data Warehouses to Transformation Hubs - A Conceptual Architecture

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0404.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Data Warehousing, Business Intelligence, Architecture, Information management</td>
</tr>
</tbody>
</table>
FROM DATA WAREHOUSES TO TRANSFORMATION HUBS –
A CONCEPTUAL ARCHITECTURE

Kemper, Hans-Georg, University of Stuttgart, Breitscheidstraße 2c, 70174 Stuttgart, Germany, kemper@wi.uni-stuttgart.de
Baars, Henning, University of Stuttgart, Breitscheidstraße 2c, 70174 Stuttgart, Germany, baars@wi.uni-stuttgart.de

Abstract

Originally, Data Warehouses (DWH) were conceived to be components for the data support of controlling and management. From early on, this brought along the need to cope with extensive data preparation, integration, and distribution requirements. In the growing infrastructures for managerial support (“Business Intelligence”), the DWH turned into a central data hub for decision support. As the business environment and the underlying technical infrastructures are fostering an ever increasing degree of systems integration, the DWH has been recognized to be a pivotal component for all sorts of data transformation and data integration operations. Nowadays, the DWH is supposed to process both managerial and operational data – it becomes a transformation hub (TH). This article delineates the relevant motives that drive the trend towards THs and the resulting requirements. The logical composition of a TH is developed based on data transformation steps. Two case studies exemplify the application of the resulting architecture.

Keywords: Data Warehousing, Data Transformation, Business Intelligence, Enterprise Systems.
1 INTRODUCTION

Business Intelligence (BI) denotes integrated approaches to decision support (Baars & Kemper 2008). In recent years, BI has been increasingly recognized as a pivotal subject for IT management (Gartner 2009). One of the subjects in the domain of BI that is currently been focused both by BI practitioners and BI researchers is how to extend BI infrastructures to tactical and operational levels and how to achieve a closer coupling of operational and BI systems. There is a puzzling variety of intertwined concepts that can be subsumed under this general theme: Business Activity Monitoring (DeFee & Harmon 2004), Operational Business Performance Management (Golfarelli & Rizzi & Cella 2004), Embedded BI (Hashmi 2004, Klawans 2008), Real Time Analytics (Raden 2003), Active and Real Time Data Warehousing (Akbay 2006, Brobst 2002, Raden 2003), and Operational BI (Chemburkar & Keny 2007, Marjanovic 2007). Each of those terms represents a different facet of the general trend to either utilize BI infrastructures for the support of operational decisions or to further bind operational and managerial systems together (hence the term “operational BI” (Eckerson 2007)). There is also one commonality on the architectural side: They are all built upon core components for the integration and exchange of data (White 2005).

The established centre for data integration, storage, and exchange in BI environments is traditionally the Data Warehouse (DWH). Inmon defines a DWH as “...a subject-oriented, integrated, time-variant and nonvolatile collection of data in support of management’s decision-making processes...” (Inmon 2005). Although this conception still captures the heart of most DWH installations, it becomes somewhat diluted when the DWH is utilized in operational environments. While still fulfilling the tasks known from traditional management support, the DWH additionally needs functionality for storing and exchanging real-time transactional data. To reflect the changed role of such “enhanced DWHs” in the enterprise, this paper introduces the term “Transformation Hub” (TH). Here, a TH is understood to be a logically central component that concentrates functions for data integration, enrichment, and exchange. It is designed to serve for managerial, analytical and operational applications alike.

As the TH differs in content and structure from classical DWHs, the underlying models, structures, and approaches also need to be modified and be rearranged under consideration of state-of-the art components and architectural designs. This paper addresses the respective issues. It specifically focuses on the derivation of a conceptual architecture of a TH which captures relevant functional components and their interplay from an application oriented view. About 20 years ago, IT adopted the term “architecture” and applied it to almost every structural aspect of hard- and software systems – including DWHs (Hammergren 1996).

As this paper takes a more business oriented perspective, it concentrates on a conceptual system structure. The resulting architecture is designed to act as a starting point for the delineation and design of system components and their interplay solely based on application needs. This differentiation is of particular importance in the realm of current TH infrastructures where the essential logical design easily comes out of focus as it is superimposed by more realization-driven design layers, e.g. for the data exchange and middleware architecture, for data-feed-approaches (i.e. data consolidation, federation, or propagation) (Brobst 2002, White 2005), for event handling mechanisms (push vs. pull) (Brobst 2006), etc. Notwithstanding the importance of those aspects, they are secondary in nature in the conceptual design which runs down to selecting functional components for defined business contexts and to specifying how they are supposed to work together.

The objective of this paper is to derive a conceptual architecture for a Transformation Hub. The course of the discussion is as follows: After relevant concepts in the vicinity of the TH are put into context, its actual role is pinpointed by distinguishing and discussing motives for the implementation of THs. This is the foundation for a derivation of the architecture which is mostly based on logical steps of data transformation as known from the data warehousing domain. For each of the different transformation
steps, the peculiarities of the TH are discussed and matched with the concepts introduced before. The application and the relevance of the framework are illustrated with two case studies.

2 RELATED CONCEPTS

There is a long tail of research on the architecture design of a Data Warehouse with some influential contributions coming from Inmon (Inmon 2005) and Kimball (Kimball & Ross 2002). The former stands for a strongly centralized, application-independent approach while the latter proposes a more decentralized data management that is bound together semantically by the use of shared dimensions (“dimensional bus”). Empirical research shows that both approaches can be found in similar numbers in current enterprises and that both are to be preferred over less-coordinated approaches (Ariyachandra & Watson 2006). Besides this, a distinction needs to be made between application-specific data excerpts – usually known as Data Marts – and a central and logically integrated data storage – the “Core Data Warehouse” (Inmon 2005).

Providing near time, transactional data is one of the most significant modifications to classical core Data Warehouses (Inmon 1999). To handle the different access profiles, reliability requirements, and update time-frames, dedicated components have been proposed which have become known as “Operational Data Stores” (ODS). Inmon introduced the concept of the “Information Factory” for ODS-enhanced DWHs (Inmon & Imhoff & Sousa 1997, Kelley & Moss 2007). Being a more technical and data storage oriented concept, the literature on Information Factories can serve as an outline for the realization options of the data management components of a TH.

There are several contradicting definitions for the term “ODS” (Inmon 1999, Sherman 2005) and not all of them are suited for the TH approach, e.g. an ODS as a mere replication of operational data base tables. In this paper, a ODS is conceptually understood to be a component that provides integrated data (Kelley et al. 2007) and for this purpose enables bringing together transactional data from multiple sources. This involves data cleansing and integrity checking. Here, the ODS, the DWH, and the Data Marts are conceived as the main data provision roles that a TH has to incorporate.

ODS-enhanced DWH architectures allow building Closed-loop and Active Data Warehousing solutions. In Closed-loop Data Warehousing, results from analytical processes are directly fed back into DWHs or operational systems (Brobst 2002). Active DWH systems automatically trigger actions based on defined data constellations. As the respective application scenarios frequently go along with the need for current data, “Active and Real Time Data Warehousing” is often combined to a fixed phrase (Akbay 2006, Raden 2003). In Closed-loop and Active Data Warehousing the DWH can turn into a data source for the operational systems.

From an application oriented standpoint, a DWH is by its very nature a component for a data driven application integration. Recently, however, some scenarios are proposed that leave this data centric approach: It has been suggested to complement a DWH with components that focus on providing data processing functionality rather than actual data. These components are usually conceived to be based on the paradigm of service oriented architectures (SOA). The name that has been coined for this approach is Embedded BI. It is “embedded” as the BI functions are conceived to be seamlessly integrated into operational systems, working directly against the local, transactional data, rather than against the integrated repository of a DWH (Hashmi 2004, Klawans 2008).

The above discussed components abstract from the applications they actually support. One of those with immediate relevance for a TH is Business Activity Monitoring (BAM). BAM is built upon the idea of providing software for the near-time monitoring of the status and the results of business processes. For this purpose, data from the involved application systems needs to be extracted, integrated and presented in a meaningful way (DeFee et al. 2004, Golfarelli et al. 2004, Melchert & Winter 2004). BAM heavily focuses on the data presentation aspect, especially in the form of “dashboards” and “cockpits” and can be applied both within the confines of an organization or across enterprise borders (Eckerson 2006), e.g. for logistics or production processes.
The idea of BAM is closely interlinked with Business Process Management. Business Process Management highlights the integrated management of business processes. In this regard, BAM can be understood as a tool that supports a subset of the tasks for Business Process Management (DeFee et al. 2004, Golfarelli et al. 2004, Melchert et al. 2004, Verner 2004). Further reach full-fledged “Business Performance Management” approaches which aim at the integrated, strategy-oriented steering of an entire organization based on consistent indicator systems for all managerial levels (Eckerson 2006, Golfarelli et al. 2004). By interlinking the complete reporting hierarchy, they reach well beyond Business Process Management, although the latter can be smoothly embedded within the former. If realized that way, Business Process Management can be understood as Operational Business Performance Management (oBPM).

3 APPLICATION DOMAINS FOR TRANSFORMATION HUBS

A closer look at the literature reveals a variety of interdependent motives that drive the evolution from the classical DWH to the TH:

Integrated management concepts, especially within the realm of Business Performance Management as discussed above. Approaches like the “Balanced Scorecard” (Kaplan & Norton 1996) or “Value Based Management” (Grant 2003, Rappaport 1998) are built upon the idea that a system of interdependent key performance indicators (KPIs) can facilitate the consistent steering of an entire organization. A consequence of pursuing such approaches is the need for an operational, process-level indicator gathering and communication, especially based on Business Process Management solutions. This entails the need to closely couple operational, tactical, and strategic decision support and to provide a consistent data socket (Eckerson 2006, Golfarelli et al. 2004).

Information logistics. With the diffusion of integrated applications, e.g. for Customer Relationship Management (CRM) or Supply Chain Management (SCM), and the demand for application-spanning services like security-checks, pro-active fraud-detection, or on-site localization of service issues, the potential of Business Process Management and BAM solutions becomes apparent: respective solutions require infrastructures for data exchange, integration, harmonization, and distribution (Furness 2004, Chemburkar et al. 2006, Nguyen 2005, Raden 2003, Stefanovic & Radenkovic & Stefanovic 2007, Watson 2005, Baars & Kemper & Lasi & Siegel 2007). As THs are innately designed for mass data processing, they are ideal building blocks for large-scale integration solutions. In contrast to the integrated concepts discussed before, these approaches are focusing more on data exchange along processes (Bucher & Dinter 2008) than across managerial levels. Closed-loop and Active Data Warehousing are often following information logistics initiatives, as they draw additional value from the achieved data integration.

Analytical access for lower managerial levels. In a turbulent business environment there is value in an analytical access to historical data even for lower managerial levels – which can be provided more conveniently with flexible, BI based reporting and analysis tools (Klawans 2008, Marjanovic 2007).

Centralization and utilization. A powerful rationale for utilizing a TH in an operational context is the centralization of data transformation: The TH is identified as a centre for all kinds of mass data processing activities. A particular concept that can be subsumed under this rationale is Embedded BI. This efficiency based argumentation not only permeates all motives discussed above. It also transcends it, and some of the supported applications can hardly be considered to be “Management Support” or even “Decision Support”. The concentration of all administration, monitoring, and resource related tasks for data transformation facilitates unlocking economies of scale and utilizing learning curves that are further fostered by dedicated organizational units. Such BI competence centres have already become widespread in the realm of BI (Unger & Kemper & Russland 2008).

Figure 1 illustrates the role of the TH in this conglomerate of applications: It connects analytical and operational systems, feeds BAM and/or Business Process Management solutions with near time data, extracts and harmonizes data for information logistics purposes, binds together key performance
indicators based on Business Performance Management concepts, and provides transformation functionality for further applications.

Figure 1: Role of the Transformation Hub

4 THE TRANSFORMATION HUB - ARCHITECTURE

The TH architecture proposed in the following section is designed to support all application settings discussed in the above sections. It is conceived to be vendor-neutral and invariant to physical design choices in order to allow for a sustainable mapping not only of already realized and projected solutions but also of possible development trajectories. It enhances a conceptual ODS enhanced DWH framework that has been iteratively developed over a course of over a decade and that already incorporates a large body of study results (Baars & Kemper 2008, Kemper 2000, Kemper & Finger 2006). Figure 2 depicts the architecture. It consists of the following main components:

- The transformation components (Filtering, Harmonizing, Enrichment, and Aggregation) form the heart of the TH. All application domains discussed above are aiming at some kind of data processing – be it a simple integration operations for monitoring tasks or complex enrichment and harmonization procedures for Business Performance Management.
- Associated repositories for data storage and data access play the roles of the ODS, a Data Warehouse and/or Data Marts. They can be logically differentiated by the respective subsets of transformation steps the data needs to have undergone.
- Interfaces for service provision, i.e. for embedded BI solutions. In fact, the whole range of transformation functionality can be made available for third applications.
- Administration interfaces which allow for a secure and documented access to relevant data for the user, support the configuration and monitoring of the operational data upload, and provide documentation.
- Meta data management that delivers the contextual glue that binds all involved components and contents together and ensures both an efficient technical maintenance as well as a consistent usage of the TH contents. It addresses both technical meta data and semantic meta data.

The next paragraphs discuss the fundamental transformation steps and the associated data access components and match them with the related concepts from section 2. The following main types of transformational tasks are distinguished (Kemper 2000):

- Filtering encompasses the data extraction and the correction of defects in syntax and semantics.
- Harmonizing is the process of granularity adaptation and merging data to defined subject areas.
- Aggregation addresses the summary of data to predetermined levels of detail
- Enrichment adds calculated indicators to the data.
4.1 Filtering and pre-access components

The filtering process starts with the extraction of data from operational data repositories. From a conceptual viewpoint it is imperative to check the quality of the operational data sources (systems, fields) that are considered as inputs for the TH (English 1999).

Regarding the actual extraction process, there are basically two basic options (Akbay 2006):

- Periodical updates that adhere to a classical Extract-Transform-Load (ETL) logic or
- Continuous, real-time streams (“trickle feed”) that directly mirror changes in the operational systems.

The two approaches need to be distinguished conceptually because of the consequences regarding timeliness, costs, performance impact on the operational systems (Brobst & Morrey 2002), and data quality (handling incomplete/wrong transaction data, time buffer for quality checks). As the TH is supposed to feed both real time applications like BAM and data analysis applications based on historic data, components for both might be necessary (Brobst & Rarey 2002).

When actually conducting the extraction, a variety of syntactic and semantic incompatibilities or defects needs to be taken care off. In a near-time setting, the options to tackle semantic errors that cannot be automatically corrected or not even be detected are naturally limited. A way to narrow down the respective error handling times is to implement tightly defined work flows which ensure a minimal timeliness of the respective activities and that might be tracked and controlled by a workflow management system (Bartel & Schwartz & Strasser 2000).

For applications that rely on near-time, local data, the results of the extraction and the automated error corrections might already be relevant. In this paper, components to read out the respective intermediate data are labelled “Pre-Access” as the data is still not consistent across different sources.

4.2 Harmonizing, the ODS, and BAM support

Systems supporting operational processes rely on data that normally differs in granularities, definitions, time periods, etc. (Kemper 2000). Clearly, harmonization is not a mere technical issue but
rather a conceptually challenging task which needs business support (Berg & Heagele 1997). Bringing heterogeneous data together is one of most fundamental tasks of a TH.

The respective merging operations pertain to both syntactical and semantic inconsistencies:

- **Syntactic harmonization** means the coding of keys and attributes on the one hand and the handling of homonyms (attributes which carry identical names but have different meanings) and synonyms (attributes which are referred to by different names but have the same meaning) on the other. Usually this type of issues can be managed on a basis of defined routines and is of low criticality.

- More serious are cases of **semantic harmonization**. A common cause for the occurrence of such predicaments is of historical nature: In many enterprises a plethora of homonymous business terms is used, each with varying local definitions and connotations. Seemingly clear time-periods differ, and so do the definitions for performance indicators like sales, revenue, profit etc. Merging data together in a TH demands the harmonization of these terms across unit borders (English 1999). This type of harmonization activities can easily result in severe political and cultural problems (Kemper 2000) and needs to be backed by a tight meta-data management.

Providing filtered and harmonized data is a core feature of a TH. Following the definition in this paper, a data access system that ensures these two layers in near-time and on transactional level is an ODS. This kind of data provision also forms the heart of BAM- and oBPM applications.

### 4.3 Aggregation, enrichment, Data Marts, Data Warehouses, and BI support

On the third layer of transformation, the filtered and harmonized data is further refined by implementing hierarchy structures and calculating business indicators.

The hierarchy structures of the *aggregation* extend the idea to centralize the definition of semantics to the specification of pathways along which granular data should be analyzed (e.g. from store to country to world region). This can include parallel dimensions: The dimension “product group” can for instance be alternatively summarized over product categories (customer based) or profit centres (organization based) – both hierarchies stand for valid analysis corridors.

A centralized calculation of indicators (enrichment) guarantees the consistence of the business terms on the basis of homogenous definitions for entire fields of applications. A DWHs can and Data Marts do include aggregated and enriched data. This type of enrichment violates the paradigm of a separation between logic and data: There is a deliberate built-up of redundancy for purposes of performance and usability. As Data Marts are understood to be application driven, a simple and immediate provision of aggregated data is in fact often one of the main strengths of the Data Mart.

By incorporating the roles of the DWH and the Data Marts, the TH also takes over the function of the focal component for the support of a BI approach.

The powerful features for mass-data enrichment make the TH also interesting for a variety of operational systems: This functionality can be just as well utilized for operational data that does not require any form of harmonization, e.g. in a shop floor environment. Some of those applications lie well beyond the confines of even the widest conceptualization of BI.

### 4.4 Mapping transformation steps and application domains

Coming back to the application domains from section 3, their concrete relationship with the TH and with the transformation steps can be laid out (cf. Table 1).

**Integrated management concepts** aim at a vertical integration with a consistent set of indicators. This requires rigid data harmonization as well as a logically centralized indicator definition, calculation, and aggregation. It can thereby be concluded that special attention needs to be devoted to a meticulous definition and maintenance of meta data.
Information logistics driven applications are built on data exchange and harmonization. By the very nature of those approaches, they need to weave together multiple heterogeneous systems and therefore require strong extraction interfaces and respective filtering components.

When it comes to analytical access on operational levels, a Data Mart based access to historical data is needed that is seamlessly interlinked with current (near-time) data.

The centralization/utilization motive captures a whole range of applications. In essence, this demands for a flexible usability of the different components and their functions. The resulting requirement is that the architecture is defined in a modular fashion with building blocks that at best can be immediately used as self-contained, service-oriented units.

<table>
<thead>
<tr>
<th>Integrated management concepts</th>
<th>Information logistics</th>
<th>Operational analytics</th>
<th>Centralization / utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus</td>
<td>Data exchange and harmonization</td>
<td>Data aggregation and analysis for lower mgmt. levels</td>
<td>Consolidating data processing and analysis tasks</td>
</tr>
<tr>
<td>Examples</td>
<td>Value-Driven Trees, Balanced Scorecard</td>
<td>Supply Chain Management, BAM in the realm of manufacturing</td>
<td>Support of decisions regarding replenishment</td>
</tr>
<tr>
<td>Relevant components</td>
<td>Harmonization, aggregation, and enrichment, Meta data and meta data administration</td>
<td>Extraction, Filtering, Harmonizing</td>
<td>Aggregation (over time), Data mart based access to historical data</td>
</tr>
<tr>
<td>Particular requirements</td>
<td>Functionality for shared meta data management</td>
<td>Multiple interfaces and flexible extraction routines</td>
<td>Options to interlink low-aggregated transactional and historical, aggregated data</td>
</tr>
</tbody>
</table>

Table 1: Applications and Transformation Hub requirements

5 CASE STUDIES

Two cases are presented to illustrate the ongoing trend towards Transformation Hubs and to highlight relevant discussion areas. In both cases research projects have been conducted jointly with the respective organization. This allowed for deep insights into the solutions and approaches. Data gathering methods have primarily been workshops and qualitative document analysis.

5.1 Stock Exchange

The first case highlights the approach of a former national stock exchange that has over time turned into one of the largest stock exchange organizations in the world. By the end of the last century, the company not only faced new environmental conditions like increased dynamics in the verge of internet-based trading but also increasing requests of internal and external stakeholders for precise, trustworthy and real time decision support content. The developed BI solution is a real time TH. It acts as a hub for information distribution to a worldwide financial community and permits internal and external users a prompt analysis of market related mass data.

The solution of the stock exchange delivers both classical DWH services (aggregated, historical data for management analysis) as well as direct feeds into operational systems which are relying on harmonized financial data. The latter part is heavily information logistics driven.
Depending on the structure of the source systems and the time variance of the data, internal and external source systems are connected either in real-time or via periodical updates. The data transformation is completely meta data driven. Automatic filtering and harmonization are realized in a separate “validation layer” while unstructured data and semantic defects are dealt with based on a workflow approach. The (approved) results are gathered in a container with the label “single point of truth”. Eventually, there is a dedicated “analytical business layer” which is in fact a manifestation of the aggregation and enrichment layers discussed here. The results – denoted as “information products” – are delivered in a multitude of forms, including data streams to operational downstream systems, cubes prepared for data analysis, website content, reports, spreadsheet files, or even text messages.

![Diagram](image)

**Figure 3: The TH of Case 1 (under consideration of Detemple & Feidieker & Münch 2006)**

The mapping of the solution with the developed framework is visualized in Figure 3. Noteworthy are the systematic integration of manual activities and the consequently layered approach.

### 5.2 Manufacturing

In the second case, the move from classical BI to a full-fledged TH infrastructure is still in its conceptual phase, although it clearly shows the centralization forces at work.

The respective (large) manufacturing enterprise has already enforced a physically centralized BI approach, based on one obligatory data warehousing product that is used for all sorts of data refinement tasks. The results are used by a multitude of systems, among them diverse reporting solutions. Due to the historical development, a variety of independent Data Mart based solutions has grown that is not yet logically integrated: The solutions run on separated instances of the software and are individually customized. Over time, the solutions grew in number, data volume, and business relevance. Furthermore, multiple interdependencies between the different systems became apparent that were sources for blatant redundancies. To address this situation, the company has set up a large-scale consolidation initiative that addresses both inconsistencies between the diverse data repositories as well as the respective redundancies in the data transformation activities.

Among the various solutions in place quite a few are not directed at classical managerial support. Those applications often utilize just a subset of the functionality of the DWH solution, e.g. only the data extraction from the ERP system or the (efficient) data enrichment features.
Especially from the side manufacturing and logistics, first near time data transformations have been implemented – predominantly aiming at enrichment. This type of application is actually prone for an Embedded BI approach: Only local, unprocessed data is enriched and the data storage on the system is just temporary – the main strengths of the classical DWH environment are underused.

The logical structure of the TH is depicted in Figure 4.

![Figure 4: The TH of Case 2](image)

Currently it is investigated which types of solutions can indeed be embedded coherently in the conceived TH, which ones need to be handled with different (but logically integrated) components and which ones should be utterly separated into distinct systems for either technical or economical reasons. This development goes along with the implementation of an overall governance approach that safeguards the efficient and strategy-conform usage of the systems.

6 DISCUSSION AND OUTLOOK

The cases illustrate the development to extend classical Data Warehouses towards an integrated Transformation Hub. They also highlight the need to couple technical and organizational design – by thoroughly defining, embedding, and supporting processes for manual data transformations on the one hand and by purposefully implement overarching governance concepts to ensure the efficiency and adequacy of the solution on the other. A relevant best practice of the first case is the strict layer based approach that follows the transformation logic.

Clearly, additional research is necessary to unravel the interplay of the discussed conceptual architecture and lower level structures. Of particular interest should be research on the role of SOA in TH environments. Besides, further empirical research is needed that sheds light on the relative significance of the discussed developments and how they evolve in concert.

Conceptually, the discussed trends entail the necessity to redefine the future role of Business Intelligence – and subsequently for the organizational units that support them as well: the metamorphosis towards mission critical Transformation Hubs reaches well beyond issues of technology design.
References


RECONSTRUCTING THE GIANT: ON THE IMPORTANCE OF RIGOUR IN DOCUMENTING THE LITERATURE SEARCH PROCESS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0566.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Literature review, Research methodology, IS Journals, IS Community</td>
</tr>
</tbody>
</table>
RECONSTRUCTING THE GIANT: 
ON THE IMPORTANCE OF RIGOUR IN DOCUMENTING 
THE LITERATURE SEARCH PROCESS

vom Brocke, Jan, Martin Hilti Chair of Business Process Management, Institute of Information Systems, University of Liechtenstein, Fuerst-Franz-Josef Strasse 21, 9490 Vaduz, Principality of Liechtenstein, jan.vom.brocke@hochschule.li

Simons, Alexander, Martin Hilti Chair of Business Process Management, Institute of Information Systems, University of Liechtenstein, Fuerst-Franz-Josef Strasse 21, 9490 Vaduz, Principality of Liechtenstein, alexander.simons@hochschule.li

Niehaves, Björn, European Research Center for Information Systems, University of Muenster, Leonardo-Campus 3, 48149 Muenster, Germany, bjoern.niehaves@ercis.uni-muenster.de

Riemer, Kai, European Research Center for Information Systems, University of Muenster, Leonardo-Campus 3, 48149 Muenster, Germany, kai.riemer@ercis.uni-muenster.de

Plattfaut, Ralf, European Research Center for Information Systems, University of Muenster, Leonardo-Campus 3, 48149 Muenster, Germany, ralf.plattfaut@ercis.uni-muenster.de

Cleven, Anne, Institute of Information Management, University of St.Gallen, Mueller-Friedberg-Strasse 8, 9000 St.Gallen, Switzerland, anne.cleven@unisg.ch

Abstract

Science is a cumulative endeavour as new knowledge is often created in the process of interpreting and combining existing knowledge. This is why literature reviews have long played a decisive role in scholarship. The quality of literature reviews is particularly determined by the literature search process. As Sir Isaac Newton eminently put it: “If I can see further, it is because I am standing on the shoulders of giants.” Drawing on this metaphor, the goal of writing a literature review is to reconstruct the giant of accumulated knowledge in a specific domain. And in doing so, a literature search represents the fundamental first step that makes up the giant’s skeleton and largely determines its reconstruction in the subsequent literature analysis. In this paper, we argue that the process of searching the literature must be comprehensibly described. Only then can readers assess the exhaustiveness of the review and other scholars in the field can more confidently (re)use the results in their own research. We set out to explore the methodological rigour of literature review articles published in ten major information systems (IS) journals and show that many of these reviews do not thoroughly document the process of literature search. The results drawn from our analysis lead us to call for more rigour in documenting the literature search process and to present guidelines for crafting a literature review and search in the IS domain.

Keywords: literature review, literature search, rigour, IS journals, IS community, research methodology
1 INTRODUCTION

Reviewing the literature has been proposed as a scientific profession a long time ago (Garfield 1977; as cited in Garfield 1987, p. 113), since it represents an “essential first step and foundation when undertaking a research project” (Baker 2000, p. 219). A literature review seeks to uncover the sources relevant to a topic under study and, thus, makes a vital contribution to the relevance and rigour of research: On the one hand, relevance is improved by avoiding the reinvestigation of what is already known (cf. Baker 2000, p. 219). On the other hand, rigour is derived from an effective use of the existing knowledge base (cf. Hevner et al. 2004, p. 88). Hence, it is undisputed that literature reviews generally play a central role in scholarship (cf. e.g., Cooper 1988, pp. 104ff.) and in information systems (IS) research in particular (cf. e.g., Levy and Ellis 2006, pp. 181f.; Webster and Watson 2002, pp. xiii f.). However, as the term ‘review’ is one of the “more ambiguous” terms in scholarship (Garfield 1987, p. 114), there is still a significant confusion about the structure and format of literature reviews (Webster and Watson 2002, p. xiv). Nonetheless, it seems clear that in particular the process of literature search plays a fundamental role in crafting a thorough review on a topic (Zorn and Campbell 2006, p. 174).

According to Webster and Watson (2002), a literature search in essence comprises the querying of scholarly databases using keywords and backward or forward searches on the basis of relevant articles. Whereas backward search means reviewing the references of the articles yielded from the keyword search, forward search, in turn, refers to reviewing additional sources that have cited the article (ebenda, p. xvi; also cf. Levy and Ellis 2006, pp. 190ff.). The search process is a challenging part of an IS literature review, as it should include “all sources that contain IS research publications” (Levy and Ellis 2006, p. 183). However, searching for literature is extremely complicated in an emerging field such as IS, because an incredible (and still increasing) number of articles are published in a wide range of sources every year: For example, in 2003 Peffers and Ya identified 326 journals that publish IS research (p. 65), while the Index of Information Systems Journals lists 647 active IS journals today. Moreover, Peffers and Hui (2003, p. 168) found that between 1997 and 2001 about 38% more articles were published in ten “pure” IS journals (following Walstrom and Hardgrave 2001, p. 122) when compared to 1987–1991. As a result, conducting a literature search in IS can turn into a Sisyphean task, so that, quite necessarily, a multitude of work has to be omitted in the search process (cf. Cooper 1988, p. 114).

Against this backdrop, we argue that the process of excluding sources (and including respectively) has to be made as transparent as possible in order for the review to proof credibility. Only then are readers able to assess the exhaustiveness of a review and other scholars in the field can more confidently (re)use the results in their own research. Therefore, we set out to explore the methodological rigour of literature reviews in the IS domain. In doing so, we apply typical evaluation criteria, such as reliability and validity, for analysing review articles published in the ten top-ranked IS journals, according to the consolidated list provided by the Association for Information Systems (AIS). We find a surprisingly large number of review publications that do not provide any detail on the underlying literature search.

The remainder of this paper is structured as follows. In the next section, we give an overview of our study and spell out our research approach. In section 3, we then present the results of our analysis of selected literature reviews published in the IS domain. In doing so, we particularly point out shortcomings referring to the documentation of the literature search process. In section 4, we discuss our results and derive three propositions that aim at providing explanations for our results. Implications from our

---


research are subsequently presented in the form of both guidelines for conducting IS literature reviews, in particular the process of searching the literature, and a plea for more rigour in crafting and documenting a literature search (section 5). We conclude with a short summary and discussion of our work (section 6).

2 STUDY OVERVIEW

A literature review can be defined as “a summary of a subject field that supports the identification of specific research questions” (Rowley and Slack 2004, p. 31). The primary objective of this paper is to explore the methodological rigour in conducting literature reviews as part of IS research. More specifically, we focus on the process of literature search (cf. Levy and Ellis 2006, pp. 185ff.), which involves both the identification of high quality papers and the evaluation of their applicability to the study. Similar to ‘traditional’ research methods or studies, the term ‘rigour’ refers in particular to the reliability and validity of the search process (cf. Levy and Ellis 2006, pp. 183ff.; Patel and Giaglis 2004, p. 304). As such, validity characterises the degree to which the literature search accurately uncovers the sources that the reviewer is attempting to collect. The basic question is ‘Does the reviewer search right?’, it reflects decisions such as the selection of databases, publications (i.e. journals, conferences or books) and keywords, as well as the period covered, the articles considered in the literature search, and the application of backward or forward searches (Cooper 1988, p. 114; Levy and Ellis 2006, pp. 185ff.; Torraco 2005, p. 360; Webster and Watson 2002, p. xvi). Reliability, in turn, describes the replicability of the search process, hence, making it substantial for any review article to comprehensively document the literature search. “The literature is the data of an integrative literature review” and “learning about [it] and how it was obtained, including the keywords and databases used, is of particular interest to readers, who may wonder if the literature they are familiar with was examined” (Torraco 2005, p. 360). It is also obvious that evaluating the exhaustiveness of a literature review is complicated without a thorough documentation of the search process. That being said, other researchers in the field can hardly ground their own work on the review without sufficient knowledge on where and how authors have already searched for literature.

Consequently, the guiding research question of this paper is ‘Do IS researchers comprehensively document the process of literature search in their review articles?’ and, therefore, we investigate the degree of methodological rigour, i.e. the replicability and evidence, of IS reviews. In other words, we review IS review literature. Our own process of literature search and the results drawn from the analysis are described in the next section.

3 REVIEWING IS REVIEW LITERATURE

As commonly recommended, in the following we focus on review articles of high quality (Rowley and Slack 2004, p. 32). However, identifying high quality IS literature is complicated as there is not only a vast amount of potentially relevant sources but also a great deal of literature of diverse quality (Levy and Ellis 2006, pp. 183 & 185ff.). To this end, IS scholars are frequently evaluating the quality of IS journals in order to provide rankings, e.g. Ferratt et al. (2007, p. 716), Hardgrave and Walstrom (1997, pp. 121f.), Lowry et al. (2004, pp. 52ff.), Peffers and Ya (2003, p. 70), and Willcocks et al. (2008, pp. 165f.). A number of those rankings have been synthesised by Carol Saunders for the AIS, which resulted in a comprehensive list of more than 100 top-tier IS journals. We decided to include in our literature review the ten top-ranked, peer-reviewed IS journals according to this consolidated list. However, as Harvard Business Review (HBR) (#8) is not a peer-reviewed journal and because the IEEE Transactions (#9) subsume various journals of differing quality and relevance for IS research (e.g. the IEEE Transactions on NanoBioscience), these journals were not considered in our review. This selection led us to explore the ten journals displayed in Figure 1. The figure also outlines the accessed databases as well as the searching functionality applied and the period covered in the review. The key phrase we used in all searches was ‘literature review.’
As outlined in the previous section, a literature search involves both identifying and evaluating scholarly literature and, accordingly, we applied a similar procedure in our own review. As the search phrase ‘literature review’ has been used in a range of contributions, many of which cannot be labelled as review articles, the contributions identified by keyword search (‘hits,’ cf. Figure 1) have subsequently been evaluated, based on their abstracts, in order to assess their relevance for this study. The then remaining articles became the basis of our review (‘reviewed’). Subsequently, we analysed whether the identified review articles meet, i.e. document, the requirements explained in the previous section, namely: the (number of) articles considered in the review, the period covered, the (number of) journals and databases explored, the keywords used for the database and/or journal search, and, finally, whether a backward and/or forward search were conducted. The results drawn from our analysis are summarised in Figure 2; they show that 6 articles – or one fourth of the examined review literature – do not provide any information on the underlying search process. At least, 15 publications accurately document how many articles were included in the review. However, among these, there are only 11 reviews that state precisely, which articles were included. Whereas, more or less, half of the examined articles explicate the examined period of time, the number of databases queried has merely been stated in 7 articles. However, in many cases it again remains unclear, which databases were actually accessed. For example, Melville et al. (2004, p. 322) searched journal databases which “included” Business Source Premier and JSTOR and Pateli and Giaglis (2004, p. 304) explored “several” sources such as ScienceDirect, JSTOR, and InterScience. Perhaps even more notably, many reviewers conducted a journal search instead of a database search, though there is “no justification for searching by journal instead of searching by topic” (Anonymous; cited in Webster and Watson 2002, p. xvi), except when the goal of the review gives sufficient reasoning for such a proceeding (e.g. Jasperson et al. 2002, p. 403, or the review at hand).

All in all, some inaccuracies in documenting the search process can be found in all sources. For example, Leidner and Kayworth (2006) – whose review is still among the best-documented ones – examined “the leading journals in our field dating back to the early 1990s.” They apply search phrases

3 These numbers have been checked and double-checked on 2009-01-15. However, when finalising this paper and reviewing the results once more on 2009-03-15, it turned out that some of them have considerably changed, in particular when again searching Informs (ISR and MS) and ScienceDirect (DSS). Referring to ScienceDirect, these differences can, at least partly, be reasoned by new issues released in the meantime. However, the results gained from the two Informs queries differ so significantly (ISR: 30 vs. 46 hits; MS: 108 vs. 228 hits) that we are likely to assume that the underlying searching mechanisms have somehow been changed in recent times (since our license contracts with both providers definitely stayed the same). Therefore, please note that the numbers described in Figure 1 have been gained on 2009-01-15.

4 Since 2001, the MISQ provides a special review section. Therefore, keyword search has not been applied for searching MISQ. However, all articles published in the MISQ review section have been considered in this review.
“such […] as ‘IT culture’, ‘information systems culture,’ and ‘IT values’” and conducted a search of ABI/Inform and Business Source Premier and “similar searches” in ScienceDirect. Furthermore, they “looked through” the references of “key articles” to not “overlook other articles” (pp. 360ff.). That being said, though they applied a backward search (as one of only a few articles), it is not fully replicable, as is the entire search process. However, most striking is the fact that only 2 articles comprehensively document the underlying search phrases (Gerwin and Barrowman 2002, p. 942, and Ngai and Gunasekaran 2007, p. 4). So even though we firmly believe that all the above listed contributions represent fundamental and way-leading reviews in their particular fields, we herein question whether it is in fact sufficient to state that “key words from our definition of IT business value” (Melville et al. 2004, p. 322) were used or that the search was conducted “by using relevant keywords” (Xiao and Benbasat 2007, p. 140). In the same way, Pateli and Giaglis (2004, p. 304) claim that “the selection of journals” (Fjermestad and Hiltz 1998, p. 9) is most likely to be valid, whenever the search process is only described very briefly, the search cannot be replicated and, thus, it can hardly be considered rigour. Drawing on these results, the question emerges: ‘Why is it that many IS researchers do not comprehensively document their literature search process?’

### DISCUSSION

When crafting a literature review on a topic, one should be aware of the fact that, unlike for other empirical studies, there are only few explicit methods or standardized guidelines (cf. Bem 1995; Jackson 1980, p. 440; Torracco 2005, p. 359). In 2002, Webster and Watson remarked: “as the initial senior editors for MISQ Review, we quickly learned that many IS scholars are not familiar with the structure and format of reviews” (p. xiv). This leads us to formulate a first proposition for explaining our results: (1) IS researchers cannot refer to established guidelines for documenting the literature search process.
Our second proposition provides an alternative reasoning: (2) IS researchers are not fully aware of the importance of rigorously documenting the literature search. Please note that it is not our intention to imply that IS researchers are unwilling to conduct their search process in a rigorous way – in particular referring to the outstanding contributions described above. However, conducting and documenting a literature search are two sides of the same coin and perhaps IS authors sometimes in fact prefer ‘the pears to the apples,’ i.e. they rather put emphasis on rigorously summarizing and synthesizing findings gained from the literature search than on documenting the long way they walked along for uncovering them.

Finally, our analysis also revealed that review articles in IS research differ considerably in terms of length, reaching from six pages (Glass et al. 2004) to 143 pages (Fjermestad and Hiltz 1998). Furthermore, there are many review articles that are not even longer than twenty pages. However, it is undisputed that literature reviews commonly require far more pages than ‘regular’ articles (cf. e.g., the MISQ submission guidelines). Consequently, the question arises ‘Why are so many review articles that short (and exclude the process of searching the literature), assuming that IS researchers know how (ad 1) and are willing (ad 2) to rigorously document their literature search process?’ We believe the answer is most likely to be found in the submission guidelines and page limitations provided by various IS outlets. More specifically, information on the literature search process, which exists in the beginning, might be excluded during the subsequent editing process in order to free space for addressing the reviewers’ comments. Only three of the journals we examined provide a special review section (CACM, AI and MISQ) and only another three journals (DSCI, EJIS and ISR), more or less, welcome review articles in their editorial statements. Please also note that – according to the results gained from our analysis – in the case of four of the journals in our sample no review articles were published at all. This leads us to question whether literature reviews play the key role they undoubtedly deserve in IS research and to formulate a third proposition: (3) It is the conditions of the publication process that often prevent a detailed description of the literature search. For example, Kari and Rozenberg (2008, p. 83) acknowledge this perception by stating that “the upper-bound placed on the number of [40] references” turned out to be a “real limitation” for their review published in the CACM.5

In summary, we conclude that IS reviewers may sometimes be either (1) unable to refer to adequate guidelines for rigorously documenting the literature search, (2) not fully aware of its importance or (3) hindered to transparently describe their literature search in its full extent due to editorial constraints. Here, we do not aim to evaluate our propositions, but rather formulate both guidelines for crafting an IS literature review and search (ad 1) as well as a plea for more comprehensive documentation of the search process (ad 2 & 3).

5 Note that this review could not be identified following our search strategy (limitations are discussed in the next section).

5 IMPLICATIONS

5.1 Guidelines for Literature Reviews

In response to proposition 1, we propose in the following a framework for conducting IS literature reviews, with particular focus on the process of searching the literature. The framework is displayed in Figure 3 and reflects a “circularity that exists when […] undertaking a literature review” (Baker 2000, p. 221). That being said, as things use to change and knowledge continuously grows, literature reviews often become out-of-vogue after a certain time, giving reason for an extension and update of the review (cf. Pervan 1998, p. 158).

A major challenge in reviewing the literature lies in defining an appropriate scope and flavour of the review (phase I). Reviews can be critical, interpretive, speculative, state of the art, and historical and
can vary referring to subject matter, period covered, and degree of coverage of sources (Manten 1973; Woodward 1972; as cited in Garfield 1987, p. 114). Furthermore, literature reviews can serve a wide range of, sometimes very different, purposes, reaching from gaining new and synthesising existing research outcomes to identifying research methodologies or techniques commonly used in a field (cf. Hart 1998, pp. 27ff.). In order to clearly define the scope of a review, we propose to draw on an established taxonomy for literature reviews presented by Cooper (cf. Cooper 1988, pp. 109ff., and Figure 4 in the follow-up). Cooper’s taxonomy is comprised of six constituent characteristics, each containing certain categories, some of which are mutually exclusive (perspective and coverage), while others can be combined (audience, organisation, goal, and focus). The focus (1) of a literature review is concerned with what is of utmost importance to the reviewer. Most literature reviews focus on research outcomes, research methods, theories, and/or applications (cf. Bem 1995; Torraco 2005, p. 361). Common goals (2) of literature reviews include summarising, criticising, and/or integrating findings (cf. Jackson 1980, p. 438). For organising a literature review (3), Cooper suggests a historical, conceptual or methodological structure. The perspective (4) of a review reflects whether a certain position is espoused or not; the audience (5) particularly determines the writing style of the author(s) (cf. Bem 1995, pp. 173ff.).

Figure 3. Framework for literature reviewing

This paper is based on the perception that in particular the degree of coverage of sources (6) is crucial for reviewing the literature on a topic. According to Cooper, four levels of coverage can be distinguished, namely: exhaustive (including the entirety of literature on a topic or at least most of it), exhaustive with selective citation (considering all the relevant sources, but describing only a sample), representative (including only a sample that typifies larger groups of articles), and central (reviewing the literature pivotal to a topic) (Cooper 1988, pp. 110ff.).

![Figure 4. Taxonomy of literature reviews (following Cooper 1988, p. 109)](image)

While the above framework does not provide immediate answers to the questions of literature search, its application is a necessary first step of clarification in any literature review, which bears implications for the later search process. An exemplary application of Cooper’s taxonomy is also given in
Figure 4 by highlighting categories that characterise our own literature review presented in section 2: We did not focus (1) on the research outcomes or theories described or applied in the above analysed articles, but rather on their underlying research methodology. We criticized (2) the way these reviews document the literature search process, since it was our position (4) that such documentation is crucial to ensure the replicability of the literature search. As the purpose of our review was not to summarise or synthesise, but rather to evaluate scholarly literature, the categories proposed by Cooper relating to organisation (3) can hardly be applied to our case. It is hoped that our results are of some value for the whole IS community (5) – even though we did not consider all IS review articles ever published in our study, but restricted the analysis to a small sample of literature reviews only. However, this proceeding can be reasoned by the fact that it was our explicit objective to analyse high-quality review articles that may be considered as representative (6) for the IS domain.

In a next step, attention should be paid to the fact that a review must begin with “a broad conception of what is known about the topic and potential areas where knowledge may be needed” (phase II) (Torracco 2005, p. 359). Therefore, working definitions of the key terms should be provided at this point (Zorn and Campbell 2006, p. 175). Baker (2000, p. 222) suggests that one should firstly consult “those sources most likely to contain a summary or overview of the key issues relevant to a subject,” such as seminal textbooks, encyclopaedias, or handbooks. A reasonable way for identifying key concepts is represented by concept mapping, which also provides the opportunity to uncover relevant search terms (in particular related concepts or synonyms and homonyms) that can be applied in the subsequent literature search (cf. Rowley and Slack 2004, p. 36). Accordingly, we began our study by consulting the seminal textbooks on literature reviews by Arlene Fink (2005) and Christopher Hart (1998) and introduced a working definition of the term ‘literature review’ in section 2. However, the various synonyms of the term ‘literature review,’ such as ‘meta-analysis’ or ‘research synthesis,’ have been disregarded in our study. Therefore, it is possible (or indeed even likely) that several articles which may be labelled as ‘literature reviews’ have not been included in our study, which again shows that our search strategy can hardly be labelled as exhaustive – even referring to the small sample of IS journals examined.

As we already noted, the search process (phase III) involves database, keyword, backward, and forward search, as well as an ongoing evaluation of sources (cf. Figure 5).

![Figure 5. Literature search process](image)

As they have typically been peer-refereed before publication, it is commonly recommended to focus on articles published in scholarly journals (Rowley and Slack 2004, p. 32) or proceedings of renowned conferences (Webster and Watson 2002, p. xvi). However, one should note that the quality of contributions in conference proceedings is usually considered lower and less mature than those in journals (cf. Levy and Ellis 2006, p. 187). Thus, authors who intend to include conference articles in their reviews should concentrate on the better ones (e.g., Walstrom and Hardgrave (2001, p. 121) and Willcocks et al. (2008, p. 166) provide rankings for IS conferences). Consequently, the identification of journals makes up the first sub phase of our framework – even though we would agree that it rather makes sense to query scholarly databases allowing for a topic-based search (cf. Webster and Watson 2002, p. xvi). However, one challenge definitely lies in identifying proper databases. We propose to...
search those databases (sub phase 2) providing access to the leading IS journals (identified in sub phase 1) – a proceeding which finally allows for ensuring that all the top-tier sources are included in the review. The identified databases have then to be queried on the basis of a keyword search (sub phase 3). It is commonly recommended to use a precise (set of) search phrase(s) in order to exclude contributions covering topics or research questions which are not necessarily relevant (Rowley and Slack 2004, p. 35). Querying EBSCOhost using the search term ‘literature review,’ for example, reveals hundreds of possible information sources; when adding the term ‘information systems’ however, the amount of literature is significantly reduced. By also adding the supplementary search phrase ‘Europe,’ the huge pool of potentially relevant articles is finally restricted to merely a handful of papers. This example illustrates that keyword search is crucial, since the selection and combination of the search phrases sets “the parameters of the research itself” (Baker 2000, p. 222). Thus, particularly the applied keywords have to be documented precisely, so that other scholars can evaluate whether they sufficiently match the topic under investigation. According to Webster and Watson (2002, p. xvi), the process of backward search refers to reviewing older literature cited in the articles yielded from the keyword search and forward search means reviewing additional sources that have cited the article (sub phase 4), e.g. by querying the ISI Web of Knowledge by Thomson Reuters (note that Levy and Ellis (2006, pp. 190ff.) discuss additional forms of backward and forward search in detail). Evaluation in all phases means limiting the amount of literature identified by keyword search as well as backward and forward search to only those articles relevant to the topic at hand. We therefore propose an evaluation of the articles’ contents, which may mean to analyse their titles, abstracts or even full texts. As described above, the review at hand is based on applying a keyword search (‘literature review’) to a sample of ten journals. That being said, though these journals were accessed via online databases, a ‘pure’ database search has not been conducted. Moreover, backward or forward searches were also not part of our literature search strategy.

After collecting sufficient literature on a topic it has to be analysed and synthesised (phase IV). As we focus on the search process, we herein only briefly touch upon this phase. For the analysis, a concept matrix, as developed by Salipante et al. (1982) and adapted for IS literature reviews by Webster and Watson (2002, p. xvii), can be used, which subdivides topic-related concepts into different units of analysis (cf. Figure 6). This allows for arranging, discussing, and synthesising prior research.

![Concept matrix](image)

**Figure 6. Concept matrix (Webster and Watson 2002, p. xvii)**

However, as underpinned by several IS reviews (e.g., Ahuja 2002, pp. 30ff.), the synthesis of literature is further expected to result into a research agenda (phase V), comprised of sharper and more insightful questions for future research (Webster and Watson 2002, p. xix). The research agenda provides the basis for extending the review in order for the IS community to keep up-to-date and it may be developed based on the proposed concept matrix. That being said, certain fields of the matrix, which remain ‘blank’ during a literature study, often highlight research areas that are significantly under-researched.

### 5.2 A Plea for More Rigour in Searching the Literature

We hope that the above section provides some insights on how to rigorously conduct a literature review and search in the IS domain. However, we did not intend to present ‘yet another guideline’ for reviewing the literature – we rather consider our explanations as both a rationale for our own literature study and a response to scholars who are likely to estimate our first proposition to be the most striking one (‘IS researchers cannot refer to adequate guidelines for rigorously documenting the literature...')
search’). Nonetheless, we think that most IS scholars – and in particular those very renowned ones cited above – indeed have a good understanding of how to conduct a literature search. Moreover, we firmly believe that the literature search process underlying (at least most of) the analysed articles is certainly a highly rigorous one. However, it is simply not documented rigorously enough, leading us to question the role the literature search process plays in the IS domain. Thus, one may also consider this paper as ‘yet another call’ for more rigour in IS research; in this case, for documenting a literature search. In the following, we appeal to the authors of IS articles (proposition 2) as well as to the editors of IS journals (proposition 3) to not only conduct and write (or support the writing of) literature reviews that are of high quality, i.e. reviews that are thoroughly crafted, useful and interesting, but also to allow for adequate documenting of the review process, including the specifics of the search process.

When writing a literature review, we advocate authors to invest in planning and describing the literature search process in detail. In the majority of cases, literature reviews serve as the means to reveal open research gaps and are part of a larger research endeavour. Planning and accomplishing the literature search process in a rigorous manner will help to discover similar research endeavours early and prevent the researcher from doing redundant work. Moreover, it will allow for a better understanding of where to find relevant articles in the jungle of different sources available and how to retrieve them later. Rigorously documenting the literature search in particular means to share these experiences with the community. As a result, a well-documented review will be gladly used, extended and cited by other researchers, which will in turn positively affect scientific impact.

We hope for more editors to consider establishing a dedicated literature review category within their journals. Moreover, the rigorous accomplishment of a literature review should not be impeded by article length or reference restrictions. One possible solution may be the publication of two versions of the same review – a shorter one that contains the major findings and is published in the printed version and a comprehensive one that outlines the whole literature search process and states how and where the sources were discovered and which is published online (see Kari and Rozenberg (2008) for an example). Moreover, we suggest demanding the documentation of the literature search process as an important review criterion for judging the quality of literature reviews.

6 CONCLUSION

Research is a collaborative endeavour, since each researcher builds on what has been developed and worked out before. Thus, literature reviews play an important role in scholarship. In this paper we argued that documenting the literature search process is a crucial part in any review article. However, drawing on the results gained from an analysis of literature reviews published in ten of the most important IS outlets, we showed that many of these reviews do not thoroughly document the process of literature search, e.g. by not comprehensibly explicating the queried databases, the applied keywords or the examined period of time. This is why we presented both guidelines for conducting a literature search and review as well as a plea for more rigour in documenting the search process.

Certainly, most IS scholars know what it actually takes to craft high-quality reviews, and we would agree that our suggestions are likely to further impede this challenging task. But our guidelines do not intend to imply that conducting a literature search means to uncover and analyse all sources ever published. A review that considers only five research papers, but sufficiently states which ones were chosen for whatever sensible reasons, may be of more value to both its authors and the community than a review that analyses a broad range of contributions, without providing sufficient information on where, why and what literature was obtained, hence, making it hard to judge its quality and the scope of its contribution. In addition, a well documented literature search offers the potential to being extended and transferred to, for instance, other domains, additional journals, or newer/older volumes. Against this background, search results are better comparable and a well-documented search process thus provides the basis for a review article to contribute to a cumulative effort of reviewing literature.
References


AN EVALUATION OF USER ACCEPTANCE OF A CORPORATE INTRANET

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0216.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>knowledge transfer, Web Site Analysis, User Satisfaction, Technology Acceptance Model (TAM)</td>
</tr>
</tbody>
</table>
AN EVALUATION OF USER ACCEPTANCE OF A CORPORATE INTRANET

Barnes, Stuart, Norwich Business School, University of East Anglia, Norwich NR4 7TJ, UK, stuart.barnes@uea.ac.uk
Vidgen, Richard, School of Management, University of Bath, Bath BA2 7AY, UK, richard@vidgen.com

Abstract

Intranets represent an important organisational resource for knowledge sharing. However, as yet, there has been little research into the quality of intranets and the impact of quality on intranet user acceptance. In the study reported in this paper, an intranet quality assessment tool comprising the dimensions usability, design, and information quality, is combined with perceived usefulness and social influence from the technology acceptance literature to create an intranet acceptance model. The model is applied to the sales and marketing division of an international manufacturing company. Data is collected via a Web survey (n=131, response rate = 65.5%) and tested using the partial least squares approach to structural equation modelling. The results show that intranet quality is a significant factor in determining behavioural intention to use, although it is less important than perceived usefulness and social influence. Comments collected from respondents are used to illustrate the findings and provide an insight into user behaviour. The discussion considers the implications, future research (e.g., the role of social influence in intranet usage) and limitations. The paper rounds off with a short summary.
1 INTRODUCTION AND BACKGROUND TO THE RESEARCH

While the Internet allows organizations to reach new markets and new consumers the corporate intranet represents a key medium for access to organizational memory and to company and product information. Skok and Kalmanovitch (2005) see intranets in the context of the larger issue of knowledge sharing, identifying three views of the role of the intranet: cognitivistic, connectionist, and autopoietic. The cognitivistic view sees the intranet as a culturally and socially neutral repository of explicit information. The connectionist sees the intranet as a tool that brings communities together to share and interpret information, making hidden organizational information locatable. The autopoietic view recognizes that much knowledge is tacit and the role of the intranet is to provide data that helps identify relevant staff to contact. Notwithstanding the fact that researchers may debate the finer distinctions of data, information, and knowledge the intranet, as a technical artefact, undoubtedly has a role to play in knowledge sharing and the quality of the intranet would be expected to contribute to intranet acceptance and ultimately to the effectiveness of knowledge sharing.

There is a surfeit of research literature with respect to evaluating the quality and user acceptance of Web sites on the public Internet and e-commerce applications, adopting a range of approaches and theoretical bases. For example, Liu and Arnett (2000) highlight four factors in Web site quality: information and service quality, system use, playfulness, and system design quality. Yoo and Donthu (2001) developed the SITEQUAL scale to measure the perceived quality of an online shop, leading them to identify four dimensions: ease of use, aesthetic design, processing speed and security. Barnes and Vidgen (2002) developed the eQual scale with five dimensions: usability, design, information, trust, and empathy. Loiacono et al. (2002) created the WebQual™ scale, composed of thirty-six items and twelve dimensions: informational fit to task; interactivity; trust; response time; design appeal; intuitiveness; visual appeal; innovativeness; flow (emotional appeal); integrated communication; business process; and substitutability. Further, Parasuraman et al. (2005) developed the e-SQ scale with seven dimensions of service quality: efficiency, system availability, fulfilment, privacy, responsiveness, compensation and contact.

A range of theoretical bases for Web quality evaluation have been drawn upon in the salient literature, including the technology acceptance model, service quality/marketing, and information systems success. These approaches are typically supplemented with items from ancillary areas such as human computer action and information quality. Typical dependent variables are intention to use, intention to purchase or re-purchase, loyalty, and overall satisfaction.

Compared with the research on Web site and e-commerce quality and acceptance there is, in comparison, a dearth of research on the evaluation of intranet effectiveness. Following many years of studying intranets Stenmark (2005) finds that intranets differ from the public Internet insofar as content is provided top-down by a small group of professionals, content is objective and business-related, and content is organized consistently according to a corporate taxonomy. Stenmark (ibid.) says that the latter is a source of weakness in intranets, which can be difficult to search and navigate as they are seldom built from the consumer’s perspective. The intranet exists in an environment with a formal organization structure and a management hierarchy that is absent in an e-commerce application and we would thus expect social influence to play a significant role in intranet use. Coming from a quality perspective, Leung (2001) draws on the ISO model and evaluates the intranet as a software artefact using the dimensions of functionality, reliability, usability, efficiency, maintainability, and portability.

Previous research on intranet quality and acceptance is clearly very limited. Thus, we perceive an important gap that deserves further research attention. The aim of this paper is to show how the effectiveness of an intranet can be assessed in terms of outcomes, principally in performance effectiveness (i.e., how useful is the intranet in supporting organizational work) and intention to use. Our approach is the development and testing of a model of intranet acceptance. The model is
subsequently tested in a real case study setting of a major multinational organisation. Data is collected using a survey and analysed using the partial least squares approach to structural equation modelling.

In the next section we explicate a model and theory developed to contribute understanding in this very new area of investigation. This is followed by a discussion of the methodology adopted for the study. Consequently, the research results - quantitative and qualitative analysis of respondent comments - are examined in section four. The findings, implications for research and practice, and limitations of the research are discussed in section five. Finally, the paper rounds off in the last section with a summary.

2 MODEL AND HYPOTHESES

In this section we describe the theoretical model adopted for the study. We discuss this in two parts: (1) intranet quality assessment; and (2) intranet acceptance and associated constructs. Let us examine each of these in turn.

2.1 Intranet quality

The eQual instrument (Barnes and Vidgen, 2002) identified five dimensions of Web quality: usability, design, information, empathy, and trust. Of these five dimensions, empathy consistently scores lower in terms of importance in e-commerce applications and is also unlikely to play a major role in a corporate setting where there will be at least a degree of uniformity of corporate culture. Although trust is a key factor in e-commerce, in a corporate setting trust is implied through the corporate and closed setting of the intranet. We thus focus on usability, design, and information quality in constructing an instrument to assess the quality of an intranet.

The usability dimension draws from literature in the field of human computer interaction (Davis, 1989, 1993, Nielsen, 1993) and information systems quality (Bailey and Pearson, 1983) while the design dimension draws on Web usability (Nielsen, 1999, 2000, Spool, 1999) and service quality (Parasuraman et al. 1991, Zeithaml et al., 1990). Information quality follows Strong et al. (1997) with provenance in the work on IS quality and IS success by Bailey and Pearson (1983) and DeLone and McLean (1992). In summary, intranet quality is expressed using the dimensions usability, design and information quality (see Barnes and Vidgen (2002) for further details of the provenance of the intranet quality items). We expect, as from previous research, for the quality construct to appear as a second-order combination of these three first-order constructs and we hypothesise the following:

H1. Intranet quality is positively related to behavioural intention to use an intranet.

2.2 Intranet acceptance

Although intranet quality would be expected to impact behavioural intention to use the intranet it is clear from the literature that other factors will likely play a part. In his seminal paper on technology acceptance, Davis (1989) defines perceived usefulness as “the degree to which a person believes that using a particular system would enhance his or her job performance” (p. 320). In a later version of the model, TAM2, Venkatesh and Davis (2000) included the construct “subjective norm”, which considers social influence as the extent to which an individual perceives that people important to him or her think he or she should or should not perform a behaviour, such as using an intranet. Latterly, the unified theory of acceptance and use of technology (UTAUT) brings these strands together in a large and wide-ranging instrument (Venkatesh et al., 2003). Despite its grand scale the explanatory and predictive power of the UTAUT in a broad range of contexts has yet to be fully demonstrated.

Part of our research agenda is to produce an instrument suitable for use in the corporate world where lengthy surveys are likely to be ignored. We therefore happily accepted the requirement that the instrument be parsimonious in terms of the number of items and we are, in a way, getting back to the
simplicity and elegance of the original TAM. First and foremost, an intranet must be useful to employees in doing their jobs. Second, we argue that social influence in the form of senior managers, peers, and co-workers will also play a key role in intention to use. From the TAM/UTAUT we therefore hypothesise:

**H2.** Social influence is positively related to behavioural intention to use an intranet.

**H3.** Perceived usefulness is positively related to behavioural intention to use an intranet.

Finally, behavioural intention is expected to have an influence on actual usage, hence:

**H4.** Behavioural intention is positively related to actual intranet use.

The above hypotheses contribute to the construction of the research model, as shown in Figure 1.

![Research Model](image)

**Figure 1.** Research Model

### 3 METHODOLOGY

To test the research model we worked with the sales and marketing division of a multi-national manufacturing company, which will be referred to pseudonymously here as SalesOrg. The company had recently employed a Web design agency to build a new intranet site to support their sales and marketing operations. The old intranet was fragmented in terms of design and content and was not widely adopted by the sales and marketing teams. The organization now wished to create a single and enhanced repository of sales and marketing data containing material about SalesOrg’s values, processes, and methods along with tools and templates, case studies, best practice examples, news items, resources (e.g., product information), and e-learning units. In sum, the new intranet is to be a single point of access for all the formalized resources that all personnel working in sales and marketing need to draw on to do their job effectively. The survey reported here was conducted three months after the initial intranet implementation.

The survey was implemented online and pilot tested with a small group to check that the questions were comprehensible and the survey easy to complete. Responses to the pilot survey showed that it was relevant and comprehensible and SalesOrg agreed to roll the survey out to a wider audience in February 2008. The survey questions are shown in the Appendix. For the items in Table 1 five-point Likert scales were used with the anchors “1 = strongly disagree” and “5 = strongly agree”. Respondents were allowed to select “not applicable” to avoid forcing responses. Actual use was measured using a single item, “In an average week, how much time would you say you spend connected to the Intranet (for any kind of service and counting all the possible sessions over the week)?”, with responses as “1=Don't use it at all”, “2=Less than 15 minutes”, “3=Between 15 minutes and 30 minutes”, “4=Between 30 minutes and 1 hour”, “5=Between 1 and 2 hours”, “6=Between 2
and 4 hours”, “7=Between 4 and 10 hours” and “8=More than 10 hours” (mean=3.66, standard error=0.139). The survey was implemented online and publicised by email to the 200 members of the SalesOrg division. A total of 131 usable responses were received, representing a respectable response rate of 65.5%.

Data analysis was performed using a variance maximization approach to structural equation modelling (SEM) and associated statistics for validity and reliability. More specifically, we used the partial least squares (PLS) technique with reflective indicators in Smart-PLS 2.0 (Ringle et al. 2005). The PLS technique has become increasingly popular in information systems research, marketing and in management research more generally in the last decade or so, influenced by its flexibility; indeed, PLS does not have the same distributional assumptions of normality for data and is able to handle small- to medium-sized samples (Chin 1998, Compeau & Higgins 1995).

SEM caters for two types of analysis: construct and measurement. The construct model is used to assess the structure of the observed and latent variables (e.g., do the three survey items capture the latent variable “Perceived Usefulness”?). The measurement model considers the hypothesised relationships, or paths, between the variables and latent variables, e.g., is the relationship between intranet quality and behavioural intention significant and if so how strong is the relationship?

4 RESULTS

4.1 Characteristics of survey respondents

Of the 131 usable responses to the survey, 63.36% were male (36.64% female), with approximately half being in the median age range of 36 to 45 years of age (48.09%). The respondents were very experienced employees in the sales and marketing profession with 69.47% having more than 10 years experience (90.84% over 5 years). Similarly, the employees had been with the company for quite some time, with more than half over 10 years and 81.68% over 5 years. The respondents were only moderate users of the Internet, with a median of 4-10 hours per week spent connected to the Internet in an average week. The use of the intranet in an average week was generally quite low, with around three-quarters of users using it for less than an hour (74.81%) and a median of 15 to 30 minutes.

4.2 Tests for Validity and Reliability of the Measures

To test the constructs we performed a confirmatory factor analysis and reliability analysis. Table 1 demonstrates that the scale items exhibit high levels of convergent validity – the extent to which theoretical scale items are empirically related. The loadings of the measures on their respective constructs in the model range from 0.756 to 0.930, with all being significant at the 0.1% level.

Table 1 also demonstrates that all of the constructs fulfil the recommended levels with reference to composite reliability (CR) and average variance extracted (AVE). All items were higher than the cut-off of 0.50 for AVE recommended by Fornell and Larcker (1981), ranging from 0.607 to 0.802. Similarly, the values for composite reliability are very good, ranging from 0.882 to 0.948, well above the reliability values of 0.70 and 0.80 that are typically advised for building strong measurement constructs (Nunnally 1978, Straub & Carlson 1989).

As an additional test for discriminant validity we utilized the cross-loading method of Chin (1998). The method prescribes a requirement for measurement items to load higher on a construct than the scale items for other constructs and for no cross-loading to occur. Item loadings in the relevant construct columns were all higher than the loadings of items designed to measure other constructs; similarly, when glancing across the rows the item loadings are considerably higher for their corresponding constructs than for other constructs. Note that the three first-order constructs for
intranet quality, usability, design and information quality all loaded on a single second-order construct, as expected. This underlines the point that intranet quality is a holistic concept in its own right.

Table 2 examines the extent to which question items measure the construct intended or other related constructs, otherwise known as discriminant validity. Fornell and Larcker’s (1981) standard test for discriminant validity was used, whereby the square root of average variance extracted for each construct is compared with the correlations between it and other constructs; discriminant validity is demonstrated if the square root is higher than the correlations. Table 2 clearly indicates that each construct shares greater variance with its own measurement items than with other constructs with different measurement items, and with a good margin of difference.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Item</th>
<th>Loading</th>
<th>Mean</th>
<th>St. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioural Intention</td>
<td>Behavioural Intention1 (BI1)</td>
<td>0.867</td>
<td>4.05</td>
<td>0.079</td>
<td>21.698</td>
</tr>
<tr>
<td>Behavioural Intention</td>
<td>Behavioural Intention2 (BI2)</td>
<td>0.889</td>
<td>4.20</td>
<td>0.075</td>
<td>29.436</td>
</tr>
<tr>
<td>Behavioural Intention</td>
<td>Behavioural Intention3 (BI3)</td>
<td>0.910</td>
<td>4.21</td>
<td>0.071</td>
<td>40.815</td>
</tr>
</tbody>
</table>

Intranet Quality (Second Order Factor; CR=0.948, AVE=0.607)

<table>
<thead>
<tr>
<th>Section</th>
<th>Item</th>
<th>Loading</th>
<th>Mean</th>
<th>St. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>Usability 1 (USE1)</td>
<td>0.878</td>
<td>4.29</td>
<td>0.064</td>
<td>33.185</td>
</tr>
<tr>
<td>Usability</td>
<td>Usability 2 (USE2)</td>
<td>0.862</td>
<td>4.26</td>
<td>0.062</td>
<td>26.555</td>
</tr>
<tr>
<td>Usability</td>
<td>Usability 3 (USE3)</td>
<td>0.852</td>
<td>4.17</td>
<td>0.071</td>
<td>22.444</td>
</tr>
<tr>
<td>Usability</td>
<td>Usability 4 (USE4)</td>
<td>0.874</td>
<td>4.19</td>
<td>0.060</td>
<td>34.851</td>
</tr>
<tr>
<td>Design</td>
<td>Design 1 (DES1)</td>
<td>0.789</td>
<td>4.15</td>
<td>0.073</td>
<td>14.135</td>
</tr>
<tr>
<td>Design</td>
<td>Design 2 (DES2)</td>
<td>0.867</td>
<td>4.07</td>
<td>0.066</td>
<td>35.333</td>
</tr>
<tr>
<td>Design</td>
<td>Design 3 (DES3)</td>
<td>0.770</td>
<td>4.19</td>
<td>0.068</td>
<td>12.634</td>
</tr>
<tr>
<td>Design</td>
<td>Design 4 (DES4)</td>
<td>0.792</td>
<td>4.34</td>
<td>0.059</td>
<td>14.357</td>
</tr>
<tr>
<td>Information</td>
<td>Information 1 (INFO1)</td>
<td>0.873</td>
<td>4.32</td>
<td>0.062</td>
<td>33.496</td>
</tr>
<tr>
<td>Information</td>
<td>Information 2 (INFO2)</td>
<td>0.807</td>
<td>3.91</td>
<td>0.078</td>
<td>22.979</td>
</tr>
<tr>
<td>Information</td>
<td>Information 3 (INFO3)</td>
<td>0.810</td>
<td>3.98</td>
<td>0.073</td>
<td>16.279</td>
</tr>
<tr>
<td>Information</td>
<td>Information 4 (INFO4)</td>
<td>0.840</td>
<td>4.21</td>
<td>0.064</td>
<td>25.893</td>
</tr>
</tbody>
</table>

Perceived Usefulness

<table>
<thead>
<tr>
<th>Section</th>
<th>Item</th>
<th>Loading</th>
<th>Mean</th>
<th>St. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Usefulness</td>
<td>Perceived Usefulness 1 (PU1)</td>
<td>0.856</td>
<td>3.91</td>
<td>0.075</td>
<td>24.260</td>
</tr>
<tr>
<td>Perceived Usefulness</td>
<td>Perceived Usefulness 2 (PU2)</td>
<td>0.930</td>
<td>4.05</td>
<td>0.074</td>
<td>45.594</td>
</tr>
<tr>
<td>Perceived Usefulness</td>
<td>Perceived Usefulness 3 (PU3)</td>
<td>0.900</td>
<td>4.29</td>
<td>0.069</td>
<td>53.889</td>
</tr>
</tbody>
</table>

Social Influence

<table>
<thead>
<tr>
<th>Section</th>
<th>Item</th>
<th>Loading</th>
<th>Mean</th>
<th>St. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Influence</td>
<td>Social Influence 1 (SOCIAL1)</td>
<td>0.804</td>
<td>3.66</td>
<td>0.096</td>
<td>15.815</td>
</tr>
<tr>
<td>Social Influence</td>
<td>Social Influence 2 (SOCIAL2)</td>
<td>0.847</td>
<td>3.72</td>
<td>0.095</td>
<td>25.166</td>
</tr>
<tr>
<td>Social Influence</td>
<td>Social Influence 3 (SOCIAL3)</td>
<td>0.756</td>
<td>3.32</td>
<td>0.097</td>
<td>11.966</td>
</tr>
<tr>
<td>Social Influence</td>
<td>Social Influence 4 (SOCIAL4)</td>
<td>0.817</td>
<td>3.78</td>
<td>0.088</td>
<td>22.475</td>
</tr>
</tbody>
</table>

Note: CR=Composite Reliability; AVE=Average Variance Extracted; n=131.

Table 1: Psychometric Table of Measurements

Overall, the results of testing for validity and reliability are very positive and provide us with a high degree of confidence in the scale items used in the study.

<table>
<thead>
<tr>
<th>ACTUAL USE</th>
<th>BI</th>
<th>PU</th>
<th>SOCIAL</th>
<th>QUALITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>n.a.</td>
<td>0.315</td>
<td>0.889</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.297</td>
<td>0.671</td>
<td>0.896</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.244</td>
<td>0.600</td>
<td>0.525</td>
<td>0.807</td>
<td></td>
</tr>
<tr>
<td>0.081</td>
<td>0.577</td>
<td>0.667</td>
<td>0.404</td>
<td>0.777</td>
</tr>
</tbody>
</table>
4.3 Test of the Research Model

The results of PLS path modelling are shown in Figure 2 (intranet quality is included as a second-order construct). A power analysis in G*Power 3.0 (Faul et al. 2007) shows that the sample size \(n=131\) has good power for explaining medium population effects \(f^2=0.15; \alpha=0.05; 1-\beta=0.95\), and is thus suitable for the testing of the model under these conditions. The results show strong support for the theoretical research model. Indeed, all paths in the model are supported at the 1 percent level of significance or higher. Specifically, the results show that H1 (the influence of quality on behavioural intention to use the intranet) is supported at the \(p<0.01\) level and that H2, H3, and H4 are supported and significant at the \(p<0.001\) level. The strongest determinant of the behavioural intention to use the SalesOrg intranet is perceived usefulness \((\beta=0.365)\) followed closely by social influence \((\beta=0.327)\) and lastly by a slightly weaker intranet quality \((\beta=0.201)\).

4.4 Analysis of comments

The quantitative data shows that the intranet was particularly well received by the users in the sample with the quality average score being 4.167 (standard error=0.051) where the maximum, 5, represents “strongly agree”. Table 1 shows that the evaluations for each of the quality scale items were well over the neutral point (neutral=3), with the three constructs averaging 4.227, 4.188 and 4.104 for usability, design and information respectively. It seems that the intranet was particularly well received as a
technical artefact and this view is supported by the open comments that respondents supplied as part of the survey. For example:

“Many thumps [sic] up for a remarkable revamp of the site”

“The site provides accurate and important information that helps to improve the training strategy it is been implemented in my country, it is friendly and easy to use.”

“Great Tool!”

Although some reservations were expressed about the quality of the site with respect to performance - “It’s slow to download files” and “site hangs once in a while from our end” - the overwhelming view of the site, as expressed through the Likert scores and accompanying comments, is positive.

The great majority of comments related to the usefulness of the site and in particular the contribution to forms of knowledge sharing. All were positive about the usefulness of the intranet and just a few are included here for illustration:

“I consider the [SalesOrg] intranet site a very comprehensive resource center for sales trainers.”

“The [SalesOrg] site has been useful resource to me in my day to day job. It helps to widen my knowledge and new ideas to implement in local countries.”

 “[SalesOrg] material is a unique contribution to our professional development. The website has finally made the workshop materials available for reference and dissemination within the local teams.”

“I really find the [SalesOrg] site a complete library of learning. It has given me so much help in crafting my training plans for my team.”

“We find the site very useful to us. We get to know more about the best practices of other countries.”

The respondents also highlighted areas for development, including translation into other languages (most notably into Spanish) and concerns were raised about the need to update the content of the intranet more frequently.

With regard to social influence, there was strong agreement amongst respondents that the site could be promoted more effectively:

“the site is not marketed very properly”.

“Site has not been promoted to me discovered it by chance because Pam mentioned it to me.”

“I think senior management is not truly aware of [SalesOrg] website. There are many programs in Sales and Marketing areas that could be helpful, but in my own experience, outside my department (Training and Development), nobody has commented me something, about this site.”

“I think in our area marketing the site need to be taken more seriously and sessions for clarification for users should be conducted plus that senior managers should indorse this initiative and ask their teams to use.”

Some of the respondents were first time or relatively new users and the quality survey proved to be a useful way of promoting the SalesOrg intranet site:

“We find the site very useful and attractive, however we must admit that we did not use it much due to lack of time. Our intention is to increase the time of navigation in order to take advantage of all of the contents of the site.”

“I have not used the Site, feel it can add value, will endeavour to use in the future.”
5 DISCUSSION

The results show that the quality construct is robust and is a reliable measure of intranet quality. Unsurprisingly, perceived usefulness is the strongest indicator of intention to use the intranet, followed by social influence. Quality is significant but does not have as large an impact on intention to use as do social influence and perceived usefulness. We speculate that once a quality threshold is surpassed attention turns to contextual factors such as whether the site is useful in doing a job. It seems reasonable to conjecture that a poor quality site would have a greater impact on behavioural intention because a poor quality site is a barrier to use while a high quality site merges into the background allowing usefulness and social influence to emerge to become of central concern. Social influence is likely to be a stronger factor in intranet applications than it is in general e-commerce applications due to the existence of formal management and organizational structures as well as powerful (but often hidden) social networks in organizations (Cross et al., 2001). Each organization will have a different balance and mix of formal and informal networks and this mix needs to be considered in any in-depth analysis of social influence through techniques such as social network analysis (Vidgen et al., 2007).

The relationship of behavioural intention and actual usage although highly significant is not strong. This is likely due to actual usage data being collected at the same time as intention data. Actual usage is a lagged effect and needs to be assessed at a different time than intention. Ideally, actual usage would be measured by objective data such as web server logs rather than being self-assessed. However, this is not straightforward since it requires that respondents are tracked through a unique code which means that the survey can no longer easily be administered anonymously.

There are numerous implications for practice and SalesOrg. First, an intranet acceptance survey can be a good way of promoting a new intranet site to new users. Second, the impact of social influence suggests that the site must be promoted and supported by senior management in order to build usage. This deployment of the formal organizational structure should be complemented by a study of the informal social networks (Cross and Parker, 2003) that support knowledge sharing in organizations (Cross et al., 2001), possibly leading to the nurturing of communities of practice (Wenger et al. 2002). Third, the content needs to be updated frequently to maintain user interest and repeat visits, and to bolster perceived usefulness and information quality. Fourth, the qualitative comments suggest that international aspects such as multi-lingual capabilities and performance/bandwidth constraints should be considered.

Although the findings of the research are significant, it is worth considering some of the limitations of the current study. First, although the sample has good explanatory power for medium-sized effects \( \beta^2=0.15; \alpha=0.05; 1-\beta=0.95 \), a larger sample could help to reveal smaller population effects. A further limitation of our research model is the absence of demographic and other factors, such as age, gender, geographical location, experience, use intensity, and so on. Since we may have a heterogeneous sample of a global nature, capturing and analyzing information on the sample may help to understand the behaviour of different groups of global employees. Second, the model was tested with a specific case study. Although the model is parsimonious, further work is needed to test whether it can be more widely applied to other contexts. For example it would be useful to examine corporate intranet sites that score low on the quality dimension to see if low quality has a relatively greater impact on behavioural intention. Third, the research is limited in terms of the measure of actual usage. The research is limited in that it adopts a single self-reported and unlagged measure of usage as a proxy for actual usage. This is not uncommon in similar studies. However, to tackle this deficiency, we hope that, in the future, a longitudinal study could be designed to measure actual usage over time.

6 SUMMARY

This study has successfully tested a model of intranet acceptance in the context of the sales and marketing department of a large multinational manufacturing company. The model is robustly
supported by the data and shows that behavioural intention is strongly driven by perceived usefulness, social influence and intranet quality, a second-order intranet quality construct comprising usability, design quality and information quality. In summary, this study has shed light on the considerably under-researched topic of intranet quality assessment and user acceptance. We hope that this research paves the way for others to begin investigating aspects of user behaviour and knowledge sharing surrounding this very important type of Web-enabled corporate information system.

References


APPENDIX: SURVEY INSTRUMENT

Intranet Quality

Usability
Learning to operate the site is easy for me
My interaction with the site is clear and understandable
I find the site easy to navigate
In general, I find the site easy to use

Design quality
The site has an attractive appearance
The design of the site is appropriate for this type of site
The site conveys a sense of competency
The site creates a positive experience for me

Information quality
The site provides accurate information
The site provides timely information
The site provides complete information
The site provides easy to understand information

Acceptance

Social Influence
People who influence my behaviour think that I should use the site
People who are important to me think I should use the site
The senior management have been helpful in the use of the site
In general, the organization has supported the use of the site

Perceived usefulness
Using the site enables me to accomplish tasks more quickly
Using the site increases my productivity
Overall, I find the site useful in my job

Intention to use
I intend to use the Site on a regular basis
I predict I will continue to use the site on a regular basis
I plan to use the site on a regular basis
Integrating Value-Driven Feedback and Recommendation Mechanisms into Business Intelligence Systems

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0512.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Intelligence, Data Warehousing, Decision support systems, Design research</td>
</tr>
</tbody>
</table>
INTEGRATING VALUE-DRIVEN FEEDBACK AND RECOMMENDATION MECHANISMS INTO BUSINESS INTELLIGENCE SYSTEMS

(Research in Progress)

Kolodner, Yoav, Ben Gurion University of the Negev,
P.O. Box 653, Beer-Sheva, 84105, Israel, kolodner@bgu.ac.il

Even, Adir, Ben Gurion University of the Negev,
P.O. Box 653, Beer-Sheva, 84105, Israel, adireven@bgu.ac.il

Abstract

Business intelligence (BI) systems and tools are broadly adopted in organizations today, supporting activities such as data analysis, managerial decision making, and business-performance measurement. Our research investigates the integration of feedback and recommendation mechanisms (FRM) into BI solutions. We define FRM as textual, visual, and/or graphical cues that are embedded into front-end BI tools and guide the end-user to consider using certain data subsets and analysis forms. Our working hypothesis is that the integration of FRM will improve the usability of BI tools and increase the benefits that end-users and organizations can gain from data resources. Our first research stage focuses on FRM based on assessment of previous usage and the associated value gain. We describe the development of such FRM, and the design of an experiment that will test the usability and the benefits of their integration. Our experiment incorporates value-driven usage metadata - a novel methodology for tracking and communicating the usage of data, linked to a quantitative assessment of the value gained. We describe a high-level architecture for supporting the collection, storage, and presentation of this new metadata form, and a quantitative method for assessing it.

Keywords: Business Intelligence, Data Warehouse, Decision Support Systems, Metadata.
1 INTRODUCTION

Data repositories, along with the information systems (IS) utilizing them are critical organizational resources. While in the past the primary goal of managing data was to enable business operations, recent years have witnessed a transition toward extended use of data for business analysis and decision support, as firms attempt to gain competitive advantage by developing advanced data-analysis capabilities [Davenport, 2006]. Our research investigates the integration of feedback and recommendation mechanisms (FRM) into business intelligence (BI) system, which support activities such as data analysis, managerial decision making, and business-performance measurement. We define FRM as textual, visual, and/or graphical cues that are embedded into front-end BI tools and guide the end-user to consider using certain data subsets and analysis forms. The working hypothesis of our study is that the integration of FRM into BI tools will improve their usability and increase the benefits that end-users and organizations can gain from data resources.

BI involves acquisition, interpretation, and analysis of data to support managerial decision making. The software market offers a plethora of commercial platforms for supporting BI activities. Such platforms typically offer a variety of presentation capabilities (e.g., tables, charts, statistics, and advanced analytics), rapid-development utilities, and administrative tools. BI tools permit different forms of data usage such as reports, spreadsheets, OLAP (On-Line Analytical Processing), digital dashboards, and data mining. This variety of presentation and analysis forms confers the flexibility to use the same data resource for supporting different analytic tasks and to adapt the presentation style to end-users’ capabilities and skills. BI solutions often use a data warehouse (DW) as an infrastructure. The DW stores historical data about past business behavior, patterns and trends, covering a broad range of business perspectives and activities. In a typical DW, datasets are imported from internal organizational IS, such as enterprise resource planning (ERP) systems [March and Hevner, 2005], and/or from external sources, such as commercial data vendors or the Internet [West, 2000]. The imported datasets are being cleansed, transformed, consolidated, and stored in a centralized repository. This DW infrastructure is then used for creating smaller databases (also known as data marts) that can accommodate different analytical needs and thus serve as a platform for supporting BI activities.

The number of firms engaged in DW/BI implementations and the volumes of data managed in DW/BI environments have grown immensely in recent years. The increasing popularity of DW/BI can be attributed to benefits such as gaining broad business coverage, leveraging data-collection investments, and shortening implementation cycles [Counihan et al., 2002; March and Hevner, 2005]. BI systems enable analytical data usage toward supporting important decisions such as evaluation of corporate strategies [Cooper et al., 2000], optimization of financial investments [West, 2000], and customer segmentation [Even and Shankaranarayanan, 2008]. Yet, exploiting DW/BI environments is challenging both technically, due to the many components and the complexity of configuration decisions involved [Shankaranarayanan and Even, 2004], and organizationally, due to the substantial managerial support and financial resources needed [Wixom and Watson, 2001]. Moreover, DW/BI design and configuration decisions are often associated with substantial cost-benefit tradeoffs [Even et al., 2006]. So far, despite the increasing popularity of the DW and the BI concept in recent years, so far these concepts have attracted only limited academic research aimed at the challenge of increasing the effectiveness of DW/BI utilization from the end-user’s perspective.

A major limitation of current BI systems is that the common end-user, in search of an answer to a business question, often finds complex DW repositories too difficult to navigate for reaching the right data, and BI tools too difficult to use for answering the question. Furthermore, it is even not uncommon for end-user to know neither the right business question to ask, nor the full range of capabilities offered by DW repositories and BI tools. This limitation exists in current BI solutions, more so with sophisticated interactive tools, which offer advanced visual and analytical capabilities for dynamic and flexible investigation of data, and less so with simple static tools, which offer “snapshot” views in forms such as pre-defined dashboards, reports, or charts. The former classes of tools are
geared toward addressing the needs of the data analyst and often require a high level of expertise and an in-depth understanding of the data-resource analyzed, whereas the latter are geared toward supporting the novice user. In terms of economic tradeoffs [Even at al., 2006], sophisticated interactive BI tools offer higher benefit potential, but are costlier in terms of licensing fees and learning curves, whereas simple static BI tools are easier to implement and learn, but offer limited capabilities, and hence, lower benefit potential.

We suggest that FRM capabilities can facilitate more effective and efficient navigation by helping to reveal undiscovered potential of unused data and analysis forms, and thus add business value. In the reminder of this paper, we present the concept of integrating FRM into BI tools and highlight a few possible approaches for generating them. We focus on FRM that are based on assessments of previous usage of the data resource, and the associated value gains. To generate this form of FRM, we propose a novel methodology for tracking the use of data resources, termed as value-driven usage metadata, which integrates in assessments of both the frequency of use and the value gained. We describe architecture for supporting the collection, storage, and presentation of this new metadata form and a quantitative method for assessing it. We then describe the design of an experiment that will test the usability and benefits of FRM integration. To conclude, we highlight the potential contributions of the new concepts that we present – the integration of FRM into BI systems, and the collection of value-driven usage metadata - and discuss directions for future research.

2 FEEDBACK AND RECOMMENDATION MECHANISMS (FRM)

In this study, we propose to integrate FRM capabilities into BI systems in a manner that would maintain simple and easy-to-learn BI functionality, while highlighting new usage directions with a high benefit potential. We define FRM as textual, visual, or graphical cues that are embedded into BI tools, providing the end-user with feedback on the actions that s/he has taken so far, and guiding him/her to consider further actions – e.g., to use certain data subsets and/or to apply certain analysis forms. Providing recommendations is a common tool used in commercial website for enhancing the end-user’s experience. Such recommendations can be generated by other users, or by automatic agents, and they have been shown to have great influence on the end-user’s decisions [Adomavicius and Tuzhilin ,2005]. We suggest that similar enhancement of BI systems may have important contribution to a better usage of the BI tools, and improve the decisions made.

Figure 1 offers a simplified illustration of integrating FRM capabilities into a BI tool. The original tool (on the left-hand side) lets the end-user navigate through sales data along certain dimensions (customer, location, date, etc.). This BI tool treats dimensions equally in the sense that it offers access to all dimensions and leaves navigation decisions to the user. In the FRM-enhanced version of the BI tool (on the right-hand side), navigation decisions are still left to the users, who are now provided with some additional visual cues. The cue that Figure 1 demonstrates, for example, is a color-coding that suggests giving higher navigation priority to certain dimensions. Obviously, there are other possible
forms for visualizing an FRM besides color-coding, such as textual or graphical pop-up messages and side bars. Such FRM forms could indicate, in addition to the actual recommendations, the level of confidence and relevance of each recommendation based on the parameters that construct it.

We suggest that FRM, when being integrated into BI tools, can facilitate more effective and efficient navigation. This, in turn, may help revealing undiscovered potential of unused data and analysis forms thereby may increase the effectiveness of DW/BI utilization and add business value. In a preliminary assessment we have identified, at a high level, a few possible methods for constructing FRM:

(a) **Value-driven usage tracking:** Information resources contribute value through usage and experience. In the DW/BI context, the value can be conceptualized, for example, as an objective measure of usage success (e.g., in terms of revenue gained and/or costs saved), or subjectively via an assessment that reflects user satisfaction and willingness to pay [Ahituv, 1981]. Quantitative assessments of the value associated with the use of data resources have been applied to optimize data processes [Ballou et al., 1998], configure DW datasets [Even et al., 2007], and develop data quality metrics [Even and Shankaranarayanan, 2008]. The latter study has highlighted the importance of recognizing inequality in the value of data, suggesting that data objects (e.g., tables, attributes, and records) may vary significantly in their value contribution. Further, by evaluating a large real-life data resource, that study shows that quantitative assessments of inequality (e.g., Gini’s index), have important implications for key data management decisions, such as the prioritization of data quality improvement efforts. We suggest that tracking data usage and the associated value can be used to construct FRM in BI tools, toward directing data analysis and exploration, and improving decision outcomes – this by providing users with feedback on the outcome of their own usage, as well as an opportunity to benefit from learning how other users have gained value from using the same data.

(b) **Task and user characteristics:** The same DW/BI environment can be used to support a plethora of business processes and tasks, each with very different data usage needs. FRM capabilities can take into account such needs, by creating either task profiles that capture specific task characteristics or by asking expert users to identify certain data elements or analysis results that are more useful and relevant for a given task. For example, referring to the illustration in Figure 1, an FRM-enhanced BI tool could direct the end-user to slice the sales data along certain dimensions (e.g., customer and product) upon designing a promotion campaign, versus focusing on other dimensions (e.g., location and date) upon optimizing distribution policies. Likewise, an FRM-enhanced tool could recommend using certain presentation styles or navigation forms that can be driven by individual preferences, e.g., presentation in textual style versus graphical visualization, or navigating data systematically and incrementally versus using heuristics. Such preferences can either be explicitly identified by asking the user to state his or her preferences or implicitly inferred via a learning process, tracking and understanding the user’s interaction and identifying more-often used data and presentation formats.

(c) **Data mining:** FRM capabilities can also be driven by analysis of the data using algorithmic data-mining techniques. Data mining algorithms typically explore the information contained in data sources automatically and provide end-users with quantitative assessments and to enable them to extract and evaluate knowledge from this information. Data mining can suggest alternatives to decisions and actions that are about to be taken and allow users to re-consider these decisions and actions. Notably, some data-analysis tools include functionality that actively guides users through large data repositories, based on some statistical analysis of the data resource (e.g., Bissantz-DeltaMaster, [http://www.bissantz.com/deltamaster](http://www.bissantz.com/deltamaster)). A possible drawback of using statistical analysis and data mining methods to derive FRM is the risk that the recommendations, in this case, will be based solely on the data itself, not taking into account the context in which the data is used. This limitation is addressed to an extent by the two former methods, and should be considered when further researching the use of data-mining techniques for developing FRM.

Initially, our research will focus on the first approach for generating FRM (denoted (a) above), while the exploration of other two will be deferred to later stages. This form of FRM will be based on a novel methodology for tracking the usage of data resources, described in the following section.
3 VALUE-DRIVEN USAGE METADATA

Data environments are often described as a complex manufacturing process, consisting of interconnected acquisition, processing, storage, retrieval and usage stages [Ballou et al., 1998]. These processes can be conceptualized as having two high-level stages – data administration versus data consumption (Figure 2) - each associated with different stakeholders, goals, motivations and tasks [Lee et al., 2004]. Data administration addresses technical aspects – providing the ICT capacity needed to store and process data (e.g., hardware, databases, and back-end processes), and the tools for implementing information products. Conversely, data consumption would seek to transform information products into business value, through their usage. As the typical goal of data consumers is gaining business benefits and increasing profitability, they would be more focused on the value gained by effective use of data resources and less on the technical aspects associated with managing them.

Usage metadata - tracking the usage of data objects (e.g., tables, attributes, and records) and applications - has been identified as an important form of metadata [Shankaranarayanan and Even, 2004]. Usage tracking utilities are offered by some specialized commercial solutions and, to an extent, by DBMS and BI platforms. We term the common approach implemented by today’s solutions as frequency-driven usage metadata (the left-hand side of Figure 2). This approach is based on tracking queries and, by analyzing them, identifying the data objects being most-frequently used. The assumption underlying this approach is that frequent usage reflects higher importance. Accordingly, usage-tracking results may affect the configuration and the administration of data resources - e.g., transferring less frequently-used data to archives, and/or giving it lower quality-improvement priority.

Illustrative Example (part 1) - Assessment of Frequency-Driven Usage Metadata

<table>
<thead>
<tr>
<th>#</th>
<th>Customer</th>
<th>Gender</th>
<th>Income</th>
<th>Children</th>
<th>Status</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Abraham</td>
<td>Male</td>
<td>High</td>
<td>0</td>
<td>Single</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Sarah</td>
<td>Female</td>
<td>Low</td>
<td>1</td>
<td>Married</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Isaac</td>
<td>Male</td>
<td>Medium</td>
<td>2</td>
<td>Married</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Rebecca</td>
<td>Female</td>
<td>Low</td>
<td>0</td>
<td>Single</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Jacob</td>
<td>Male</td>
<td>Medium</td>
<td>3</td>
<td>Married</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Lea</td>
<td>Female</td>
<td>High</td>
<td>2</td>
<td>Married</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>Rachel</td>
<td>Female</td>
<td>Low</td>
<td>4</td>
<td>Single</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Frequency</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Queries</td>
<td>WHERE condition</td>
<td>Attributes Used</td>
<td>Records Retrieved</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gender = ‘Male’ and Children &gt; 0</td>
<td>Gender, Children</td>
<td>[3], [5]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gender = ‘Female’ and Children &lt; 3</td>
<td>Gender, Children</td>
<td>[2], [4], [6]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gender = ‘Female’ and Status = ‘Married’</td>
<td>Gender, Status</td>
<td>[2], [6]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Income = ‘High’</td>
<td>Income</td>
<td>[1], [6]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Illustrative Example (part 1) - Assessment of Frequency-Driven Usage Metadata
To illustrate the creation of frequency-driven metadata, we use a simplified example above (extended later). The table in this example is used by marketing associates to decide which customers will be approached when promoting a new product. An associate would use a BI tool to investigate previous sale transactions, and the tool will generate queries directed to the customers tables, such as those demonstrated, to specify the subset of customers that will be targeted. Each query can be analyzed to detect which records and attributes are used to specify the selection (e.g., by parsing the WHERE clauses in the SQL statement) and, accordingly, the frequency of usage can be calculated. Frequency-driven metadata collection may provide important inputs to the data administrator, toward improving system design and prioritizing administration efforts. It is common in databases, as highlighted by the example above, that some records and attributes are accessed more frequently than others. In a larger real-world databases, this differentiation may lead to a decision to grant the more frequently used records and attributes a higher priority in terms of data quality maintenance – i.e., watch these database objects more closely, detect and correct defects, and make sure to keep them up-to-date.

While seeing the merits of collecting metadata on usage frequency for data administration, we question - does it truly address the needs of data consumers? One could argue that, to an extent, frequent usage reflects higher significance of certain data components to data consumers; hence, higher value-contribution potential. On the other hand, we suggest that frequent usage may reflect certain stagnation and a tendency to "dig into the same well" - re-using certain data subsets repetitively, while possibly ignoring unused subsets with high contribution potential. Therefore, a potential risk with basing data management decisions solely on frequency-driven metadata (e.g., due to a removal of data that is less-frequently used from the active repository into an archive) is a possible loss of opportunity to benefit from records and attributes that data consumers have neglected to use so far, which may permit new forms of data usage.

There is possibly no "clear cut" answer to this question, as it largely depends on the business context and the usage tasks. However, we suggest that important insights can be gained from tracking and considering not only actual data usage, but also the associated value gains. The benefits gained from the use of information products have been conceptualized as *utility* [Ahituv, 1980]. Utility assessments have been used to optimize the configuration of data processes and resources [Ballou et al., 1998; Even et al., 2007] – tasks typically associated with *data administration*. We suggest that, beyond the benefit offered to data administration, collecting quantitative assessment of the business-value gained as a form of metadata can improve *data consumption* as well. Business value can be measured, for example, in terms of decision outcomes (e.g., production increase, customers’ purchase intent), revenues and profitability. Organizations capture such value measurements today, but rarely link them to the data resources and the decision-support tools that were used in the process of value generation.

Value-driven collection of usage metadata (Figure 2) aims at establishing such a link. To demonstrate this approach, we have successfully implemented a working prototype of a module that captures and stores value-driven usage tracking as a metadata layer. In its base, the module applies a similar approach to the one described earlier for collecting frequency-driven metadata - capturing the queries directed at a data resource, and parsing them into specific components. However, the module also collects different types of value measures (e.g., throughput, performance, and income), which are associated with a specific decision task. In certain cases, value assessments can be based on the same data resource (sale transactions, for example, can often be linked to a specific marketing campaigns that were based on a certain analysis of previous sales). In other cases – such assessments may use other information resources such as CRM and accounting systems. The module associates the value with certain decision tasks and then, through a mechanism of inference (e.g., by comparing the user-name and the time stamp), to the queries that have supported each tasks. Establishing this link between decision tasks and the underlying queries permits the creation of integrated metadata that associate business value with specific data components. An API (Application Programming Interface) can provide this form of usage metadata upon demand through function calls. Such metadata can be integrated into front-end tools, enhance the presentation, and communicate important information on the frequency of usage and on the associated value to both data consumers and administrators.
Once the link between decision tasks and queries is established, different methods can be considered for attributing value to specific data objects. For illustration, we describe here a relatively simple method, which assumes that value is attributed to the last in a sequence of queries that support a decision task. We assume that to support a certain decision, users query repetitively a certain tabular dataset with \( N \) records indexed by \([n] \) and \( M \) attributes indexed by \([m] \). We consider \( Q \) queries indexed by \([q] \), each associated with a business value \( V^q \). The binary indicator \( R_{n}^q \) indicates whether record \([n] \) was retrieved by query \([q] \) (\( R_{n}^q = 1 \)), or not (\( R_{n}^q = 0 \)). Similarly, \( R_{m}^q \) indicates whether attribute \([m] \) participated in query \([q] \) or not. The value of a certain query \((V^q)\) is attributed between the participating data items, using a certain value-attribution function \(V_{n,m}^q = u(V^q, R_{n}^q, R_{m}^q)\), such that \( V^q = \sum_{n,m} V_{n,m}^q \). For simplification, we use here an equal attribution of value among all participating data items. Accordingly, the overall value of a certain data items \( V_{n,m}^q \) is given by:

\[
(1) \quad V_{n,m}^q = \sum_{q=1}^{Q} V_{n,m}^q = \sum_{q=1}^{Q} u(V^q, R_{n}^q, R_{m}^q) = \sum_{q=1}^{Q} V^q / \left( \sum_{n=1,N} \sum_{m=1,M} R_{n}^q R_{m}^q \right),
\]

- \( Q, q \) - The number of queries performed, and the corresponding index, respectively
- \( M, N \) - The number of attributes (indexed \([m] \)) and records (indexed \([n] \)), respectively
- \( V_{n,m}^q, R_{n}^q, R_{m}^q \) - Query \([q] \) value, its attribution to data item \([n,m] \), and the attribution function used
- \( R_{n}^q, R_{m}^q \) - Binary indicators of the participation \((=1)\) of record \([n] \) and attribute \([m] \) in query \([q] \)

To demonstrate the value allocation described above, we extend the previous example. We assume that each query has led to a certain promotion campaign in which a group of customers has been approached. Some customers may have responded to the campaign by making certain purchases, and the overall value attributed to a query is the total purchase amount. As illustrated, this value proxy may significantly vary among queries. We now use the allocation (Eq. 1) to assess the relative value of each data object. As illustrated by color-coding – some records, and attributes may turn out to have significantly higher value than others, and the value attribution “map” may look significantly different than the one when basing the attribution of usage frequency.

**Customers**

<table>
<thead>
<tr>
<th>#</th>
<th>Customer</th>
<th>Gender</th>
<th>Income</th>
<th>Children</th>
<th>Status</th>
<th>...</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Abraham</td>
<td>Male</td>
<td>High</td>
<td>0</td>
<td>Single</td>
<td></td>
<td>1000</td>
</tr>
<tr>
<td>2</td>
<td>Sarah</td>
<td>Female</td>
<td>Low</td>
<td>1</td>
<td></td>
<td></td>
<td>510</td>
</tr>
<tr>
<td>3</td>
<td>Isaac</td>
<td>Male</td>
<td>Medium</td>
<td>2</td>
<td>Married</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>Rebecca</td>
<td>Female</td>
<td>Low</td>
<td>0</td>
<td>Single</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Jacob</td>
<td>Male</td>
<td>Medium</td>
<td>3</td>
<td>Married</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Lea</td>
<td>Female</td>
<td>High</td>
<td>2</td>
<td>Married</td>
<td>1510</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Rachel</td>
<td>Female</td>
<td>Low</td>
<td>4</td>
<td>Single</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>515</td>
<td>2000</td>
</tr>
</tbody>
</table>

**Queries**

<table>
<thead>
<tr>
<th>WHERE condition</th>
<th>Attributes Used</th>
<th>Records/Value</th>
<th>Total Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender = ‘Male’ and Children &gt; 0</td>
<td>Gender, Children</td>
<td>[3], [5]</td>
<td>100</td>
</tr>
<tr>
<td>Gender = ‘Female’ and Children &lt; 3</td>
<td>Gender, Children</td>
<td>[2], [4], [6]</td>
<td>30</td>
</tr>
<tr>
<td>Gender = ‘Female’ and Status = ‘Married’</td>
<td>Gender, Status</td>
<td>[2], [6]</td>
<td>1000</td>
</tr>
</tbody>
</table>

**Illustrative Example (part 2) - Assessment of Value-Driven Usage Metadata**

We can potentially gain important insights by analyzing the value distribution, along with the assessment of frequency of use. For example, the **Income** attribute, which was not frequently used, is associated with the highest value, while the **Children** attribute, which was more frequently used, is associated with lower value. Insights as such can be transformed into valuable recommendations for a marketing associate the next time s/he plans to run a similar campaign. The decision value of each of the participants is being saved in a metadata repository. Using the API, a BI tool, which was designed to access to the value-driven usage metadata can now demonstrate value differentials and distribution, toward improving the decisions made, as demonstrated in the experimental design described next.
4 EXPERIMENT DESIGN

This section describes the design of a lab experiment, currently under final preparation stages, that tests the integration of FRM based on previous usage and the associated value.

4.1 Research Model and Working Hypotheses

The first research stage will be directed by the theoretical model shown in Figure 3. Some model variables will be measured with the test experiment tool described in the following section, while others will be assessed using a previously-tested questionnaire.

**Dependent Variable – Performance:** The dependent variable, reflecting the ability of a user to effectively perform a task with tool support, will take one of two forms:
- Objective – actual decision outcome and time it takes to complete the decision task
- Subjective – perceived usefulness and ease of use

Previously tested models, such as TAM (technology acceptance model [Davis, 1989]), have suggested that a higher sense of usefulness and ease of use increase the likelihood of user acceptance. While this study intends to focus mainly on performance, the experiment described below will permit assessing acceptance and validating the anticipated link between performance and acceptance.

**Independent Variables:** The independent variables will be Experience and FRM inclusion:

a) **Experience** can be measured in terms of:
- Learning curve – the time a user spends using and mastering the tool
- Familiarity – the extent to which the user has previously used similar tools in the past

It is reasonable to assume that an experienced user (in terms of learning curve and/or familiarity) will perform better than a non-experienced user; hence,

- **H1:** Usage Experience positively affects Performance

b) Our key assumption is that the inclusion of an FRM will offer a major improvement in the usability of BI tools and therefore in user performance; hence,

- **H2:** FRM inclusion positively affects Performance

As discussed in the previous section, we suggest that value-driven collection of usage metadata is superior to frequency-driven collection; hence:

- **H2a:** The Performance effect of FRM that are based on value-driven metadata will be superior to the effect of FRM based on frequency-driven usage metadata alone
It is reasonable to assume a possible synergistic effect between the two independent variables, i.e., that the overall effect of Experience and FRM inclusion is higher than the effect of each alone. Hence,

- **H3: The interaction effect between Experience and FRM inclusion is positive**

**Moderating Variables:** It is reasonable to assume that certain user characteristics will moderate the effect of Experience and FRM inclusion on Performance. The moderating variables that will be tested are Motivation, the user’s motivation to perform well, and Expertise, the extent to which the user is knowledgeable in the particular task domain. Studies (e.g., Siegel and Watts-Sussman [2003]) have shown Motivation (or involvement) and Expertise to have moderation effects on the usefulness of information resources and hence on their acceptance and adoption. Hence,

- **H4: The greater the user’s Motivation, the more Experience affects Performance**
- **H5: The greater the user’s Motivation, the more FRM inclusion affects Performance**
- **H6: The greater the user’s Expertise, the more Experience affects Performance**
- **H7: The greater the user’s Expertise, the more FRM inclusion affects Performance**

**Control Variables:** The experiment will control for a few additional variables - age, gender, language fluency, and possibly others.

### 4.2 Experiment Procedures and Tool

The model and the derived hypotheses will be tested in a laboratory setting. In the planned experiment, all participants will be asked to perform a certain decision task repetitively, aided by a BI tool. The decision outcomes, as well as the actual usage of the tool and the data resources will be tracked and measured. This will enable data collection that will allow measuring some of the variables (as described later in Table 1). In addition to tracking decision outcome and actual usage, users will be asked to complete a questionnaire, which will enable data collection on remaining variables.

Due to space limitations, we do not describe here in details all the experiment preparation procedures, but rather explain the principles that guide its design.

**The decision task:** the participants will act as marketing associates on behalf of a firm that offers a certain product or service to its customers (e.g., a vacation package). To decision will be aided by a large database that includes two main tables:

- **A list of customers,** in which each associated with a given set of attributes (e.g., Income, Gender, Marital Status, and Number of Children). Based on the mix of attribute value – each customer \([t]\) will be associated with a set of likelihood numbers \(P_{t,z}(z = 0, 1, 2, \ldots)\) of purchasing \(z\) units within a given time period, such that \(\sum_{z} P_{t,z} = 1\), and \(Q_{t} = \sum_{z} z P_{t,z}\) is the mean number of purchases.

- **Purchases transactions,** based on the purchase likelihoods defined per customer. A random generator will produce a large number of purchase transactions for a broad period of time.

Given access to this database, the participants will be asked to choose a customer segment that will be targeted. Approaching a customer and offering him/her a promotion has a given cost (e.g., the mail delivery fee, or the time needed for a phone call); hence, the larger is number of customers approached – the higher is the cost. Each customer is associated with certain likelihood to purchase a certain quantity of the service and, accordingly, the overall decision value is defined as:

\[
(2) \quad V = \sum_{i=1}^{T} I_{i} (SQ_{i} - C), \quad \text{where}
\]

- \(V\) - The overall decision
- \(T\) - The number of customers (indexed \([t]\))
- \(S, C\) - The revenue per service item sold, and the promotion cost per customer, respectively
- \(Q_{i}\) - The expected number of item that customer \([t]\) will purchase \((Q \geq 0)\)
- \(I_{i}\) - A decision whether to include customer \([t]\) in the promotion campaign \((=1)\) or not \((=0)\)
A decision to include a customer may increase revenue, but at a cost. Obviously, a decision maker would prefer to include only customers for which the expected revenue is higher than the cost (i.e., only $t$ for which $SQ_t > C$). The expected quantity $Q_t$ is defined in advance, but not exposed to the decision maker explicitly. The user will be asked to infer which customers are likely to purchase the service by observing customers’ past purchases. Moreover, the decision maker will not select specific customers, but rather will be asked to define customer segments. The segment definition will be based on a given set of customer attributes and a selection of a certain criterion per attribute (e.g., “All high-income married male customers, with 2 children or more”), where the user may choose to avoid defining selection criteria in certain attributes.

To aid the decision, each participant will be provided with a BI tool, such as the one illustrated by Figure 1. The BI tool will permit exploring past transactions, analyze purchase activity, and determine the revenue associated so far with each customer segment. The tool presents the distribution of certain measures (Revenues associated with past transactions) along certain customer attribute. The visual display imitates a decision tree. Starting at the high-level node, which reflects the entire population of customers, the user may choose an attribute (e.g., Children) along which he wishes to segment the data. For a given attribute value (e.g., Children = 2), the user may choose to segment the data along another attribute (e.g., Status), and so on. Based on the different customer segmentations that are explored by using the BI tool – the user finally selects the customer segments that will be targeted. Once the selection is made – the overall value of the selection is calculated (Eq. 2), the value is attributed among the different attributes and records (Eq. 1), and the attribution is saved in the value-driven usage metadata module. As the experiment participants keep performing the decision tasks repetitively – the value-driven usage metadata is accumulated and enhanced.

The left-hand side of Figure 1 illustrates the BI tool in its basic form, which does not include FRM. The enhance form, illustrated in the right-hand side of Figure 1, includes certain FRM enhancements – indication of the total value and the value distribution (a variance measurement) associated with the different attributes, at each node. The recommendations change dynamically, depending on the node that the user selects. The FRM enhancements are based on the usage metadata that was accumulated while participants keep performing the decision task repetitively. To help testing the hypotheses, the group of participants will be divided into a few sub-groups, and some variability will be created in the tasks that each sub-group is asked to perform (Table 1).

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>How the hypotheses will be tested</th>
</tr>
</thead>
</table>
| H1: Experience affects Performance | • Participants will be asked to perform the same task repetitively  
• Some participants will provided with the same BI tool in all sessions  
• Participants will be asked about past experience with similar tools |
| H2: FRM affects Performance | • Initially, participants will perform the task using BI tools with no FRM  
• Later, some participants will be offered BI tools with FRM  
• Some participants will perform the task with no BI support at all |
| H2A: Value-Driven FRM are superior to Frequency-Driven FRM | • Some participants will be provided with FRM enhancement based on value and frequency assessments, while others will be provided with FRM enhancement based on frequency alone |
| H3: Experience-FRM interaction | • Experience/FRM – same as the above  
• Certain statistical regression methods permit testing interaction |
| H4, H5: Motivation moderation | • Experience/FRM – same as the above  
• Some participants will be offered performance-based compensation |
| H6, H7: Expertise moderation | • Experience/FRM – same as the above  
• Participants will be recruited from different expertise populations |
| Control | • The test will validate assumption that the control variables have no substantial effect in the given setting |

Table 1. Hypotheses Testing
5 CONCLUSIONS

Our research investigates the integration of feedback and recommendation mechanisms (FRM) into BI tools. The working hypothesis that guides our study is that the integration of FRM into BI tools will improve their usability and increase the benefits that end-users and organizations can gain from data resources. We have described an experiment, currently under preparation, for testing the usability and the benefits of such integration in terms of improving decision-making processes. We see this experiment as a first “proof of concept” step of testing the FRM-integration idea, toward gaining insights on its usability and benefits. The controlled lab environment, in which we intend to apply the test, will permit a more precise data collection on usage patterns and value generation – what is often hard to achieve in real-world environments. Nevertheless, testing FRM integration in real-world environments would be an important follow-up step in furthering this line of research.

Another key contribution of our study, which links to the previous, is the introduction of a novel approach for usage tracking in data environments. This approach suggests that integrating quantitative assessments of usage-frequency together with the associated value gained may offer substantial benefits to data administration and consumption. Joint frequency and value assessments can help identifying unused data subsets with high value-contribution potential, may highlight flaws with repetitive use of data and, consequently, motivate new usage forms. Further, value assessment can direct design decisions, and help prioritizing data maintenance efforts. Relying on usage frequency alone might promote usage stagnation and loss of opportunity to gain new forms of benefits. Complementing frequency assessments with value assessments may help “closing the loop”, in terms of providing feedback based on usage performance, and reducing the potential risks. First, value allocation gives higher weight to past usages with high contributions. Second, it can reflect variability in the importance of different subsets depending on the usage context. Lastly, it can help detecting data subsets with high contribution potential that have not been frequently used. Obviously, future extensions to our study will need to address some key limitations of this approach:

(a) Quantifying value – organization maintain performance measurements (e.g., productivity, income, and profitability) that can be possibly linked to decision tasks. However, decision performance may depend on other resources such as human knowledge and financial assets. Further, the value depends on the usage context, and value assessment for a certain type of usage tasks does not necessarily apply to others. Further, value is time-dependent, as data that can be used effectively at a certain point of time, might become obsolete later. We hasten to say that the value-allocation methodology, which we apply in this study, appears to be a better fit to operational environments in which decision tasks have a high degree of repetition, and causal relations between data usage and business performance are easier to establish. Promotion-campaign management, such as in our illustrative example would be a good representative for this type of decision-making. Financial-investment decisions would be another example for data-driven decisions, in which outcomes are measurable (e.g., the change in the value of the financial asset) and linkable to the data resources being used. Conversely, quantifying the value of decision outcomes might turn out to be more challenging in strategic decision scenarios, which are not repetitive in nature and often relay on information sources other that organizational data repositories.

(b) Linking value to specific queries - performance assessments are rarely linked explicitly to the data resources and tools used. Our preliminary prototype includes inference mechanisms for creating implicit links – e.g., based on the user name, and/or time proximity. Obviously, implicit links cannot be absolutely precise and might bias the value allocation significantly. Establishing explicit links will require stronger metadata integration between systems and, likely, redesign of data environments (e.g., joint codes that link each decision task and queries). One could question whether or not making such a high investment in redesigning data environments and BI tools would justify the benefits gained.

(c) Attributing value to specific data objects – the attribution system has critical impact on the results. Our prototype attributes value only to the last query in the sequence that generated the decision, and distributes the value equally between all the data that were retrieved. A different allocation method
may consider, for example, spreading the usage value along all queries and/or consider possible interactions among attributes – hence, unequal allocation.

Finally, we would suggest that future extensions of this study should further explore links to the research of recommender systems. Recommender systems are common in web-based user interfaces (e.g., rating systems in E-commerce sites), but less so in BI tools. Adomavicius and Tuzhilin [2005] identify the need to incorporate contextual information into the recommendation process. As value assessments depend on the context in which data is used [Even and Shankaranarayanan, 2008], we would suggest that the value-driven metadata approach may help such incorporation.

Acknowledgments

This project was supported by a grant from Microsoft R&D, Israel.

We wish to thank Prof. Nava Pliskin from the Department of Industrial Engineering and Management at Ben-Gurion University of the Negev for her advice and support.

References


WHY DO PEOPLE BUY VIRTUAL ITEMS IN VIRTUAL WORLDS? AN EMPIRICAL TEST OF A CONCEPTUAL MODEL

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0012.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Virtual world, Social Networking (e.g. Facebook, second life), Structural Equation Modeling, Partial Least Squares</td>
</tr>
</tbody>
</table>
WHY DO PEOPLE BUY VIRTUAL ITEMS IN VIRTUAL WORLDS? AN EMPIRICAL TEST OF A CONCEPTUAL MODEL

Yue Guo, Norwich Business School, University of East Anglia, Norwich, United Kingdom, yue.guo@uea.ac.uk

Stuart Barnes, Norwich Business School, University of East Anglia, Norwich, United Kingdom, stuart.barnes@uea.ac.uk

Abstract

In the past few years, virtual worlds – such as Second Life, World of Warcraft and RuneScape – have demonstrated the potential to be a promising online business model. Millions of paying users around the world now participate in virtual worlds and trade virtual items with each other. However, little empirical research has been conducted into players’ purchase behaviour in virtual worlds. To address this research gap, the current study develops and tests a conceptual model of purchase behaviour in the context of virtual worlds. An online survey was conducted within Second Life (n=250) and SEM-PLS (Structural Equation Models by Partial Least Squares) was used to confirm the conceptual model. Among the statistically significant paths found in the conceptual model, effort expectancy, performance expectancy, perceived value, customisation, habit, advancement and enjoyment all have strong impact on virtual world residents’ purchase intention. The study rounds off with a discussion and conclusions, highlighting implications for practice, research limitations and promising areas for future research.

Keywords: Structural Equation Model; PLS; Virtual World; Purchase Behaviour; Second Life.
1 INTRODUCTION

In the past few years, virtual worlds have demonstrated the potential to be a novel online transaction environment compared to the Web-based context; thousands of participants may not only interact with one another, but also buy and sell virtual items in a virtual world. At present, popular 3-D virtual worlds are mainly divided into two categories: game-oriented virtual worlds (e.g. World of Warcraft, RuneScape and EverQuest) and socially-oriented virtual worlds (e.g. Second Life, There and HiPiHi) (Bainbridge 2007, Guo & Barnes 2007). The major difference between the two types is that a game-oriented virtual world provides a story-lined scene and fixed contents while a socially-oriented virtual world attempts to provide near unlimited freedom in a simulated world in which participants use their imagination to create desired world contents. Regardless of virtual world type, typically virtual items can be traded for real world currencies within virtual worlds or via other third-party exchange platforms. DFC Intelligence (2007) estimated that virtual worlds will produce US$13 billion and over 40% of revenues will be attributed to trading virtual assets by 2012. In the face of such a large quantity of virtual item transactions, understanding the latent psychological processes that typically induce players’ purchase behaviour in virtual worlds becomes of obvious importance.

The majority of previous studies in this area have focused on legal issues with regard to virtual item transactions (e.g. Lastowka & Hunter 2004, Lederman 2007) or have explored the impact of virtual worlds on the real world from a purely economic point of view (e.g. Castronova 2001, 2002). Although several prominent explanatory models of behaviour-expectation, such as the Technology Acceptance Model (TAM) (Davis 1989), have been developed and applied to explaining online shopping behaviour or the adoption of information systems (IS), virtual item purchase behaviour in virtual worlds is likely to be quite different and requires new conceptualisations. Moreover, we also need to examine whether factors related to online purchase behaviour still exert considerable influence in the context of virtual worlds. A few recent studies have begun to notice virtual item transactions and offered new insights into players’ purchase behaviour in virtual worlds (e.g. Yee 2005, Lehdonvirta 2005, Oh & Ryu 2007, Nojima 2007). Nevertheless, these studies are limited in the sense that they are only based on a descriptive or theoretical analysis of virtual item transactions and they lack the rigour of empirical analysis. Their research findings need to be tested by further quantitative research using large-scale samples before making generalisations.

In order to detect factors determining purchase behaviour in virtual worlds for theory building, this research empirically develops and tests a theoretical model of virtual item purchase behaviour in virtual worlds based on established behaviour-expectation theories and previous exploratory research into virtual item transactions. The research employs a survey in the virtual world of Second Life (SL) and analyses the research model by way of a structural equation modelling approach. The outcome from this research is an extension to existing behavioural theories via new constructs validated in the virtual world domain.

The paper is organised as follows. In the next section we examine the salient literature on virtual item purchase behaviour. This is followed by an explication of the research model. Section four describes the study design and method adopted in the research, and the results are presented in section 5. Finally the paper rounds off with a discussion and conclusions, including limitations and implications for future research and practice in this very new area of investigation.

2 REVIEW OF EXTANT LITERATURE ON VIRTUAL ITEM PURCHASE

Although virtual item purchase behaviour is a relatively new transaction phenomenon and therefore research is limited, the prediction of purchase behaviour in the context of Web sites has been long studied by previous literature using prominent behaviour-expectation models, such as the Theory of
Reasoned Action (TRA) (Fishbein & Ajzen 1975), the Theory of Planned Behaviour (TPB) (Ajzen 1991) and TAM. These established theoretical models have shown high reliability and the predictive power of social behaviour across a variety of contexts. We thoroughly reviewed previous studies published since 1992 on the adoption of IS especially related to Web-based shopping behaviour. According to the literature, the most common construct variables for Web-based online purchase behaviour are performance expectancy, effort expectancy, social influence, perceived enjoyment, IS quality, facilitating conditions, trust and risk. Guo and Barnes (2007) first developed a theoretical model for explaining purchase behaviour in virtual worlds based on five prominent theoretical models such as, TAM, TPB and UTAUT (Venkatesh et al. 2003). Guo and Barnes (2009) further revised the preliminary theoretical model using focus groups and offered a deep insight into the whole psychological processes of virtual item purchase behaviour in virtual worlds from three perspectives: motivators for pursuing advanced virtual items, the reasons for using virtual world trading platforms to obtain desired virtual items, and factors with regard to making a specific purchase decision. All of the above factors were introduced into the revised theoretical model based on focus group findings—except for risk. On the one hand, many kinds of risk are not as relevant to virtual items, such as performance risk (Guo & Barnes 2009); on the other hand, the influence of risk on purchase behaviour in the context of virtual worlds is mediated by trust.

Yee (2005) focused on investigating players’ motivations for participating in massively multiplayer online role-playing games (MMORPGs). Yee’s research findings showed that customisation and achievement were players’ two major motivations for pursuing advanced virtual items in virtual worlds. Oh and Ryu (2007) analysed the item-selling based payment model used by two virtual worlds, “Kart Rider” and “Special Force”. Their research explored some factors stimulating players to pursue valuable virtual items in virtual worlds including improving game capability and quickly upgrading levels. Lin and Sum (2007) pointed out that virtual world operators may not only charge subscription fees, but also obtain significant profits by selling functional or decorative virtual items. A functional virtual item may increase a character’s overall ability to compete (e.g. advancing character levels or moving one’s avatar more quickly) while a decorative virtual item helps players to customise the appearances of their characters.

Based on Guo and Barnes’ (2009) research findings, we realise that it is necessary to distinguish between general purchase behavioural intention and special purchase behavioural intention in this research. With regard to real product/service purchase behaviour, general purchase behavioural intention refers to the subjective probability of one’s engagement in purchasing desired products/services through the Web-based purchase manner while special purchase behavioural intention can be viewed as an immediate indication of an individual's readiness to perform a given purchase from the specific Web site for a specific product or service. General purchase behavioural intention demonstrates an individual’s intention to purchase desired goods/services through the Web-based manner rather than other manners, such as offline stores, telephone and catalogue, even if the individual still has no idea where he/she will make a purchase (i.e. identifying a specific Web site). Users possibly form their special purchase behavioural intentions and immediately result in an actual purchase when they browse an online shopping Web site. Most previous studies on Web-based shopping behaviour did not explicitly distinguish between general purchase behavioural intention and special purchase behavioural intention because an individual’s belief about shopping in a specific Web site and about the Web-based shopping manner are the same in most cases. In other worlds, factors related to users’ a choice between the Web-based purchase manner and other purchase manners are the equivalent of factors identifying a specific shopping Web site, such as, convenience, price, service quality, risk and so on. For example, if a user perceives that using a specific Web site is convenient, he/she usually believes that the Web-based shopping manner is more convenient than other purchase manners. Similarly, if users consider that the Web-based shopping manner is more convenient than other purchase manners, they usually have enjoyed the convenience of shopping in a specific Web site.
However, the situation in virtual worlds is quite different and more complex because players have other virtual item obtaining manners in addition to the in-world purchase manner. For example, players may get desired virtual items in virtual worlds through in-world non-purchase manners, e.g. fighting with non-player characters (NPCs), which may randomly drop virtual items, or by finishing quests in game-oriented virtual worlds, and via self-design in socially-oriented virtual worlds, in addition to buying from other players. Furthermore, buying from other players can be achieved through Web-based platforms or face-to-face approaches (i.e. out-of-world purchasing manners). Compared to virtual items, perhaps purchasing in the real world (i.e. using real money) is the only method to get desired real products/goods and very limited in-world trading platforms have been used for purchasing real products/goods so far (i.e. using virtual currencies).

With regard to virtual item purchase behaviour, general purchase behavioural intention refers to the subjective probability of one’s engagement in obtaining desired virtual items through the purchase manner in virtual worlds while special purchase behavioural intention is only an immediate indication of an individual’s readiness to perform a given purchase from a specific vendor for a specific virtual item in virtual worlds. First, general purchase behavioural intention in the context of virtual worlds is determined by an individual’s overall belief that in-world shopping manner has greater relative advantages and lower complexity than other in-world non-purchase manners and out-of-world purchase manners. Second, special purchase behavioural intention is mainly determined by a player’s belief that a specific virtual world vendor has relative advantages than others (e.g. good reputation and reasonable price). With respect to virtual world transactions, players still have to make a series of specific decisions (e.g. identifying a specific vendor) prior to actual purchase even if they have formed a strong intention to obtain desired virtual items through buying in virtual worlds (i.e. general purchase behavioural intention). Consider the following example of a player in a game-oriented virtual world for further illustrating the differences between general purchase behavioural intention and special behavioural purchase intention. A player has planned to obtain advanced weapons for improving a character’s power through buying from other players in the virtual world, which refers to a general purchase behavioural intention. However, he or she has to identify a specific purchasable virtual item based on his/her personal conditions prior to making an actual purchase from enormous range of alternatives because every virtual item has various attributes (e.g. strength, level requirements for use or durability) even though two virtual items may be of the same item type, e.g. a virtual sword. Moreover, he or she also needs to identify a suitable vendor for the desired virtual items. In case the player has formed intention to purchase a specific virtual item from a vendor, we claim that he or she has formed a special purchase behavioural intention at this point.

In this research, we only focus on capturing key determinants influencing players’ general intention to purchase virtual items in virtual worlds rather than special purchase intention with a focus on specific purchase decisions. A clear distinction between general purchase behavioural intention and special purchase behavioural intention with a focus on players’ specific purchase decisions can keep our conceptual model parsimonious as much as possible. A parsimonious model is easier to comprehend and should obtain a high degree of generality. Factors identified by previous literature are not included this research if they relate predominantly to players’ special purchase intention, e.g. trust (Guo & Barnes 2009, McKnight 2002). It is worth noting that the term “purchase behavioural intention” used in later sections pertains primarily to general purchase behavioural intention unless otherwise noted.

3 FORMULATION OF A CONCEPTUAL MODEL OF PURCHASE BEHAVIOUR IN VIRTUAL WORLDS

Based on a thorough review of previous studies on virtual worlds and a focus on players’ general purchase intention, a conceptual research model was formulated (see Figure 1). Eight constructs appear to be significant determinants that influence players’ virtual item purchase behavioural intention and actual purchase behaviour in virtual worlds including: effort expectancy, performance expectancy, perceived value, advancement, customisation, perceived enjoyment, social influence and
habit. Moreover, performance expectancy, effort expectancy and perceived value are identified as extrinsic motivators, while perceived enjoyment, advancement and customisation are three intrinsic motivators based on Verhagen et al.’s (2008) typology of motivations for adopting virtual worlds. Extrinsic motives refer to the technical performance of an activity, which are relevant to the achievement of goals/benefits external to the system-user interaction while intrinsic motives refer to users undertaking an activity for creating pleasurable experiences, which are derived from the system-user interaction per se (Van der Heijden 2004; Koo 2009). The importance of effort expectancy, performance expectancy, perceived enjoyment and social influence has been broadly discussed in previous literature on behaviour-expectation in different contexts include general IS adoption, Web-based shopping and virtual world purchase behaviour (e.g. Venkatesh et al. 2003, Guo & Barnes 2007, 2009, Davis et al. 1992). For example, Guo & Barnes (2007, 2009) proposed that the two constructs, effort expectancy and performance expectancy are utilitarian factors contributing to explaining the relative advantages of using in-world trading platforms, which in turn influence players’ purchase channel choice when considering buying desired virtual item from others. Similarly, Venkatesh et al. (2003) suggested that effort and performance expectancies are two major beliefs with regard to system characteristics in the context of IS adoption (i.e. in-world trading platforms in this case). TPB, DTPB (Taylor & Todd 1995) and UTAUT proposed that social influence has an important impact on behavioural intention. In addition, some prior studies (e.g. Van der Heijden 2004, Dahlberg et al. 2003, Lee et al. 2006, Moon & Kim 2001, Guo & Barnes 2007, 2009) proposed the importance of hedonic factors (i.e. perceived enjoyment) for explaining users’ IS adoption behaviour. In the remainder of this section, we place emphasis on providing the theoretical justification for four new empirical constructs (i.e. advancement, customisation, perceived value, and habit).

First, we viewed “customisation” and “advancement” as two subcomponents of general achievement because they represent players’ two major desires to achieve in a virtual world. Lehdonvirta (2005) applied Yee’s (2005) motivations to analysing four different virtual worlds. According to their research findings, all motivations contributing to virtual item purchase behavioural intention can be covered by the two constructs. On the one hand, players can derive a great sense of achievement from becoming powerful, rapidly increasing game level through purchasing functional virtual items or accumulating rare virtual items or money. On the other hand, players may get self-esteem and a sense of satisfaction through personalising their character appearances or showing a unique style using decorative virtual items (Yee 2005). For example, in the famous socially-oriented virtual world, “Three.com”, “much attention is given to personal appearance, and one of the main activities in There is designing virtual clothes and selling them” (Brown & Bell 2004, p. 1). Therefore, we propose that “advancement and “customisation” should be two important intrinsic motivators for purchasing virtual items in virtual worlds in addition to perceived enjoyment.

Second, perceived value has been defined as “the consumer’s overall assessment of the utility of a product based on perceptions of what is received and what is given” (Zeithaml 1988, p. 14). Within this definition, Zeithaml (1988) proposed four kinds of meanings of value: (1) low price, (2) whatever one wants in a product, (3) the quality that the consumer receives for the price paid, and (4) what the consumer gets for what they give (p. 14). A product of perceived value is usually measured according to the difference between the perceived benefits that a product delivers and its perceived costs, or the ratio of the perceived benefits to the perceived costs. Here, costs include monetary costs (i.e. currency) and non-monetary costs (i.e. time and energy). Previous studies have proposed that there are different channels to get desired virtual items in virtual worlds. Thus, players first need to make a selection from these channels when pursuing desired virtual items through a comprehensive virtual item benefit and associated cost valuation (Castronova 2001, 2003, Guo & Barnes 2009). Consequently, it seems that the construct “perceived value” is appropriate to cover these dimensions.

Third, many previous studies proposed that habit and purchase behavioural intention may be used as two independent antecedents of actual behaviour (e.g. Mittal 1988, Tuorila & Pangborn 1988). Moreover, the moderating effect of habit on the relationship between purchase behavioural intention and actual purchase behaviour is added to the conceptual model. Limayem et al.’s (2007) research
findings proposed that the relationship between behavioural intention and actual IS continuance behaviour may be weakened when an individual has established a strong habit to use it. Here, we place research emphasis on one-off purchase behaviour rather than continuance behaviour because the latter is quite difficult to measure in the context of virtual worlds. Unlike other IS adoption behaviour, it is impossible that a player never uses virtual world trading platforms in virtual worlds especially so in a new virtual world. Inexperienced players are relatively unaware of outside trading channels (e.g. Web-based platforms or face-to-face). During the early stage of virtual world participation, Second Life users may purchase basic virtual items for daily life in virtual worlds. Moreover, an individual may obtain many free virtual items through Second Life trading platforms. Therefore, we believe that Second Life players may have established the habit of using virtual trading platforms to some extent even if they first spend considerable money on purchasing valuable virtual items for meeting their intrinsic motivators. These basic or free virtual items have low market values as they can only satisfy players’ basic survival requirements, which are not our research target. In this research, we only consider players’ commercial purchase behaviour for achieving their goals (i.e. buying valuable virtual items for advancement and customisation) in a virtual world. Thus, there is no conflict when adding habit into the conceptual model even if we focus on initial purchase behaviour towards valuable virtual items.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Conceptual Definition</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effort expectancy</td>
<td>The degree to which an individual believes that purchasing virtual items in virtual worlds would be free of physical and mental effort.</td>
<td>+\rightarrow Intention</td>
</tr>
<tr>
<td>Performance expectancy</td>
<td>The degree to which an individual believes that using virtual world trading platforms would help him or her to attain desired virtual items in virtual worlds.</td>
<td>+\rightarrow Intention</td>
</tr>
<tr>
<td>Perceived enjoyment</td>
<td>The degree to which fun can be derived from participating in virtual worlds.</td>
<td>+\rightarrow Intention</td>
</tr>
<tr>
<td>Perceived value</td>
<td>An individual’s perception of net benefit gained in exchange for sacrifices while purchasing virtual items in virtual worlds.</td>
<td>+\rightarrow Intention</td>
</tr>
<tr>
<td>Advancement</td>
<td>The degree to which an individual tends to derive satisfaction from becoming powerful and well-known or accumulating virtual world resources and personal fortune.</td>
<td>+\rightarrow Intention</td>
</tr>
<tr>
<td>Customisation</td>
<td>The degree to which an individual tends to derive satisfaction from customising the appearance of their characters.</td>
<td>+\rightarrow Intention</td>
</tr>
<tr>
<td>Habit</td>
<td>The degree to which an individual tends to perform behaviour (i.e. using a virtual world trading platform to purchase virtual items) automatically.</td>
<td>+\rightarrow Intention</td>
</tr>
<tr>
<td>Social influence</td>
<td>The degree to which an individual perceives that important others believe he or she should obtain desired virtual items through purchasing from other players in the virtual world.</td>
<td>+\rightarrow Intention</td>
</tr>
<tr>
<td>Purchase intention</td>
<td>An individual’s own estimated probability that he or she would like to obtain desired virtual items through buying from others in virtual worlds.</td>
<td>+\rightarrow Actual</td>
</tr>
</tbody>
</table>

Table 1. Constructs and definitions used in this research.

Note that, in this research, perceived enjoyment is defined as the extent to which fun can be derived from participating in virtual worlds instead of using virtual item trading platforms (systems). Guo & Barnes (2007, 2009) suggested that when a player perceives that participation in a virtual world is an enjoyable experience, he or she will have an incentive to obtain advanced virtual items for achieving personal goals thereby possibly purchasing virtual items in virtual worlds. Therefore, we suggest that perceived enjoyment should significantly influence a player’s behavioural intention to purchase virtual items in virtual worlds.
4 STUDY DESIGN AND METHOD

In this research, Second Life was selected as the target virtual world. Second Life is the most popular socially-oriented virtual world at present (KZero 2008). Since opening to the public in 2003, it has grown explosively and today is inhabited by nearly 16 million residents from around the globe (Second Life 2008). An online survey was carried out in SL. The URL to the survey Web site was disseminated by means of virtual world internal communication facilities, official Web sites and professional game forums. A total of 262 completed responses were received and 12 invalid responses were deleted. Therefore, a total of 250 qualified responses were obtained for quantitative analysis. Prior to data collection, the required sample size was computed based on the power analysis technique using G*Power 3.0 (Faul et al. 2007). For our conceptual model and a medium effect size (1-β=0.8, α=0.05) the sample size should therefore be 160. Thus, we can confirm that 250 responses are enough for detecting a medium effect size and using the SEM-PLS technique.

4.1 Measurement

A major issue associated with using questionnaire surveys in social science research is how to develop valid and reliable measures of the latent constructs (i.e. the unobservable constructs) (Churchill 1979). In this research, multi-item scale measurement was employed. For each construct, items were drawn from previous empirical studies related to the adoption of IS and behaviour-expectation. Theoretical guidance and judgment was used to select the items that best reflect the dimensions of constructs developed in this research. Although, where possible, we encapsulated items examined in previous studies to maintain consistency, some items were reworded to relate specifically to this research. The final version of questionnaire contained 43 questions (13 general questions and 30 scale items). Perceived enjoyment and perceived social status were measured using 7-point semantic differentials. All other items on the questionnaire were scored on a 7-point Likert scale with a score of 7 indicating ‘strongly agree’ and a score of 1 indicating ‘strongly disagree’. A 7-point scale excels at capturing smaller differences, which is an important condition leading to a higher reliability (Nunnally 1978). In addition, the research used self-reported previous purchase frequency and money spent as a proxy for actual purchase. Prior to the formal online survey, questionnaire pre-testing was carried out among 20 players who have high levels of participation experience in different virtual worlds; the aim was to further enhance content validity by assessing the differences between the understanding of participants on measurement items and the definitions of the variables that they are intended to measure.

4.2 Data Analysis

The hypothesised relationships developed in this study reflect general propositions to theoretically explain virtual item purchase behaviour. The proposed model and hypothesised relationships among constructs were evaluated using SEM-PLS modelling in SmartPLS 2.0M. Partial Least Squares (PLS) is a component-based structural modelling technique that has become popular in modern business research, particularly since it has specific advantages including minimal demands on measurement scales and sample size (Chin 1998a). What is more important is that the PLS technique excels at causal-predictive analysis in which hypothesised relationships are complex and few theoretical bases have been established (Gefen et al. 2000). This research is an initial attempt to empirically test a conceptual model of virtual item purchase behaviour in virtual worlds grounded in few theoretical foundations on behaviour-expectation in the context of virtual worlds.

4.2.1 Convergent validity and reliability

To test the convergent validity and reliability, two metrics were used: average variance extracted (AVE) and composite reliability (CR). AVE measures the amount of variance that a latent construct
captures from its indicators relative to the amount of variance from measurement error (Chin 1998a, p. 321); an AVE of 0.5 or higher is usually acceptable, meaning that 50 percent or more variance of the indicators is accounted for (Chin 1998a, Fornell & Larcker 1981). The second criterion used was the composite reliability, which indicates the internal consistency of measurement items of one latent construct; typically CR should be greater than the benchmark of 0.7 to be considered adequate (Fornell & Larcker 1981). As illustrated in Table 2, all the values of AVE and composite reliability for all constructs are considered satisfactory, with composite reliability at 0.746 or above and average variance extracted at 0.584 or above. It thus can be deduced that the measurement items used in this research are converging on the same latent construct.

<table>
<thead>
<tr>
<th>Construct</th>
<th>AVE</th>
<th>Composite Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Purchase</td>
<td>0.725</td>
<td>0.839</td>
</tr>
<tr>
<td>Purchase Intention</td>
<td>0.884</td>
<td>0.958</td>
</tr>
<tr>
<td>Social Influence</td>
<td>0.682</td>
<td>0.810</td>
</tr>
<tr>
<td>Advancement</td>
<td>0.696</td>
<td>0.902</td>
</tr>
<tr>
<td>Customisation</td>
<td>0.584</td>
<td>0.847</td>
</tr>
<tr>
<td>Performance Expectancy</td>
<td>0.638</td>
<td>0.839</td>
</tr>
<tr>
<td>Effort Expectancy</td>
<td>0.662</td>
<td>0.886</td>
</tr>
<tr>
<td>Perceived Enjoyment</td>
<td>0.755</td>
<td>0.925</td>
</tr>
<tr>
<td>Perceived Value</td>
<td>0.613</td>
<td>0.746</td>
</tr>
<tr>
<td>Habit</td>
<td>0.709</td>
<td>0.878</td>
</tr>
</tbody>
</table>

Table 2. Item convergent validity measurement.

### 4.2.2 Discriminant Validity

One way to test discriminate validity is to examine whether the square root of the AVE for each factor is much larger than the correlations between it and all other constructs (Fornell & Larcker 1981, Chin 1998b). To assess discriminant validity, we developed a matrix of correlations between constructs with reflective measures in which we replaced the diagonal with the square root of the AVE (see Table 3). We found that the squared root of AVE for each construct is higher than the elements off the diagonal (i.e. the correlations between it and all other constructs). According to Churchill (1979), we can further assess discriminant validity of constructs through making a comparison between the loadings of an item with its associated construct and its cross-loading on other constructs. For our model, we found that all items loaded on their corresponding constructs much more strongly than their cross-loadings on other constructs (details not shown here due to space constraints). Therefore, the above evidence provides strong empirical support for the reliability and validity of the constructs used in the research model.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>AP</th>
<th>PI</th>
<th>SI</th>
<th>Ad</th>
<th>Cu</th>
<th>EE</th>
<th>PE</th>
<th>Ha</th>
<th>PEx</th>
<th>PV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Purchase (AP)</td>
<td>0.852</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purchase Intention (PI)</td>
<td>0.446</td>
<td>0.940</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Influence (SI)</td>
<td>0.189</td>
<td>0.187</td>
<td>0.826</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advancement (Ad)</td>
<td>0.250</td>
<td>0.144</td>
<td>0.248</td>
<td>0.835</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customisation (Cu)</td>
<td>0.412</td>
<td>0.420</td>
<td>0.197</td>
<td>0.494</td>
<td>0.764</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effort Expectancy (EE)</td>
<td>0.248</td>
<td>0.528</td>
<td>0.280</td>
<td>0.117</td>
<td>0.256</td>
<td>0.814</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Enjoyment (PE)</td>
<td>0.150</td>
<td>0.307</td>
<td>0.103</td>
<td>0.020</td>
<td>0.171</td>
<td>0.278</td>
<td>0.869</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habit (Ha)</td>
<td>0.484</td>
<td>0.690</td>
<td>0.338</td>
<td>0.259</td>
<td>0.367</td>
<td>0.476</td>
<td>0.476</td>
<td>0.199</td>
<td>0.842</td>
<td></td>
</tr>
<tr>
<td>Performance Expectancy (PEx)</td>
<td>0.192</td>
<td>0.500</td>
<td>0.185</td>
<td>0.090</td>
<td>0.228</td>
<td>0.600</td>
<td>0.228</td>
<td>0.440</td>
<td>0.799</td>
<td></td>
</tr>
<tr>
<td>Perceived Value (PV)</td>
<td>0.256</td>
<td>0.334</td>
<td>0.249</td>
<td>0.266</td>
<td>0.244</td>
<td>0.226</td>
<td>0.129</td>
<td>0.355</td>
<td>0.143</td>
<td>0.783</td>
</tr>
</tbody>
</table>

Table 3. Correlations between constructs with reflective measures.
5 RESULTS

We calculated t-statistics and path significance levels for each of the hypothesised relationships through the bootstrapping method (with 1000 resamples). Path coefficients and $R^2$ values were obtained by running the PLS algorithm to assess the predictive performance of the structural model. The construct for purchase intention had an $R^2$ value of 0.453 indicating that the research model accounts for 45.3% of the variance in SL players’ general purchase intention. Moreover, the research model explained some variance in SL players’ actual purchase behaviour ($R^2=0.258$). Overall, the empirical results strongly confirm the explanatory power of the developed research model over players’ virtual item purchasing behaviour in virtual worlds. The important research findings of the empirical test are summarised below.

Figure 1. The research model with empirical results.

Overall, as we can see in Figure 1, there was strong support for most of the relationships in the research model. The extrinsic motivators in the model were all strong determinants of purchase behavioural intention at the 0.1 percent level of significance. The intrinsic motivators also showed significant relationships. The influence of customisation on purchase behavioural intention was particularly strong – the path coefficient being the largest of all direct relationships on purchase behavioural intention in our research model ($p<0.001$). Perceived enjoyment was also significantly related to purchase intention, albeit at a lower level of significance ($p<0.05$). However, our research findings showed that the path between advancement and purchase behavioural intention in SL was negative and significant at the 5 percent level. Compared to game-oriented players, SL players may freely design and create any virtual items they imagine – which is a major attraction of socially-
oriented virtual worlds. It seems that SL players intend to design rather than purchase desired virtual items for achieving advancement, which is contrary our expectation. The path between social influence and purchase intention was not supported and the path coefficient was close to zero. As mentioned above, this research focuses on explaining players’ general intention to purchase virtual items rather than a specific purchase decision. Thus, our research findings suggested that social influence had little effect on forming players’ general purchase intention, which does not demonstrate that social influence has no impact on making a specific decision relating to virtual item purchase (i.e. special behavioural intention), such as identifying a vendor.

It is worth noting that the research findings showed that although habit exerted a strong direct and positive influence on actual purchase behaviour (p<0.001), habit did not act as a moderating variable of the relationship between behavioural intention and purchase behaviour in the context of SL, which was inconsistent with previous published research (Limayem et al. 2007). Unfortunately, existing studies on exploring key factors that determine the relationships among habit, purchase behavioural intention and actual behaviour are quite limited and we believe that further research on this area is needed to solve the issue.

6 DISCUSSION AND CONCLUSIONS

This research has attempted to gain an understanding of the factors influencing virtual item purchase behaviour in virtual worlds. From a theoretical perspective, our findings indicate that the extant literature on online shopping behaviour does not adequately explore factors that either discourage or encourage players to purchase desired virtual items when participating in virtual worlds. To our knowledge, this research is the first explanatory study of virtual item purchase behaviour, which is the heart of virtual world transactions. A major gap in the existing body of knowledge regarding virtual item transactions is a lack of explanatory model with practical measures. The research has introduced a model for the measurement of virtual item purchase behaviour in virtual worlds by using previously developed constructs as well as newly developed constructs. The conceptual model was empirically tested based on a large-scale survey in SL. The statistically tests provide good support for the proposed conceptual model and a significant contribution to the body of knowledge regarding behaviour prediction in the context of virtual worlds. The rigorous testing of the conceptual model will not only extend the application of established theory models (e.g. TAM) to a new environment, but also offer a theoretical foundation for further studies on virtual item purchase behaviour.

If the research findings were to be successfully replicated in other virtual world settings, they could be of value in assisting virtual world developers in the development and design of virtual worlds. The knowledge of factors influencing players’ virtual item purchase behaviour, coupled with effective items of measurement are useful for virtual world developers to prioritise their resources in terms of manpower, investment, time and allocation, in the most effective and efficient way. Effort expectancy is an important factor that influences intention to purchase virtual items in virtual worlds. Hence, virtual world developers must ensure that they provide detailed guidance on virtual world trading platforms, design effective information search functions and simplify purchasing procedures in virtual worlds. According to our results, transaction platforms should be integrated into gameplay in an intuitive and enjoyable manner and provide clear benefit and value to the players. Players should also be afforded the freedom to easily customise their own avatars, which provides an important driver for purchase behaviour – as individuals will seek to purchase items to fulfil this desire. Developers should use interface and game design that is easy to use that enforces habitual transaction behaviour. Although the freeform nature of SL appears to suggest more individualistic behaviour it is possible that other game-oriented virtual worlds encourage more collectivistic and socially-oriented purchases.

It is necessary to recognise the limitations and possible future directions of the research. First, we cannot claim that the results obtained here hold equally well in other virtual worlds especially in
game-oriented virtual worlds because characteristics unique to specific virtual worlds can potentially alter the relative importance of factors influencing players’ purchase intention and actual purchase behaviour. The conceptual model and constructs should be further examined and revised in game-oriented virtual worlds, such as World of Warcraft (WoW), RuneScape or EverQuest. Second, another potential limitation of this research is that it adopts self-reported measures of previous purchase frequency and money spent as a proxy for actual purchase. To date, researchers still cannot find valid and reliable measures of actual purchase behaviour in the context of virtual worlds. To tackle this deficiency, we hope that, in the future, a longitudinal study could be designed to measure players’ purchasing behaviour over time. As shown in Figure 1, in this research, all seven independent variables have been modelled as a direct determinant of behavioural intention expect for habit. Thus, it do not minimise the explanatory ability of the conceptual model even if we lack appropriate items for measuring actual purchase behaviour in view of placing emphasis on general purchase intention. However, care should be given in interpreting or generalising from these findings when applying them to the overall population (i.e. all virtual worlds). Third, although the research developed most measurement items based on previous studies, this is the first time that these items have been examined in the context of virtual worlds. Moreover, most research on virtual world transactions were carried out in the context of game-oriented virtual worlds. The ability of these items to reflect the complexities of players’ perceptions and intentions in virtual worlds has not been fully explored. Therefore, the measures for the conceptual model should be viewed as preliminary and future research should be targeted at more fully developing and validating appropriate measurement items with an emphasis on new constructs lacking previous empirical testing, such as advancement, customisation and perceived value.

References


Koo, D.M. (2009). The moderating role of locus of control on the links between experiential motives and intention to play online games, Computers in Human Behavior, 25(2), 466-474


ADOPTION OF CROSS-COMPANY RFID: AN EMPIRICAL ANALYSIS OF PERCEIVED INFLUENCE FACTORS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0482.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>RFID, interorganizational systems, Adoption, Empirical study</td>
</tr>
</tbody>
</table>
ADOPTION OF CROSS-COMPANY RFID: AN EMPIRICAL ANALYSIS OF PERCEIVED INFLUENCE FACTORS

Abstract

Although the use of RFID in the supply chain still lags behind expectations, its appeal to practitioners and researchers remains unabated. Apart from technical challenges, the profitability of an RFID deployment is a major concern for potential customers. A promising way to increase an RFID solution’s profitability is to use RFID transponders in several companies along the supply chain and share the cost and implementation effort. This paper represents a first attempt to identify the factors affecting the perceived likelihood that cross-company RFID is adopted. Our empirical results indicate that profitability is a key influence factor in this context. Related important factors are the uncertainty of costs and returns and the possible imbalance of costs and returns among the supply chain participants. The influence of organizational factors, such as power, leadership and experience, is negligible in comparison.

Keywords: RFID, Inter-organizational Information Systems, Empirical Study.
1 INTRODUCTION

The potential of Radio Frequency Identification (RFID) for boosting supply chain efficiency has been stressed repeatedly by practitioners and researchers alike (Niederman et al. 2007). However, the actual adoption of the technology in supply chain monitoring applications has been unexpectedly low to date. In 2006 the biggest share of RFID transponders produced worldwide (556 million) went to applications such as ‘smart’ cards, keys, passports and tickets (IDTechEx 2007). Only 388 million transponders were sold for the purpose of identifying goods including drugs, tools, books, apparel and other consumer products (153 million) and logistical units like packages, cases and pallets (235 million). Industry experts expect the RFID market to take off as soon as item level tagging in logistics applications becomes economically feasible for selected industries (such as the apparel industry). The current market price of passive RFID transponders (about 7 Eurocents) is the pivot parameter of most profitability calculations. The efficiency gains achievable in the logistics operations of organizations such as labour cost savings and prevention of process errors has to outstrip RFID transponder and infrastructure costs. Otherwise, the vision of pervasive RFID tagging is unlikely to become a reality.

Given today’s increasing division of labour, physical items usually enter and leave the control spheres of several organizations during their lifetimes. Enabling as many supply chain participants as possible to use the same RFID transponders attached to these physical items is a straightforward measure to increase the overall financial return on tag investment. The backend-infrastructure required for the processing of RFID data can also be shared by several organizations. The final vision is the Internet of Things: the worldwide interconnection of all databases containing object-related data. Industry consortia like the Auto-ID labs and EPCglobal continue to work on the technical standards that are supposed to form the basis of the Internet of Things (Brock 2001).

Unfortunately, the introduction of cross-company IT systems has always been a difficult and time-consuming task. A well-known example of such systems from the supply chain domain is Electronic Data Interchange (EDI) which took decades to be successfully introduced. Among other things, asymmetric costs and benefits, different risk attitudes and capabilities across the supply chain participants can complicate the adoption and efficient usage of inter-organizational information systems (Scala and McGrath Jr. 1993).

In this paper we focus on the factors related to the adoption of inter-organizational RFID systems. We identify a number of candidates and empirically test the corresponding hypotheses using data from a recent survey. To the best of our knowledge, this is the first empirical work conducted on non-technical influence factors of cross-company RFID systems.

In Section 2 we provide an overview of related literature. Our hypotheses and a conceptual model are presented in Section 3. In Section 4 we describe the employed methodology and the results of an empirical study. Section 5 discusses managerial implications of our work. Section 6 concludes.

2 RELATED WORK

The academic literature on RFID and its use in supply chain management is already substantial. Recent literature reviews include Ngai et al. (2008). Most of the publications directly related to our research topic can be allocated to one of two groups: (i) conceptual and empirical research on RFID adoption on the company level, and (ii) analytical research on the distribution of RFID benefits between a prototypical manufacturer and retailer.

A number of papers belonging to the first group concentrate on collecting the views of practitioners concerning the benefits of RFID within their organizations with different emphases. Leimeister et al. (2007) investigated the perceived strategic importance of RFID among IT decision makers. They found that the perceived strategic importance is correlated with industry affiliation and company size. Seymour et al. (2007) have developed a framework of possible factors of RFID adoption based on
several accepted theories on technology adoption and diffusion, e.g. Bakry’s (2003) e-readiness model and Rogers’ (2003) Innovation Diffusion Process. Sharma et al. (2007) propose a model for RFID adoption on the company level that is among other things grounded in the literature on inter-organizational systems; they adopt a number of factors form research on the adoption of Electronic Data Interchange (EDI), in particular Chwelos et al. (2001) and Teo et al. (2003). Madlberger (2008) investigated the influencing factors on the introduction of RFID in supply chain management applications and found that internal process improvements, inter-organizational benefits, technical advantages and the costs of RFID, but not the company size, have an influence on the introduction of RFID. Using data from a survey among 146 German companies Gille and Strüker (2008) measured how the type and sophistication of benefit and performance analyses conducted before and after RFID introduction impact the productivity gains achieved by RFID. They found that the frequent use of particular measurement methods is strongly correlated with to the improvement of target variables such as lead time and labour cost. In another paper, the same authors address specific aspects of small and medium sized companies (Strüker and Gille 2008). They found that RFID adoption is easier in smaller enterprises.

Some researchers have begun to investigate the distribution of benefits across prototypical supply chains; this automatically leads to the question how RFID transponder costs should be shared optimally in case benefits are distributed unequally. Gaukler et al. (2007) investigated this research question using an analytical model. They showed that sharing the tag costs results in overall profit maximisation if the manufacturer is the more powerful than the retailer. However, if the retailer is more powerful, there is a need for sharing tag costs in order to realize the maximum profit. Unfortunately, the analytical model is based on a highly stylized supply chain model and only captures information benefit and tag cost while ignoring the benefits resulting from labour and error cost savings.

This paper empirically investigates the factors that determine the adoption of cross-company RFID. The focus on the entire supply chain instead of single companies sets it apart from previous empirical business research on RFID. While using an empirical approach makes it easier to grasp real world conditions, our findings leave more room for interpretation than results obtained from analytical models.

3 HYPOTHESES AND CONCEPTUAL MODEL

Our conceptual model consists of one major dependent variable, namely the perceived likelihood of cross-company RFID adoption. The model is designed to reveal the influence of several factors on this dependent variable. These independent variables include the expected degree of RFID profitability across the supply chain, the uncertainty of RFID benefits, the uncertainty of RFID costs, the asymmetry of RFID profitability across the participants, the existence of a driving organization that takes the initiative in planning cross-company RFID deployment, the existence of a dominant supply chain participant who can force the introduction of cross-company RFID, and the existing RFID experience in the supply chain. In addition to this greater model we also investigate how expected degree of RFID profitability is impacted by two more independent variables, namely the depth and the breadth of the inter-organizational RFID implementation. We propose the following related hypotheses:

H1: The expected profitability of cross-company RFID positively influences the perceived likelihood of adoption.

By this hypothesis we imply that higher stakes provide an incentive for better coordination. In other words we hypothesize that if the supply chain participants expect a higher profit for the supply chain as a whole they will be more motivated to collaborate in order to realize (and possibly redistribute) this return. This factor profitability appears in many studies on technology adoption; however it is usually decomposed into benefits and costs (cf. Sharma et al. 2007). We focus on profitability since we are interested in the ‘size of the pie’ to be distributed among all the supply chain participants.
The current uncertainty involved in estimating the costs and benefits of RFID may cause risk-averse decision makers to not participate in or bail out of cross-company RFID projects. Transaction cost theory suggests that in situations where the outcome of a joint investment is highly uncertain and the assets are highly specific, the emerging negotiation, monitoring and legal costs can be significant (cf. e.g. Williamson 1979). Furthermore, the fear of opportunistic behaviour can result in a complete failure to coordinate on technology adoption. We therefore hypothesize that the uncertainty of RFID benefits as well as the uncertainty of the eventual cost of the RFID implementation have a negative effect on the perceived likelihood of adoption.

H2: The uncertainty of the benefits provided by cross-company RFID negatively affects the perceived likelihood of adoption.

H3: The uncertainty of the costs of cross-company RFID negatively affects the perceived likelihood of adoption.

With regard to the consumer products industry it has often been argued that RFID will provide higher benefits for retailers than for manufacturers (Byrnes 2003). Whereas the former can use it for various purposes on the shop floor, the latter may not be able to reap substantial benefits (Weber and Jensen 2007, p. 34). Without efficient and incentive-compatible methods to redistribute RFID costs, a concerted deployment of RFID along the supply chain will be hard to achieve. Although its importance has been repeatedly stressed, RFID cost redistribution remains an open issue (Bensel et al. 2008). The more asymmetrically profitability is distributed among the supply chain participants, the more incentives in different forms have to be provided by those participants who gain more. Due to the company-centred vantage point of the existing literature, this factor has no been considered in previous work. Against this background we formulate the following hypothesis:

H4: The asymmetry of RFID profitability in the supply chain has a negative effect on the perceived likelihood of cross-company RFID adoption.

One or several supply chain participants can play a crucial role in initiating and supporting the RFID introduction process. Regarding cross-company RFID there are two prominent examples for such leading organizations: Wal-Mart and Metro. Although their methods of fostering RFID introduction differ, they both stand out as main supporters of the technology in the respective supply chains. Wal-Mart took unilateral action in planning the deployment of RFID and issued mandates to their suppliers (Romanow and Lundstrom 2003). Metro actively involved suppliers and other companies by starting the future store initiative (Loebbecke 2005). In particular, they offered non-monetary compensation to their suppliers, including the timely communication of relevant sales data. If they exist, such companies usually take the lead in coordination and standardization activities. We hypothesize that the existence of such RFID ‘leaders’ has a positive impact on the perceived likelihood of adoption, irrespective of the means that they apply to foster the adoption process.

H5: The existence of a RFID leader in the supply chain has a positive influence on the likelihood of cross-company RFID adoption.

We would like to stress that a RFID leader does not necessarily have to be powerful in the sense that it can ‘mandate’ the supply chain wide adoption of RFID (irrespective of the mentioned examples). However, the existence of a powerful player in the supply chain can have an influence just as crucial as the existence of a RFID leader (cf. Sharma et al. 2007). From an economics point of view, power asymmetries can intensify incentive problems regarding the adoption of shared information technology. If a supply chain participant is economically dependent on another one, it will fear that this other organization will impose its will when the parties disagree on some issue during the implementation process. In joint projects, such as cross-company RFID introduction, this dependency can be reinforced by large upfront investments. In anticipation of such opportunistic behaviour, the weaker party may refuse to cooperate right from the start. Against this background one would expect adoption failure if there are power imbalances in the supply chain. In the empirical IS literature, however, the contrary hypothesis seems more common, namely that the (potential) influence of a
The existence of a powerful player among the supply chain participants has a negative impact on the adoption of cross-company RFID.

If the stakeholders involved in a cross-company RFID project have gained experience with the technology upfront, they should also have a more realistic view of the cost-benefit tradeoffs and technical challenges of its inter-organizational use. This in turn should make them more confident to avoid pitfalls in the planning and implementation phase. We therefore expect that decision makers estimate the probability of cross-company RFID adoption to be higher if there is more extant knowledge about the technology in the supply chain.

A higher degree of RFID experience in the supply chain positively affects the perceived likelihood of cross-company RFID adoption.

The more details about the movement of goods through the supply chain can be obtained from RFID-enabled information systems, the higher the potential for supply chain process automation. Experts in the field have long argued that benefits are likely to increase when moving from pallet to case and case to item-level tagging (Michael and McCathie 2005). At the same time, the more processes are restructured and adjusted to each other in order to effectively use the additional data capturing capability provided by RFID, the higher the ROI of the transponders becomes. However, increasing the depth of an RFID implementation does not necessarily improve its profitability since it comes with higher implementation and integration costs. Just think of the additional transponders required to tag single products instead of cases or the effort involved in redesigning all supply chain processes instead of just one or two. Although this cost-benefit trade-off is non-trivial, we hypothesize that the depth of the inter-organizational RFID implementation has a positive impact on the expected profitability.

The depth of the inter-organizational RFID implementation has a positive effect on its expected profitability.

Similar to the depth of a cross-company RFID implementation, its breadth should have a positive effect on the expected profitability. Breadth denotes the number of different organizations participating in the RFID application. If each additional supply chain organization that participates in the inter-organizational RFID application can benefit from the RFID transponders moving through the supply chain, the overall ROI of the transponders should increase with the number of participants. The following hypothesis reflects this reasoning.

The breadth of the inter-organizational RFID implementation has a positive effect on its expected profitability.

Figure 1 shows the conceptual model and the corresponding hypotheses graphically.
4 EMPIRICAL STUDY

4.1 Survey Design and Sampling

In order to test the hypotheses outlined in the previous section we developed scales that measure the different variables. In the forefront of the survey several industry experts and representatives were interviewed in a small workshop. The wording of the questionnaire was discussed in order to make sure that the all terms and formulations were clear and interpreted correctly and equally by the practitioners.

The developed questionnaire was appended to a more general survey on RFID in logistical applications. Its online completion was possible in German or English language. Sample collection took place from April 1st to September 2nd 2008. 947 personal invitations were emailed. Additionally, the survey was announced in a number of popular logistics and RFID publications and forums. After several weeks the invited persons were reminded of the survey by phone.

The data collection effort resulted in 153 answered questionnaires. 107 out of the 153 were responses to the personal invitations, the rest (46 responses) were filled out anonymously in response to the public announcements. Not considering the anonymous responses, the response rate was 11.3%.

<table>
<thead>
<tr>
<th>Revenues in mln. €</th>
<th>% of companies</th>
<th>Employees</th>
<th>% of companies</th>
<th>Industry</th>
<th>% of companies</th>
<th>Supply chain role</th>
<th>% of companies</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 100</td>
<td>14.7%</td>
<td>&lt; 500</td>
<td>29.4%</td>
<td>Electrical</td>
<td>16.2%</td>
<td>Supplier</td>
<td>10.3%</td>
</tr>
<tr>
<td>100-1,000</td>
<td>13.2%</td>
<td>500-5,000</td>
<td>22.1%</td>
<td>IT</td>
<td>14.7%</td>
<td>Manufacturer</td>
<td>41.2%</td>
</tr>
<tr>
<td>&gt; 1,000</td>
<td>27.9%</td>
<td>5,000-50,000</td>
<td>27.9%</td>
<td>Retail</td>
<td>11.8%</td>
<td>Retailer</td>
<td>17.6%</td>
</tr>
<tr>
<td>n. a.</td>
<td>44.1%</td>
<td>&gt; 50,000</td>
<td>20.6%</td>
<td>Logistics</td>
<td>8.8%</td>
<td>Logistics SP</td>
<td>8.8%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Automotive</td>
<td>7.4%</td>
<td>IT SP</td>
<td>22.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Consumer goods</td>
<td>5.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Engineering</td>
<td>5.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Others</td>
<td>29.4%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. General sample characteristics.
In order to evaluate the representativeness of the sample, we analyzed its distribution with respect to four general company profile indicators: size of revenue, number of employees, industry affiliation, and the role that the respondent’s company plays in the supply chain. Table 1 summarizes the general sample description. None of the indicators exhibits any unexpected concentration. The companies represented in the sample take on different roles in the supply chain: as the corresponding sample description data indicates, our data reflects the opinion of managers who represent suppliers, manufacturers, retailers, logistics service providers and IT service providers. This ensures that questions referring to the supply chain as a whole are answered from different vantage points and therefore increases the validity of our results.

In addition to the basic company profile, we provide descriptive statistics on RFID implementation progress indicators. These indicators include the number of respondent organizations that have implemented RFID pilots and running applications, the duration of RFID usage and the budget of a typical RFID project within the respective company. Table 2 summarizes the RFID-related statistics. They show that the sample is dominated by respondents who have gained substantial experience with the RFID technology.

<table>
<thead>
<tr>
<th>Type of RFID use</th>
<th>% of companies</th>
<th>Duration of RFID use</th>
<th>% of companies</th>
<th>Typical RFID project budget in €</th>
<th>% of companies</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>23.5%</td>
<td>0</td>
<td>23.5%</td>
<td>0</td>
<td>23.5%</td>
</tr>
<tr>
<td>Pilot</td>
<td>26.5%</td>
<td>&lt; 1 months</td>
<td>11.8%</td>
<td>&lt; 100,000</td>
<td>25.0%</td>
</tr>
<tr>
<td>Running system</td>
<td>50.0%</td>
<td>1 - 6 months</td>
<td>14.7%</td>
<td>100,000 - 500,000</td>
<td>17.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7 - 12 months</td>
<td>22.1%</td>
<td>500,000 - 1 mln.</td>
<td>20.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 - 2 years</td>
<td>11.8%</td>
<td>&gt; 1 mln.</td>
<td>13.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt; 2 years</td>
<td>16.2%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. RFID-related sample characteristics.

### 4.2 Descriptive Results

Table 3 provides descriptive statistics of the obtained survey data. A darker cell background indicates a higher concentration of the respective answer. As the results show, the majority of survey participants judge the profitability of cross-company RFID (EP) positively. Slightly more participants believe that cross-company RFID will be adopted (LA) in their supply chain; almost half of the respondents are undecided. Surprisingly, more participants indicated that they find it rather easy to estimate costs and benefits of cross-company RFID ex-ante.

<table>
<thead>
<tr>
<th>Measured Variable</th>
<th>5 Point Scale (1 = very low, 5 = very high)</th>
<th>Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation Depth (ID)</td>
<td>9%</td>
<td>31%</td>
<td>35%</td>
</tr>
<tr>
<td>Implementation Breadth (IB)</td>
<td>19%</td>
<td>31%</td>
<td>1%</td>
</tr>
<tr>
<td>Expected Profitability (EP)</td>
<td>3%</td>
<td>9%</td>
<td>22%</td>
</tr>
<tr>
<td>Benefit Uncertainty (BU)</td>
<td>6%</td>
<td>34%</td>
<td>35%</td>
</tr>
<tr>
<td>Cost Uncertainty (CU)</td>
<td>4%</td>
<td>29%</td>
<td>38%</td>
</tr>
<tr>
<td>Asymmetry of Profitability (AP)</td>
<td>16%</td>
<td>28%</td>
<td>28%</td>
</tr>
<tr>
<td>RFID Leader (RL)</td>
<td>3%</td>
<td>13%</td>
<td>26%</td>
</tr>
<tr>
<td>Powerful Player (PP)</td>
<td>9%</td>
<td>16%</td>
<td>22%</td>
</tr>
<tr>
<td>RFID Experience (RE)</td>
<td>22%</td>
<td>16%</td>
<td>38%</td>
</tr>
<tr>
<td>Likelihood of Adoption (LA)</td>
<td>4%</td>
<td>13%</td>
<td>46%</td>
</tr>
</tbody>
</table>

Table 3. Descriptive statistics (cell shading indicates concentrations).
4.3 Statistical Methodology and Results

The purpose of our statistical analysis is to test the hypotheses presented in section 3 using our data sample. In this section we outline and justify the applied statistical methodology.

The subset of the collected data used to test the hypotheses of our conceptual model consists of 10 times 68 values encoded on a range from 1 to 5. These values represent ordinal measurement points of the defined indicators. Our aim is to determine whether the independent variables in the model have a significant effect on the dependent variables and if so whether the effect is positive or negative. A suitable statistical methodology to estimate these effects is ordinal logistic regression. We apply this method to the general and the subordinate model with the dependent variables LA and EP respectively. In a first step, we conducted regressions of all independent variables onto the corresponding dependent variables. All statistical computations related to this paper were done using the Zelig library for the statistical software R (Imai et al. 2008). Table 4 summarizes the results of the monovariate ordinal logistic regressions.

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>Indep. Var.</th>
<th>Est. coef.</th>
<th>Std. error</th>
<th>t-value</th>
<th>Sig. level</th>
<th>Residual dev.</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP</td>
<td>ID</td>
<td>0.982</td>
<td>0.233</td>
<td>4.210</td>
<td>99.9%</td>
<td>166.91</td>
<td>176.91</td>
</tr>
<tr>
<td>EP</td>
<td>IB</td>
<td>0.142</td>
<td>0.154</td>
<td>0.923</td>
<td>-</td>
<td>186.17</td>
<td>196.17</td>
</tr>
<tr>
<td>LA</td>
<td>EP</td>
<td>0.679</td>
<td>0.247</td>
<td>2.748</td>
<td>99%</td>
<td>171.21</td>
<td>181.21</td>
</tr>
<tr>
<td>LA</td>
<td>BU</td>
<td>-0.410</td>
<td>0.241</td>
<td>-1.696</td>
<td>90%</td>
<td>175.98</td>
<td>185.98</td>
</tr>
<tr>
<td>LA</td>
<td>CU</td>
<td>-1.125</td>
<td>0.281</td>
<td>-4.008</td>
<td>99.9%</td>
<td>161.12</td>
<td>171.12</td>
</tr>
<tr>
<td>LA</td>
<td>RL</td>
<td>-0.090</td>
<td>0.203</td>
<td>-0.442</td>
<td>-</td>
<td>178.69</td>
<td>188.69</td>
</tr>
<tr>
<td>LA</td>
<td>AP</td>
<td>-0.670</td>
<td>0.219</td>
<td>-3.979</td>
<td>99.9%</td>
<td>161.24</td>
<td>171.24</td>
</tr>
<tr>
<td>LA</td>
<td>PP</td>
<td>-0.332</td>
<td>0.184</td>
<td>-1.809</td>
<td>90%</td>
<td>175.53</td>
<td>185.53</td>
</tr>
<tr>
<td>LA</td>
<td>RE</td>
<td>0.201</td>
<td>0.211</td>
<td>0.950</td>
<td>-</td>
<td>177.97</td>
<td>187.97</td>
</tr>
</tbody>
</table>

Table 4. Results of monovariate ordinal logistic regressions.

The statistical effect of the IB variable on the EP variable and the effect of both RL and RE on LA turned out to be insignificant at the 90% level.

Further, the results of the monovariate regressions reveal a significant positive effect of the ID on the EP variable. The variables EP, BU, CU, AP and PP all have a statistically significant effect on LA. While the impact of ID and EP is positive, the impact of BU, CU, AP and PP is negative. The effect of the variables BU and PP is only significant at the 90% level whereas the other variables’ significance reaches higher levels.

In order to compare the relative explanatory power of the independent variables, we also employed multiple regression analysis.

When conducting multiple regression analysis so-called multicollinearity can cause problems (Mason et al. 1975): a high degree of correlation between the independent variables in a multivariate regression model can cause the regression coefficient estimates to vary erratically in response to small changes in the model or the data. In particular, the regression coefficients of all independent variables may change drastically depending on which variables are included or left out of the model. Thus, using the regression output without controlling for the adverse effects of multicollinearity on model estimation can lead to wrong interpretations.

In order to identify potential sources of multicollinearity in our data we calculated the correlation of the independent variables of the two multivariate regression models using Spearman’s rank correlation coefficient. Table 5 summarizes the results of this analysis.

The correlation coefficients and the significance levels provided in Table 5 reveal a number of significant correlations between the independent variables used in the model. In particular, there are strong positive correlations between the variables DI and BI, BU and CU, AP and EP, PP and CA, as well as between PP and AP.
Table 5. Results of correlation analysis (Spearman’s rhos and p-values, cell shading indicates strong and significant correlation).

Multicollinearity can be removed in a number of ways. One is to simply remove independent variables one by one until there is no correlated predictor variables left in the model. This process has to be conducted carefully; variables should only be removed if they cause an intolerable degree of multicollinearity.

In a first step we removed the highly correlated variables BI, RL and RE. The results of the monovariate regressions indicate that the influence of these variables on the respective dependent variables is insignificant; therefore they do not represent important predictors anyway.

The remaining correlations between independent variables include the one between CU and BU, AP and EP, and between PP and AP. In order to test whether these correlations cause problems with respect to parameter estimation, models without one variable out of each correlated variable pair were estimated and the corresponding regression coefficients were compared with the estimates from the initial model. These computations revealed that whereas the correlation between the CU and BU variable causes problems when estimating their regression coefficients within the same model, the correlation between AP and EP was unproblematic with respect to parameter estimation.

The impact of the PP variable turned out to be insignificant in the multivariate model and did not contribute to the model fit measured by the Akaike Information Criterion (AIC); we therefore dropped it from the model which eliminated the correlation between PP and AP.

Table 6. Multivariate ordinal logistic regression results after controlling for multicollinearity.
In summary, only the correlation between the BU and the CU variable could not be eliminated without significantly decreasing the explanatory power of the model. In order to resolve this problem, we estimated two models each containing one of the two variables.

The regression results for the final multivariate models after controlling for multicollinearity are provided in Table 6.

As stated earlier, the results of the multivariate regressions allow for a comparison of the different independent variables regarding the strength of their impact on the dependent variables. The estimated regression coefficient (abbreviated by Est. coef. in Table 6) indicates the direction (positive or negative) and the strength of the impact (absolute value). The obtained significance levels are sufficiently high (95-99.9%) to assume the existence of the postulated relationships with some confidence.

The PP variable becomes insignificant when it is estimated in the multivariate model. This suggests that its explanatory power is dwarfed by the other independent variables contained in the model. The AP variable has a stronger effect on LA than EP both if BU or CU is included in the model. If BU is included in the model, its impact is stronger than EP’s but weaker than AP’s. If CU is included, its impact is about twice as high as EP’s and AP’s.

As the results of the regression analysis indicate, hypotheses H1, H2, H3, H4, and H8 are supported by our data. The support for hypotheses H3, H4 and H8 is particularly significant. The statistical support for hypotheses H5, H6, H7, and H9 is not significant. Possible implications of the statistical results are discussed in the following section.

5 DISCUSSION AND MANAGERIAL IMPLICATIONS

The support for H1 suggests that the respondents who expect a high financial return on the introduction of cross-company RFID in the supply chain are more confident with respect to its realization. It implies that higher stakes in the form of unrealized profit make coordination on collaborative RFID introduction more likely. This result appears intuitive. More intriguing is the strength of the statistical support for H1 compared to the effect that uncertainty (BU and CU) and asymmetry of profitability (AP) have on the perceived likelihood of adoption. Higher uncertainty of both benefits and costs reduces the likelihood of adoption more strongly than the expectation of higher overall profitability. The same applies to the impact of a more unequal distribution of profitability (AP); the negative impact of asymmetric profitability on the perceived likelihood of cross-company RFID introduction is stronger than the positive impact of expected profitability. These results suggest that the adoption of cross-company RFID can be seriously threatened by both the uncertainty of its profitability and the imbalance of the financial returns realized by the different participants – even in situations where the overall profitability of such applications is judged very positively.

Our results indicate that cost uncertainty has a much stronger effect on the likelihood of adoption than benefit uncertainty. In other words, uncertain costs make decision makers more pessimistic regarding the introduction of cross-company RFID than uncertain benefits. Given the current problems of accurately quantifying RFID benefits this result comes as a surprise. However, it could be explained by a possible bias towards cost-based assessment of RFID applications in general. An indication speaking for this theory is the finding of Gille and Strüker (2008) that the costs of RFID applications are currently quantified more frequently than their benefits because they can be quantified more easily.

The lack of statistical support for hypotheses H5, H6 and H7 indicates that compared to the examined cost-benefit factors, the considered organizational factors have a less crucial influence on the perceived likelihood of cross-company RFID adoption. Neither does the promotion of RFID by a leading organization seem to play a decisive role (H6), nor do our results indicate a substantial influence of dependencies due to the existence of powerful supply chain participants (H7). The degree of extant RFID experience in the supply chain does not seem to be relevant either (H8).
The lack of statistical significance of H7 corresponds with the results of Chwelos et al. (2001) who tested the influence of dependency on intent to adopt EDI. However, we believe that it is still interesting since it encourages further research on the usefulness of RFID mandates such as the one issued by Wal-Mart.

The strong statistical support for H8 suggests that the benefit gain achieved by increasing the depth of cross-company RFID implementations is steeper than the corresponding cost increase. H9 can neither be supported nor rejected based on our data: the breadth of an inter-organizational RFID application in terms of participating organizations does not seem to affect its expected profitability. The economic network effect implied by this hypothesis would justify more upfront funding of initiatives that develop and standardize scalable system architectures for the Internet of Things: if profitability increased with the number of participants, the emergence of large clusters of companies that are connected by a common RFID backend infrastructure would become more likely. We believe that more research on possible economic network effects related to the use of RFID is definitely warranted.

Summarizing our interpretations of the statistical results, the main adoption hurdle of cross-company RFID implementation is the unequal distribution of profitability and the difficulties involved in estimating costs and benefits on the supply chain level. The influence of the considered organizational factors was dwarfed by the considered economic factors. In the light of high expectations regarding the profitability of cross-company implementation of RFID (see Table 1), our results suggest that in order to advance here, managerial efforts should now concentrate on the development of adequate cost sharing arrangements and tools that support more accurate ex-ante cost and benefit estimation.

6 CONCLUSIONS

Our research has lead to a number of insights regarding the factors that are perceived to influence the adoption of cross-company RFID. The impact of all considered cost-benefit factors – namely the expected overall profitability of RFID across the supply chain, the uncertainty of costs and benefits, and the asymmetry of profitability – was statistically significant. The influence of the considered organizational factors – in particular the existence of a powerful player, the existence of an RFID ‘leader’ and the extant RFID experience in the supply chain – could not be proven. Our results therefore indicate that the role of organizational factors may be overstated in the given context – at least in direct comparison to cost-benefit factors. However, the collection of more empirical data and the application of more sophisticated statistical methods are warranted in order to legitimately draw conclusions along those lines.

Technical challenges related to the use of RFID in supply chain operations, in particular the often criticised lack of technical standards for RFID hard- and software, have not been considered in our model although they could be an important determinant for RFID adoption. Follow-up research should therefore explicitly address technical issues and evaluate their impact on the adoption and success of cross-company RFID.

Based on our results we recommend that future non-technical RFID research should focus on effective and more reliable ways to estimate and measure RFID costs and benefits across the supply chain and to share RFID technology costs in an incentive-compatible way.

References


Proceedings ECIS 2009


An Event-Driven Approach to Dynamic Situation Detection

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0496.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Supply networks, interorganizational business processes, Service oriented architecture (SOA), Ubiquitous systems</td>
</tr>
</tbody>
</table>
AN EVENT-DRIVEN APPROACH TO DYNAMIC SITUATION DETECTION

Jacob, Ansger, Universität Hohenheim, Information Systems 2, Schwerzstraße 35, 70599 Stuttgart, Germany, ansger.jacob@uni-hohenheim.de
Müller, Marcus, Universität Hohenheim, Information Systems 2, Schwerzstraße 35, 70599 Stuttgart, Germany, marcus.mueller@uni-hohenheim.de
Kirn, Stefan, Universität Hohenheim, Information Systems 2, Schwerzstraße 35, 70599 Stuttgart, Germany, stefan.kirn@uni-hohenheim.de

Abstract

To cope with the emerging demand for individual products and services, new approaches for handling the individualization in dynamic supply chains are needed. Our object of investigation is to establish Situation-awareness (SA) in an assistant information system for supply chain partners to cope with individualization-driven demands. Our contribution is an event-driven architecture design to enable situational planning using the Situation Calculus upon distributed contextual information. Sensors raise events among supply chain partner’s information systems, which are combined to context events in order to derive situations on partner level. Our works follows the Design Science in Information Systems Research to design and evaluate the architecture design. The preliminary descriptive evaluation uses life stock transports as an example.

Keywords: Event-driven architecture, context-awareness, Situation Calculus, product individualization, supply chain management
1 INTRODUCTION

After a period of mass production, the demand for individualized products with prices similar to mass products is growing (Piller 2003; Pine II 1993; Reichwald and Piller 2006). In order to produce individualized goods, supply chain partners with special capabilities get together in adaptive multilevel supply chains. They work together for a limited period of time or for fulfilling the demand of only few customers. Afterwards that they participate in other supply chains to bring in their special production capabilities there (Dietrich et al. 2007). New mechanisms are needed to help combining the right production capabilities quickly and to ensure a high level of product quality as well as supply chain efficiency. Individualization research is about these adaptive mechanisms for participating organisations, products and supporting information systems in dynamic multi-tier supply chains (Kirn et al. 2008). Fleisch et al. (2005) show the past and ongoing development of information technology usage in companies from using IT for single functions to individualized 1:1 coordination with customers across company borders. In logistics individualized goods lead to an atomisation of delivery with more frequent deliveries and smaller quantities (Freitag et al. 2004). To support these future needs, real-time-management based on ubiquitous computing technologies and identification of individual products will be needed. Our work focuses on an information systems concept supporting the product individualisation process while using context and situation interpretation. Stated by Mertens (2004) Situation-awareness of information systems evolves into a major field of interest for Business Information Systems (BIS) research.

Our research is motivated by a lack of information system architectures supporting multi-tier product individualization in an autonomic manner. Object of investigation is an event-driven architecture which implements situation-based planning. The situation planning is done for each partner on top of shared context information in a dynamic multi-tier value system. The perspective taken is the decoupling of context generation and its interpretation through situation logic. This allows to access varying context information based on partner specific situation logic. Partners can join the value network and collect available context information relevant for their value transformation from all other partners to adjust their planning based on their own needs. We argue that situation logic on top of distributed context information in BIS supports individualized monitoring of the production, storage and logistics in dynamic value networks. We will present an instantiation of the architecture in life stock transports to monitor the transport for different supply chain participants. Our work follows the Design Science in Information Systems Research proposed by Hevner et al. (2004).

Our work uses a livestock transport example for evaluation of the proposed framework. Long journey livestock transports need monitoring of different parameters during transport. Architectures today, as shown in the technical specification proposed by the European Union (2006), use static rules causing alarms based on sensor events recorded by the transport company. To show the individualization supporting potential of our approach, we propose a company border crossing approach enabling situation planning for all partners in the livestock supply chain. This allows the optimization of business processes for the involved partners according to the goods situation. The situation can be set to the individual customer needs of a product, for example meat from well treated animals.

The rest of the paper is organized as follows: In the second section related work in mass customization, context and situation information systems research is shown. In section three we describe our architecture and in section four present an evaluation of the theoretical approach in life cattle transports fulfilling the European decree for long journey transports (2004). Finally we give an outlook on future work in section five.
2 THEORETICAL FOUNDATION

After decades of unifying production and products in order to achieve Economies of Scale and Scope, approaches in Mass Customization aim to support individualization and to keep the cost advantages of mass production (Piller 2003; Pine II 1993). Differentiation from competitors becomes even more important in a global economy. In future there will be more options for individualizing goods which are more suitable for differing customer needs worldwide. An example in shoe industry brings up promising results for customized goods engineered in multitier supply chains (Dietrich et al. 2007; EwoMacs 2007). To achieve this, significant changes in information system design for supply chain information systems have to be applied (Dietrich et al. 2006). On top of the findings of Mass Customization, Kirn et al. (2008) use a value system centric approach to individualization of goods and services. According to Porter (1985) a value system couples single value adding activities across more than one company with certain roles like suppliers, OEM, retailer, service provider or customer, which can be described using the Porter Value Chain model. The hypothesis in individualization research is that individualization of goods and services itself is an exclusive value for the customer that he or she is willing to pay for. Context and situations obtain a major role in building BIS for supporting pervasively adaptive value systems. Goods can be produced and monitored individually to support individualization in mostly automated manufacturing processes in an economic manner. Value system partners can be involved according to the needs of a special good without high cost efforts and the costs for individualization will decrease enormously. To achieve this, adaption of the value system in time, space and economy is implied. We realize adaption of value systems using invoked actions based on detected context and situations in our architecture.

In 1999 Dey, Salber and Abowd (2001) present the context toolkit. It offers a solution for rapid prototyping of context-aware applications and shows an integrated approach to context-awareness. The six main features of the context toolkit are encapsulation of sensors, access to context data through a network API, abstraction of context data through interpreters, sharing of context data through a distributed infrastructure, storage of context data and basic access control for privacy protection. To provide these functionalities, the following objects are considered in the context toolkit. Context widgets give access to context information for applications using the context. Interpreters can raise the abstraction of a piece of context. Aggregators are used to collect multiple pieces of context information that are logically related into a common repository. Services execute actions on behalf of applications. In contrast to a context widget retrieving state information a service changes state information in the environment via an actuator. Finally the Discoverer maintains a registry of context capabilities (widgets, interpreters, aggregators and services) actually existing in the framework. The context toolkit is a framework for dealing with context elements on different levels of abstraction, but Dey and Abowd do not explicitly define situations. The term situation can be derived from their definition for context (Dey and Abowed 1999): “Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.” Following this definition a situation is seen as a combination of context elements and reflects the detection of these elements. It does not cope with the requirements of a situation which may occur in future or try to work with special requirements a situation may bring up by its dynamic state in time. The Context Toolkit integrates important functionalities for dealing with context elements, and hence we adapt some elements of their approach to handle context information in our architecture.

Some approaches aim to individualize process execution according to a given context. Schicker et al. (2007) follow the approach of Case Based Reasoning (CBR) to find similarities between processes and electronic services within the processes to enforce the reuse of already existing knowledge. Once a process similar to the actually needed is found, its structure is changing using the adaptive CBR mechanisms. The possible operations are add, delete, change order and change of structural relations to adapt the dimensions within the CBR process. The proposed architecture implies high manual effort.
in setup and maintaining. Unknown context elements are not considered in order to have a complete picture of the process. These shortcomings are addressed in our work. Instead of forming processes based on services, we change the process based on upcoming situations. A situation defines the context it needs to know.

Handling vague context information using situation mechanisms is made by using methods applied in artificial intelligence. Bachfischer et al. (2007) make use of Bayesian Networks in order to adapt user interfaces according to given context information. Their approach optimizes complex modern car interfaces for accessing entertainment, air conditioning, navigation system, telephone and more functions. The Bayesian Network is used to estimate the driver’s situation using sensor input from the car. After the interface of the car is adapted the driver can reach functions in which he or she is most probably interested in the current situation on top level. A high temperature inside the car may lead to a high ranking for the buttons to control the air-conditioning. Currently, their implementation tries to anticipate the driver’s situation and displays a set of POI1 accordingly, e.g. the display shows gas stations when the gas level is low. This solution is possible in closed systems like a car, but cannot be prewired in dynamic supply chain networks. Anagnostopoulos et al. (2007) make use of the fuzzy logic theory to cope with imperfect observations of context. They use and define situation-awareness as “the particular kind of context-awareness, where situations are viewed as logically aggregated pieces of context” (Anagnostopoulos et al. 2007). Their definition is one of only few definitions differentiating between context and situations, but finally implies the aggregation of context only. For them a situation-aware application has “to estimate the user’s current situation(s) and react appropriately” and to “autonomously adapt to the user’s current situational context” (Anagnostopoulos et al. 2007). They propose a situation estimation architecture, which applies rules for situation detection and action determination by exploiting the user’s reactions in order to become more pervasive. According to a given estimated situation of a user the system applies actions based on an ontology containing relations between situations, persons and contexts. The context is differentiated into personal, temporal, special and artefact context. Using fuzzy algorithms and learning capabilities is a considerable future enhancement for our work. But the approach Anagnostopoulos et al. take has its limits in changing sensor environments, which we will cope with.

There are also situation- and individualization-related approaches adapted from artificial intelligence methods in supply-chain research. Freitag et al. (2004) describe the autonomous logistic concepts developed within the CRC 637 – Autonomous Cooperating Logistics Processes: A Paradigm Shift and its Limitations. They establish autonomy and a new control paradigm for logistic processes. Decisions about a process can be made without human intervention using autonomous software agents, sensor networks and mobile computing in a transport entity. According to sensor data input, autonomous agents representing each involved entity make decisions within these ubiquitous surroundings to optimize the transport. The approach is one of only few approaches not being user-centered in context aware system design. In contrast to our distributed situation planning approach for each partner, they shift the control paradigm in value networks to autonomous control without human intervention. Our approach relies on similar sensor approaches to integrate the system into adaptive business processes and to rely on decisions made by a system, but does not use autonomous agent technology.

In 1969 John McCarthy and Patrick J. Hayes (1969) pick up a situation description for machines in “Some philosophical Problems from the Standpoint of Artificial Intelligence”. They propose the Situation Calculus capable of handling situations in information systems. They form situations based on fluents, causalities, actions, strategies and knowledge, and abilities. A situation is defined by McCarthy and Hayes as “... the complete state of the universe at an instant of time”. All changes to the universe are result of an action, which are applied to a fluent. Changing fluents in a current Situation leads to a new Situation. Using a model of a situation, a machine would be capable of taking

---

1 POI (Point of interest) is an overlay map containing spatial information about certain topics as the location of restaurants, gas stations, sights etc.
actions while only limited knowledge would be needed. Further work from Reiter (1991) helped solving the frame problem (sometimes) in the Situation Calculus. The extended Situation Calculus based on Reiter is a logic containing elements for describing a situation, but does not contain any reasoning or representation abilities. The programming language Golog presented by Levesque et al. (1997) bases on Reiter’s findings to model complex operations based on the Situation Calculus. Meanwhile Golog emerged in many ways. There are approaches enhancing Golog to using concurrency (Giacomo et al. 2000), incremental program execution under incomplete knowledge with interleaved action, planning, sensing and exogenous events (Sardina et al. 2004), and decision theoretic planning for robot soccer (Ferrein et al. 2005). Li and Iijima (2007) stated in their analyses of the Situation Calculus use in IS, that “despite of the existence of these problems, the Situation Calculus could be applied to business systems when some tradeoffs are made between expressiveness and reasoning”. Li and Iijima propose Web Service Composition, Context Awareness and Business Process Management as fields of application for the Situation Calculus. We use the Situation Calculus based on Reiter’s approaches to build the situation logic for planning situations according to goals and based on context information. There are different extensions considered, as for example incremental program execution under incomplete knowledge proposed by Sardina et al. (2004).

Already researchers make use of the Situation Calculus in BIS. Dong et al. (2004) use it to apply autonomous computing in ubiquitous computing services. Their approach separates available fluents into a precondition fluent base reflecting the maximum of fluents affecting a situation, decisive fluents affecting a situation and non decisive fluent not affecting the situation. The Validity Theory ensures only valid components are Compared to approaches using Finite State Machines for situation detection, the Situation Calculus allows the detection of non-predefined situations during runtime. The Validity Theory formalizes application-specific validity requirements to validity-ensured policies, on which computers can autonomously choose to serve a service request. The approach is a promising way of reasoning using the Situation Calculus and Validation Theory in ubiquitous computing, but research is still at its beginning. The authors argue that commercial development methodologies are not available for the Situation Calculus, and that it is still difficult to deal with the Situation Calculus on engineering level. The works of Dong et al. are the main starting point for our situation-related work.

Concluding the related work there is no satisfying approach to use situation logic on top of real world information represented by sensor-based context information, but there are many promising approaches in context and situation research already. We include suitable research identified from context and situation research, especially the ideas from the Context Toolkit from Dey et al. (2001) in our context layer and the Situation Calculus with its extensions in our situation layer (McCarthy and Hayes 1969).

3 AN EVENT-DRIVEN APPROACH TO DYNAMIC SITUATION DETECTION

In this section we propose an event-driven architecture which is capable of situation handling on top of distributed real world information. It enables higher-level IS to use the paradigms of context- and situation-awareness based on sensor information. The novelty of our work compared to other approaches is the combination of a Situation Calculus approach for situational reasoning with an event-driven context abstraction logic based on sensors. This enables partners in collaborative environments like dynamic supply chains to maintain their own situation logic expressing their need, whilst product-related context information can be accessed from all partners in the collaborative environment. The following is a presentation of the three-layered design of this approach.

Today context- and situation information is used mainly in optimizing the human computer interface (HCI) design. An example for this field of application is the use of information about the position of a device which minimizes the data presented to the user to the information relevant for the current
location. This context service is called location-based service as described by Wehrmann (2004) in his work about situation-dependent mobile services. Beyond location nearly any context information can be used in a way to simplify interaction with a device. By traditional means an application becomes context-aware if it causes actions according to changing context information. The definition of context-awareness presented by Dey (2001) emphasizes the need for a user: “A System is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user’s task.” In contrast to this user-centred view we follow a more system-centred one. Distributed sensors in collaborative environments offer their information in an event-driven manner. This is matched with situation logic and its reasoning capabilities. Context sources become exchangeable and situations are detected and processed upon their availability. Our architecture integrates situation reasoning based on assumptions and targets on top of dynamic context knowledge. A sketch of the architecture including two partners sharing their context information is shown in figure 1 and described in the following sections.

![Figure 1. An architecture model for dynamic situation detection](image)

At first we introduce the elements of the Sensor Abstraction Layer (SAL) operating on event level. The Sensor Services (sens) hide the specifics of a certain sensor. All communications with the sensor are leveraged to a unified XML-based format. A short history of sensor values, stored in a database of the sensor hosting node, is implemented to allow queries involving the past. Information about the sensor range and the position are stored in the database as well. It is developed similar to the concept of Widgets from the Context Toolkit and extends them in standardization using Web Service technology (Dey et al. 2001). For the description of sens we use a simplified Sensor Model Language (SensorML) as proposed by the Open Geospatial Consortium (2007). SensorML itself is a very powerful sensor description language. Beyond the sensor description itself SensorML stores metadata about the sensor, especially the sensors position, its availability or its range in space. We used SensorML to develop our corresponding database.
Besides physical sensors, already available data can also be accessed through the sensor abstraction layer. The data can be encapsulated like conventional sensors in SensorML using a standard sens services. Examples for explicit information interesting for encapsulation could be weather information, traffic information or data from a company’s database. All sensor services are part of a partners system and are accessible by other partners for the purpose of a specific supply chain need.

On the level above the SAL, the Context Layer (CL) contains Context Composition Services (ccs). The ccs contain a composition logic and combine elementary sensor information or other ccs information with higher aggregated context information. These are similar to the concept of Aggregators Dey et al. used in the Context Toolkit expect the fact a ccs can integrate other ccs (Dey et al. 2001). The ccs raise the level of context in a cascading manner. A possible combination, for example, is having one sensor detecting the current outside temperature and another sensor detecting the illumination level in the same local area. These can be combined to a ccs delivering information about day and night time more accurately. The ccs are valid depending on their relationship in space and time (for example “in the same room at the same time”) and can expire according to the availability of underlying sens or other ccs.

The Access and Registry Service (ars) gives access to one instantiated context framework on top of a cascade of ccs or even only a single sens. It contains a registry of underlying sens and ccs and is the entry point for other partners accessing one partners services. The ars can be combined with other partner’s ars. In this case all lower services become available in both ars. The combination of different ars is applied to connect different partner domains for a limited time or to reduce handling complexity in case of larger networks.

All element names from SAL, CL and SL use the concepts of a common Matching Ontology (MO). Thus we can estimate the closeness of two elements and even match similar elements of every layer.

Formally the CL is a multidimensional undirected graph, in which R are all available edges, \( \alpha : R \rightarrow (SENS,CCS,ARS) \) being the starting points and \( \omega : R \rightarrow (SENS,CCS,ARS) \) being the endpoints of an edge. Within this configuration, each \( ccs \in CCS \) can be connected to any \( sens \in SENS \) or other ccs to generate higher levels of context:

\[
CL = ((SENS,CCS,ARS),R,\alpha,\omega)
\]

Each Sensor is connected to one sens exclusively. A valid minimum CL is one sens and one ars to access the sens:

\[
CL_{min} = ((sens,ars),R,\alpha(sens),\omega(ars)) \text{ and } |R| = 1
\]

One level above the shared and common CL, the partner-specific Situation Layer (SL) is located, which is using the Situation Calculus from McCarthy and Hayes (1969) as well as further developments from Reiter (1991). The SL’s main task is to detect a situation \( s \) upon situation template axioms \( s^* \) taken from a situation template repository \( S^* \).

The Situation Calculus is a first-order language with some second-order features for representing dynamic domains. According to its definition, all changes to the world are result of an action and the situation “represents the complete state of the world at an instant of time” (McCarthy and Hayes 1969). Thus a world history is a sequence of actions and is represented by the first order term situation. The constant \( S_0 \) denotes the initial situation in which no actions have yet occurred. Objects catch everything else relevant depending on the situations domain. All \( a \in ACTION \) are specified by an action precondition axiom and a successor state. The action precondition axiom is highly important for our work, because we need to find relevant preconditions fulfilled for a certain situation according context detected:

\[
Poss(A(x_1,\ldots,x_n),s) \equiv \Pi_d (x_1,\ldots,x_n,s)
\]
Πₐ is a formula uniform in s with free variables xᵢ. To define the closeness of a situation to a success, Dong et al. (2004) use special fluents on top of Reiter’s approach to estimate the closeness to a successor state. Their concept of incremental successor states containing the distinguished binary function \(\text{do}(a, s)\) denotes the situation resulting from performing action \(a\) in situation \(s\). \(\Psi_{Fₐ}\) helps to find the way coming closer to a better situation \(s\) performing an action \(a\) instead of having a binary function like in traditional Situation Calculus:

\[
\text{Poss}(a(x), s) \supset F(\text{do}(a(x), s) = \Psi_{Fₐ}(x, F(s))
\]

Dong et al. (2004) argue that the definitions for a fluent in standard Situation Calculus would be too general, and therefore propose an extension to make it capable of comparing situations. They classify into functional fluents \((\overline{FF})\) meaning all available fluents, decisive fluents \((\overline{DF})\) meaning a minimum set of fluents for a situation, non decisive fluents \((\overline{NF})\) meaning possible fluents for a situation and a precondition fluent base \((\overline{PFB})\) containing all fluents which may affect a situation:

\[
\overline{FF} \subset \overline{PFB} = \overline{NF} \cup \overline{DF} \text{ and } \overline{DF} \cap \overline{NF} = \emptyset
\]

Decisive fluents are all atomic fluents or ones that cannot be derived by other fluents for a specific situation. Based on the differentiation from Dong et al. (2004), a comparison of situations according to fluent classification is done in our approach.

Using situation validity based on Validity Theory, Dong et al. (2004) define a minimum set of relational fluents to ensure the situation validity. For one situation a possible validity fluent is defined as:

\[
VA(\bar{y}, s): \text{Object}^* \times \text{Situation} \rightarrow \text{boolean}
\]

All \(VA\) in the vector \(y\) contain special terms in which a situation \(s\) is valid. The situation \(s\) is valid only if all \(VA(s)\) in \(y\) hold. A situation can have none, one or more validity arguments to define the state in which a situation is valid. This is used in our approach to ensure that certain situations only come true when valid configurations of events occur.

We are using the Situation Calculus in the SL to match with context information available from CL for a current situation \(s_i^*\) from all available situations \(S^* \in S^*\). The syntax of actions, objects and fluents in SL is taken from the MO. This ensures a unique syntax in CL and SL. The matching process starts within the SL matching precondition axioms to \(ccs\) or \(sens\) of the CL. The search starts at the highest level in CL, the most upper \(ccs\), going down the graph along the \(ccs\) to the \(sens\) using graph search algorithms in order to find a best match to a situation template \(S^* \in S^*\). Using the MO a possible combination of \(sens\) and \(ccs\) can lead to new \(ccs\) matching a specific precondition axiom. Detectable situations are all \(s_i^* \in S^*\) with a matching between precondition axioms and \(ccs\) or \(sens\):

\[
s_i^* = \max \{ \Piₐ(x_1, \ldots, x_n, s_i^*) \cap (CCS \cup SENS) \}
\]

Once all detectable \(s_i^* \in S^*\) have been found, the plans can be applied to a detected situation. A plan can lead to new detectable situations within the SL with new matchings to the CL. The MO is used as taxonomy in order to have a unique vocabulary to find similarities between the elements of the CL and SL. In order to prevent a time-consuming search for a best match when applied to the available \(SenS\) and \(CCS\) in CL, prebuilt graphs for all elements of the CL are generated every time there is an update in the CL. Additional functional validity fluents can be used to ensure validity of a \(ccs\) or \(sens\) in space and time in more complex environments.

In this section we presented the theoretical approach of our event-oriented architecture. It allows new sensors and services to be added and removed from the CL during runtime (Michelson 2006). If new sensors including their corresponding \(sens\) are integrated, they register with their correspondent \(ars\) which may spread the information to other related \(ars\). The \(ars\) inform current subscribers in SL about
the new service, which may book the service then. The architecture ensures the possibility to compose dynamic value systems containing many sensors from changing value system partners. Also mobile sensors which are changing their context in space and time can be handled dynamically, for example the sensors of a livestock transport can be added logically to the system of the butcher.

4 EVALUATION

The empirical evaluation of the implemented framework in livestock transports will be determined in the future. Instead we use a descriptive evaluation to detect the situation heat shock for animals in trailer for the animal transport company and heat stress for the butchery now.

At first we model the CL as a graph starting with the ars on top of the sensors and containing direct links from ars to every sens in the trailer and is aware the ars from the farm:

\[ CL = (\langle sens_1, sens_2, sens_3, ars \rangle, R, \alpha(ars), o(sens_1, sens_2, sens_3)) \]

The object to be manipulated by the situation heat shock for animals in trailer for the transport company in SL\textsubscript{transport} is the route the transporter will take. To alter an existing route (res\textsubscript{routeo}) in situation s we need the situational preconditions low water consumption before transport (con\textsubscript{low}), high temperature in trailer (temp\_hi) and no drinking water in trailer (¬wat) fulfilled:

\[ \text{Poss}(\text{res}(\text{route}_o), s) \equiv \text{con\_low}(s, \text{route}_o) \land \text{temp\_hi}(s, \text{route}_o) \land \neg \text{wat}(s, \text{route}_o) \]

Altering a route\textsubscript{o} in a situation s should lead to a new route\textsubscript{n} which either has enough water consumption before the transport detected, or has enough drinking water for the resulting driving time, or has no high temperature for the resulting route. A successor state after altering the route\textsubscript{o} in situation s to a new route\textsubscript{n} is:

\[ \text{route}_n(\text{do}(\text{res}(\text{route}_o), s) \equiv \neg \text{con\_low}(s, \text{route}_n) \lor \neg \text{temp\_hi}(s, \text{route}_n) \lor \text{wat}(s, \text{route}_n) \]

The situation validity is ensured checking against the maximum driving time \( t_{\text{max}} \) allowed for long journey livestock transports. In European law this is set to a maximum of 8 hours without a break for the animals (European Union 2004):

\[ VA(s) \equiv (t(\text{route}_n) \leq t_{\text{max}}) \]

Using an MO with values for livestock transports the elements con\_low can be matched to sens\textsubscript{3} with a minimum amount each animal has drunk before, temp\_hi is matched to sens\textsubscript{2} with a certain minimum level of temperature detected and wat is matched to a low-level-indication of sens\textsubscript{1}. The maximum journey time is set in the MO as well. Each possible route\textsubscript{n} considered does fulfil the successor state above. In case there is no route alternative, the driver is informed according to the given situation.

With the same sensor information we can conclude to a different situation for a different supply chain partner. The object to be manipulated by the situation heat stress in the butchery’s SL\textsubscript{butcher} is the heat stress for the animals transported, which can be detected using temperature and water consumption sensors. Water consumption uses the history of water levels in order to calculate the water consumption. The heat stress condition is important in order to estimate the quality of the meat. The preconditions needed only depend on a high temperature exposure high temperature (temp\_hi) and a high level of water consumption (Δwat):

\[ \text{Poss}(\text{de}(\text{heatstress}), s) \equiv \text{temp\_hi}(s, \text{heatstress}) \land \Delta\text{wat}(s, \text{heatstress}) \]

Depending on the precondition fulfilment, the rest time of the animals before being butchered can be determined by the butchery.
All fluents described in the examples are within the $\overline{DF}$, because they affect the current situation. But we could also detect the situation if we would have an additional parameter like a considered traffic jam, but no sensor to match to. Then the route would also be altered ignoring traffic jams and the traffic information is part of $\overline{NF}$. This becomes important when dealing with large amounts of fluents.

Using the proposed architecture for monitoring the transports leads to an adaption of the transport process during runtime in contrast to single alarm events like high temperature or low water in other installations of a monitoring environments for long journey livestock transports (Joint Research Centre 2008). This helps not only notifying in case something goes wrong, but preventing something might go wrong and assist finding solutions. Also other solutions don’t provide individual situations to be detected by different partners. We presented the detection of two simple situations heat shock for animals in trailer for the animal transport company and heatstress for the butcher in order to optimize the company’s business processes based on sensors of the truck and the farm. The sensors are interchangeable enabling new applications in environments with changing sensor sources.

5 CONCLUSION

We presented the concept of an architecture which uses information about the situation on top of context information for complex context-based decisions in pervasive computing. The framework is currently under development and the approach in this work is of theoretical nature. There are still open issues to solve within the design, especially developing an efficient matching of the elements in the $CL$ and the situation templates $S*$ based on the $MO$ and an efficient structure of the $MO$.

Our approach is limited to static situation execution yet. As long as parts of the situation are detected and no argument is invalid, the situation is considered true. In our future work we need to add the ability to handle uncertainty in situation detection, for example when two of three preconditions are fulfilled. Vassos and Levesque (2007) have an interesting approach using the Situation Calculus with incomplete knowledge of an agent.

A next major step is the implementation of the livestock sensor scenario using IndiGolog proposed by Sardina et al. (2004). We will use RFID for animal identification, and sensors detecting conditions in the trailer as well as external parameters. Other supply chain partners are represented by their own situation client. Within the implemented scenario we will be able to test scalability issues which will be a challenge to deal with.

In our future work a connection between the framework and a higher level transport information system will be implemented to demonstrate the advantage of situation handling in logistics. Whereas new situations based on a changing context are determined, an automated adaption of the planning process is achievable. Furthermore interesting topics cover human feedback to enhance the system making more precise situation detections or continuous sensor values substituting other values at a certain level. An integration of digital world models in our architecture could enhance the capabilities of the framework by using spatial relations. Nicklas et al. (2001) explore using spatial aware applications in digital world models. This could enhance our situation reasoning by physical distance of objects, for example causing a situation “Possible infection transfer” by two animal transports being next to each other on a rest area.

Our architecture is not limited to animal transports. It can be used whenever distributed systems share a common data in an event based manner. Thus it can be used in many setting in supply chain BIS. We are still at the beginning of discovering situation potential in supply chain BIS, and we expect an increasing level of automation and lower coordination costs.
6 ACKNOWLEDGEMENTS

This work was supported by the German Ministry for Research and Technology (BMBF) within the research project IT FoodTrace – IT supported Food Traceability (2007).

REFERENCES


Proceedings ECIS 2009


Proceedings ECIS 2009
DEDUCING DEMANDS AT BUSINESS-INTELLIGENCE-SYSTEMS BEYOND UNBUNDLING WITHIN THE EUROPEAN ENERGY MARKETS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0371.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Business Intelligence, Case Study, Data flow, Information</td>
</tr>
</tbody>
</table>
DEDUCING DEMANDS AT BUSINESS-INTELLIGENCE-SYSTEMS BEYOND UNBUNDLING WITHIN THE EUROPEAN ENERGY MARKETS

Buder, Johannes, Technische Universität Bergakademie Freiberg, Lessingstrasse 45, 09599 Freiberg, Germany, Johannes.Buder@bwl.tu-freiberg.de
Felden, Carsten, Technische Universität Bergakademie Freiberg, Lessingstrasse 45, 09599 Freiberg, Germany, Carsten.Felden@bwl.tu-freiberg.de

Abstract

The paper presents a process model to deduce utility specific requirements within the utility industry beyond unbundling. It is based on a case study. Due to the market transition, Business Intelligence (BI) and its tasks of extracting, transforming, and analyzing data at different enterprise levels becomes relevant within this industry. Whereas utilities were formerly protected by regional monopolies, they had to change to competition nowadays. Therefore, getting information in the appropriate quality, at the right time, and in every situation is important for all management levels. According to empirical studies, the BI usage within the utility industry is low. The companies invested a lot of money into hard- and software, but had no time to fulfill business requirements. One reason is that priority tasks like unbundling had to be implemented first. Such projects were time consuming so that there were no resources left for additional vertical information flow oriented activities. Moreover, current BI implementation models do not consider utility specific requirements. Due to this reason, the contribution of this paper is to present a business flow orientated process model that deduces requirements by presenting the main business processes, data volume, and complexities on a department level, especially in context of the European unbundling directive.

Keywords: Business Intelligence, Utility Sector, Unbundling
1 INTRODUCTION

The liberalization of the energy market has dramatically changed the European power market since 1998. The former monopolistic market structure has been broken and abruptly substituted with competition (EU 2007, Steger et al. 2008, p. 20). Moreover, the liberalization induces that the management have to satisfy also investors' demands. Due to this reason, supporting the management with internal and external information is of increasing importance in context of planning, forecasting, and budgeting (Appelrath & Chamoni 2007, Horvath 2006). However, traditional planning and controlling tools are neither flexible nor structured enough to support the management with information and to bridge the gap between strategic and operative business processes (Kemper et al. 2006, Oehler 2006). Therefore, Business Intelligence tools (BI) are used in order to ensure fast, reliable, and integrated planning and controlling. According to Howard Dressner, BI includes all decision support information systems,\(^1\) from collecting data to storing data and its presentation (Gartner 2008). Its core applications imply adequate tools and applications in order to improve both the decision making process and communication with the users. The goal of BI is to design and implement business applications, enabling qualitative and quantitative analyzes (Gluchowski & Dittmar 2008, pp. 89). Consequently, the contribution of BI is a flexible corporate management on all management levels. Furthermore, BI is an instrument for knowledge generation, knowledge distribution, and knowledge usage (Gluchowski & Dittmar 2008, pp. 319). However, introducing BI does not work automatically. Existing reference models for BI strategy and BI implementation, e.g. Steria Mummert (Czochter 2008) and BARC (BARC 2007) do not imply utility specific requirements. Instead, they present a multitude of criteria that are either too general or do not concern utility specific requirements. Due to this drawback it is the paper’s goal to present utility specific requirements as a basis for BI software evaluation and successful implementation. The selection and weighting of requirements serves as a guide for coming BI implementation within the utility sector.

The research objective is realized by a process model that is based on a case study with a German utility company. Additionally, the research objective is to deduce department- and business-oriented as well as unbundling-conform requirements in favor of an enterprise wide BI-system. The paper is structured as follows:

- The necessity of formulating utility specific demands is explained by describing the background of the European energy sector (section 2).
- IT strategy is explained, because the characteristics of the energy sector influence the requirements so that they are very specific and a copy&paste from other economic sectors is not adequate.
- The meaning of vertical information supply is described as well as its influence for BI in the energy sector (section 3).
- The reference framework is presented in order to substantiate the chosen research method.
- Chapter 4 deals with the explanation of the process model in order to deduce requirements. The contribution of this chapter is a business task oriented information flow from data sources to their destinations in utility companies. This is summarized in a data-flow oriented architecture (process-system-map). This enables us to determine so called gravity centers (defined by their amount of data and complexity of algorithms) to characterize the departments and to define a generalized strategy map which is done the first time within utility companies.
- Following, the BI requirements and their impact in respect of the utility industry are presented.
- Finally, the paper ends with the conclusion.

---

\(^1\) The term was mentioned by IBM in 1958, but refers to a more general sense (Luhn 1958).

\(^2\) It is not the paper’s goal to analyze the different BI definitions and describe the BI process. These definitions and explanations can be found in (Gluchowski & Dittmar 2008a; Watson & Wixom 2006; Kemper et al. 2006).
2 RESEARCH BACKGROUND

This section explains the transformation of the European energy market. It illustrates the necessity of BI within the utility sector. Moreover, the problem of formulating utility specific BI requirements is described.

2.1 The European Energy Market Transition

The majority of the European utility industry is still structured by monopolies and oligopolies (Storr 2007). In Germany and France for example, the five largest utility companies hold 95 percent of the high voltage grid and generate more than 80 percent of the total power generation (NA 2004, EDF 2007). The monopolistic structure is based on a vertical integration of the utility companies from power generation, transmission and distribution (T&D), and retail services. Nevertheless, the monopolistic structure has some advantages. Firstly, building up a second grid is expensive and wasteful, because the T&D system is a natural monopoly. Therefore, it is easy for utility companies to reach economies of scale by increasing their capacity at a progress rate to gain decreasing average cost. Secondly, a vertical integration leads to economies of scope. Because of extensive investment cost, monopolies can realize synergies by coordinating and concentrating both investment and operations (Hung-po et al. 2005). But this aspect contrasts the goal of the European Union (EU) which is characterized by economic integration in order to eliminate the economic frontiers between two or more countries (Pelkmans 2001, pp. 2). In context of the European energy sector, the EU has started a single market for electricity and gas (EU 2003) by opening the market for competition and incrementing transparency and data protection. While the progress may differ within the EU countries, the one thing they all have in common is the unbundling of the utility companies’ business activities and the highly volatile trading of commodities such as fossil fuels, electricity and even water (Felden 2002). Unbundling means that the main operation, T&D and retail in vertical integrated companies are separated from each other to enable third parties a non-discriminating access to the market. According to the directive, the member states have to implement regulatory authorities that secure competition, free market access and fair compensation fee for the unbundled grid (the grid still remains as a regulated monopoly). Moreover, the directive is distinguished into legal, operational, accounting and informational unbundling. Legal unbundling deals with the company legal entity of the Distribution Service Operator (DSO). Accounting unbundling separates the T&D from an accounting perspective, so that the market prices for electricity can be identified independently from T&D costs. Organizational unbundling ensures that the T&D unit operates independently from the utility company. Finally, the informational unbundling deals with the confidential treatment of sensitive and business oriented information. In summary, unbundling shall create equal market conditions for all grid users (EU 2003).

2.2 Problem Formulation

Integrated utilities are confronted with increasing competition whereby they had to implement the unbundling regulations first, before other demands could be fulfilled. Such an implementation leads to efforts caused by a large number of national legislations\(^3\) based on the EU directive (Nick 2006)\(^4\). Additionally, utility companies are faced with increasing market stress of competition, especially in the power generation and DSO units (Steger 2008). Hence, decision makers have a need to get the right information in context of energy trading and incentive based regulation (Nick 2007).

---

\(^3\) The German legislations are concerning grid access (both natural gas and electricity), incentive based regulation, and grid remuneration e.g. EnWG, StromNZV, StromGVV, StromNEV, GasNEV, GasGVV, EEG, Anreizregulierung.

\(^4\) E.g.: Business processes have to be checked to unbundling guidelines. Moreover, utilities are forced to prepare an unbundling report yearly.
Furthermore, the management is obligated to its investors. Both reporting oriented needs are usually not formulated, yet.

But whereas the power generation is capital-intensive, budgets are getting tighter. In this context, controlling tools like planning, forecasting and budgeting are of major importance (Horvart 2006). Surprisingly, especially these tasks are often done manually within the utility sector. It is quite evident that the considerations about unbundling determine both the business oriented horizontal and also the vertical information flow. An information flow can be described as a path from a supplier to an end user or customer independently from their locations (Felden 2008). Management decisions are based on information that is gained from the vertical information flow. With regard to unbundling and particularly competition, BI offers opportunities in order to fulfill information requirements and to handle the increasing data volume in utility companies. Empirical studies show also that the technical premises for a BI system are already established in utility companies. But BI maturity studies explain an initial status of BI usage. It is limited to distinctive departments, ad-hoc reporting is not established, and reports are created manually and redundantly (Chamoni & Gluchowski 2004). Moreover, the business requirements remain unclear, so that the actual system’s usage is low (Czotscher 2008). This defines the need for BI within the utility sector. Due to this, we set up a case study to be able to identify BI-oriented requirements of the utility sector so that a functional BI implementation is easier to handle.

3  IT IMPACT

The unbundling directive determines the IT-architecture of utility companies. Securing non discriminating access to the market and information preservation of sensitive and confidential data are the core messages of informational unbundling (§10, 3 EnWG5). In context of a horizontal information supply, business processes and its information are faced with unbundling, although they were fully integrated before6 (Abel 2005). Unbundling determines data that can be of technical, economical, or legal nature. From a technical point of view, especially ERP and metering systems are focused with unbundling. On the one hand, it is necessary to adapt the interfaces of formally integrated systems in order to comply with unbundling directives. On the other hand, the user authorization concept has to be mentioned. In this context, the Need to Know policy plays an important role. Need to Know means those users are allowed to see just the information concerning their business needs (Hartmann & Bitz 2008). This principle enables the opportunity that every DSO customer is able to access its user data that are recorded within the DSO (Nick 2007). National authorities advise the separation between data sources that are used from DSO, retailer, and a generation unit. Moreover, unbundling requires a detailed documentation of data and folder structure (§10 EnWG).

The consequences of unbundling are that the IT architecture and strategy determine the horizontal information flow. Due to the increasing competition, the management is more interested in appropriate and aggregated information at the right time in order to take decisions comprehensively (Oehler 2006). So, the unbundling within the horizontal information flow influences heavily the vertical, BI-oriented information flow. Thus, it is necessary to consider unbundling directives to identify and analyze critical business processes. The identification implicates the effect that utility companies have to think about their business processes by identifying the process operations, the responsible person, the trigger, the interfaces, and the expenses (Nick2007). The contribution of BI is the integration of measured processes performance data from operational systems like an Energy Data Management system (EDM). Consequently, unbundling and the requirement to identify business processes is a device to accelerate the implementation of internal cost allocation. In this context, BI contributes to the strengthening of competitiveness while supplying the information demand. But the collected data

---

5 EnWG is the abbreviation for Energiewirtschaftsgesetz.
6 E.g.: change of supplier, metering services, changing master data, business data query, and grid-use-billing.
at the different company levels have to comply with the defined unbundling requirements. These requirements are relevant in context of integrated planning and decision support systems. Also specifically, such requirements are not valid for other industries.

4 PROCESS MODEL FOR BI ARCHITECTURE

This section presents the research framework to explain the chosen research method. After the explanation, a process model is described in order to deduce specific requirements.

4.1 Research Framework

We are using a case study as an empirical method, in order to describe the construction of a process model for a BI-strategy to be able to deduce utility specific requirements. Yin has described the case study as an empirical inquiry, “…that investigates a contemporary phenomenon within its real-life context…” (Yin 2003). We have chosen this method, because the direct involvement into business processes was very important to us. Speaking face to face with department managers and system users in order to understand their issues and identify requirements of BI systems in the energy sector is important, because current studies (Czotscher 2008, BARC 2007) ignore usage characteristics and specific requirements. Furthermore, companies are confronted with numerous surveys so that it is disputable, if they participate with the necessary engagement, if another anonymous one is approaching (Pendse 2007).

4.2 Process Model for a BI Architecture

The current procedure is based on the project model of VDI directive 2221 (VDI 1993). This approach is appropriate, because other process models are too specific in context of a service design (e.g. Scheuing & Johnson 1989, Shostack & Kingman-Brundage 1991), are focused on software development (e.g. Royce 1970, Boehm 1988, BMI 2006), or ignore utility sector specific requirements in context of a BI implementation (Totok 2006, Knoell et al. 2006). Due to the reason that the VDI directive gives advice on an aggregated level, we are able to adapt its general rules in favor of the study. The study uses the first four steps of the VDI approach, because the focus is not the development of new software but the construction of a process model for BI architecture. These steps are:

- Defining the problem;
- Identifying the functions and structures of the determined departments;
- Searching for solution principles and its structures;
- Classifying modules;

In the first phase, the problem definition is done by interviewing managers of the strategic management and the IT department. Furthermore, the projects goal and budget are defined. In the second phase, the status quo is established via an expert survey in different departments. Its goal is to describe an actual handling of information and the department functions. The survey concentrates on controlling, IT-services, investment, customer service, DSO- and sales department managers. BI is relevant for these departments, because controlling needs standardized reporting tools whereas the other departments use complex analyzes in order to plan investments, understand customers’ requirements, and to forecast grid capacity and grid costs. The results of this phase are use cases for the business processes. In the third phase, all use cases are consolidated to a process-system-map...
that describes the main business processes within the departments. After that, a consolidated PSM of the utility company and the DSO is built. Its goal is to illustrate the main company processes by showing which business process runs, which data sources are used as input, and which data are produced as output. At the end of this phase, each department and the whole company are skipped into a maturity model. This model is based on the Mummert BI Maturity Model [biMM] (Mummert 2004). This model is chosen, because of its empirical evidence, especially in Germany (Chamoni & Gluchowski 2004). The model’s intention is to clarify the interrelation between maturity and return on investment. On the first stage, a company does not have any integrated databases. Therefore, it focuses on data integration. Within the second stage a company has reached a first level of standardization by implementing a BI tool for reporting. At the third stage, a company involves a data warehouse [DWH]. In addition, extensively BI functions are provided, e.g. OLAP. At the next stage, companies have the opportunity to use data mining algorithms in order to discover implicitly information in databases. Finally, companies which are placed on the highest maturity stage are able to use data mining results directly to improve their processes. The fourth phase uses the results from the previous phases. Its goal is to deduce department specific requirements. At first, BI gravity centers are built. The centers are symbolized as bubbles, whereas each bubble represents a department. The bubble’s surface area describes the number of users and its locality symbolizes the grade of complexity and data volume. The surface shows which department has a critical status in context of information supply. Additionally, it is a basis for the BI architecture. It combines the different management levels with functional requirements. The resulting three components (PSM, maturity model, and gravity centers) engage all dimensions that cover specific requirements. These requirements are consolidated in a BI architecture that consists of different dimensions. Finally, specific requirements are deduced from the set of results. The requirements are weighted in consideration of comprehensibility and clarity. Additionally, they constitute the present situation.

It is evident that the current procedure of deducing requirements reflects a process view that describes a strategic development process. A BI strategy is necessary in order to describe how to deal with information (Totok 2006). It defines the information supply and its meaning for the company (Horvard 2008). A utility company should implement a BI strategy that is congruent with its company strategy. The BI strategy defines functional and technical requirements by implementing a BI architecture (Totok 2006). Whereas a BI architecture describes the information supply, a process model based BI roadmap constitutes a schedule in order to realize the architecture. Such a road map is the result of this case study and it supports in different phases how to reach the architecture.

5 RESULTS

This section presents the PSM, because it describes the survey results in the second phase. Additionally, gravity centers are explained. Its function is to present the impact of BI in the different utility applications. Finally, the requirements are mentioned and the evaluation method is illustrated.

5.1 Process-System-Map, Gravity Centres and BI Maturity

The process-system-map presents the results of the expert survey within the department managers. As already mentioned, two PSM are created in respect of unbundling: one for the utility company and one for the DSO. The map itself is built upon a two dimensional one. The first dimension deals with the business processes that are faced with BI within the utility. The PSM clarifies the data flow from incoming data, data transformation and outgoing data for each BI relevant business process in the utility industry. Due to this reason, relations between business processes become obvious. On the basis

---

7 The name process-system-map belongs to the presentation of main activities of each department, although business processes run across several departments.
of the risk assessment processes the PSM and its meaning for deducing BI requirements is explained. Furthermore, it identifies the status quo and is the basis for BI maturity classifications and the definition of gravity centers.

- First, the report recipients are identified. By identifying the recipients it becomes obvious which report aggregation level is necessary. This conclusion is important to deduce requirements in context of report presentation and drill down functions. Furthermore, requirements of report distribution can be deduced.
- Second, the incoming data objectives are examined. The examination provides information about requirements of data sources. The risk assessment department has both technical and functional data sources that have to be integrated. Therefore, the BI system has to integrate local sources, need interfaces to the geographical information system (GIS) and utility specific systems.
- Third, the outgoing data object is explained by describing its report presentation and distribution format. Hence, graphical functions for presentation are enumerated as requirements. Moreover, architectural artifacts, like client platform, supported browsers and clients, play an important role in distributing reports.
- Fourth, the PSM describes the data requirements. The description distinguishes between structured and unstructured data in general. The risk assessment department enumerates financial data, geographical data and maps. Financial and geographical key figures and their meaning are mapped. This mapping facilitates the allocation of revenues to distribution stations and net customers, as well as analyzing the state of distinctive assets. Therefore, a BI system must consider techniques like data plausibility check, simulation methods to forecast revenues, and also report features like quick views and calculation within a report. Moreover, working with customer data requires a sensitive handling.
- Fifth, the PSM lists these systems supporting the business processes. In case of risk assessment, SAP BW, SAP R3, GIS and MS-Office applications are needed. Therefore, a BI system should processes relational and multidimensional data sources. Furthermore, it needs certified interfaces to SAP BW and GIS. Especially the GIS and SAP IS-U interface is an industry specific BI-requirement.
- Sixth, the purpose of the business process is described to improve the functional understanding.

**Figure 1: Process-System-Map of the DSO**
Considering the PSM, the meaning of the controlling department in context of planning, forecasting, and budgeting within the utility sector is important. It has the function to consolidate reports from utility and unbundled sales and the DSO division. Moreover, it has to distinguish between financial and action planning. In context of financial planning, the generation of planned and actual data is a core application of BI. Additionally, analytical processes, like knowledge discovery in databases (KDD), forecasting and operation research methods become more important. Thereby, it has to be distinguished between DSO and Sales. A DSO customer refers to competitors of the utility company, whereas the utility customer is defined as an end consumer. Methods like data mining afford a better customer relation management (CRM) by analyzing customer behavior in order to plan and optimize marketing activities and improve customer service programs. Furthermore, incentive based regulation becomes interesting for DSO. This leads to the need of analytical tools that enable excellent grid capacity, power consumption and peak load forecast.

The PSM emphasizes the information flow within a business process. It describes the data transformation by explaining incoming data objects, their transformation, and the outgoing data to supply information requirements. Furthermore, the presentation is the basis for the further examination of necessary interfaces, especially if data are transformed manually. So, unbundling-sensitive data are identified.

Whereas the PSM describes the data centric view on business processes to understand the operational workflows and to deduce requirements for a BI system, it does not give information about the contribution of BI within the processes directly. Due to this reason, gravity centers are built to describe both the complexity of BI requirements and the data volume for each business process. A maturity classification also clarifies the contribution of BI within an utility company.

Besides the complexity of requirements, the data volume is very extensively, too. On the one hand, the sales department deals with a big set of data; on the other hand, the DSO is confronted with capacity data, geographical data, and legal data for the regulatory authorities. Whereas in the three mentioned departments both requirements and data volume are high, the controlling department concentrates on reporting. Therefore, it is not focused on analytical methods (in the meaning of using complex algorithms), but on reporting consolidation.

The maturity classification bases on the process-system-map. In context of the case study, the utility company is placed on the second level because it focuses on reporting. But in regards to the intensifying competition, analytical methods are of increasing importance. Therefore, it is the goal to reach the fifth level in the intermediate term (Chamoni & Gluchowski 2004). Furthermore, our classification is supported by current maturity classifications of BI within the utility sector, because it shows comparable results.

Figure 2: Gravity Centers and biMM

---

8 E.g. grid capacity planning and investment-planning.
5.2 Deducing Requirements

Formulating requirements is essentially in order to establish a BI architecture. Besides, copying requirements from recent studies like BARC (BARC 2007) and TEC (TEC 2008) offer a multitude of BI criteria. Many of them are not relevant for an adoption within the utility industry. It is evident that a list of requirements does not match with industry specific requirements. The PSM and the gravity centers are the basis for deducing requirements. By giving information about the business workflow, their functional complexity, the data volume, and the requirements ensure an appropriate contribution of BI to support business processes.

Nevertheless, not every requirement has the same significance. Hence, they are ranked on a scale from 1 to 10; a scale of 1 means that the requirement is relevant for a BI implementation, but it is not as important as other requirements. A scale of 10 reveals that the requirement is very important. The scaled requirements are the result of a second survey with the department managers after presenting the deduced requirements. The following Figure 3 and Figure 4 present in different radar charts the significance of each deduced requirement. The ranking is basis for a comparison of BI software and help evaluating an appropriate BI software for the utility sector.9

The following section presents the deduced requirements which are separated in four parts. The first part deals with general requirements like user support within the implementation and usage, vendor size, administration, and reference customers. Usually, big BI vendors are synonymous with good performance, but small vendors normally show greater engagement (BARC 2007). Moreover, the price is an important criterion, whereas it is separated into implementation cost, cost of development, training cost, and maintenance cost.

The second part contains requirements concerning the administration. These criteria determine a BI architecture and its implications to performance, especially load sharing, authorization scheme, and metadata administration. In context of performance optimization, the server architecture is decisive, because load advantages can depend on partitioned servers. Furthermore, administration with its authorization scheme is very important, because of unbundling requirements. The more flexible the user authorization is, the higher is the transparency. The reason is that authorization can be constricted at the lowest level. Moreover, they treat with system stability which is a very important demand, especially in context of energy trading and net metering.

![Figure 3: Requirements at BI in context of general and administration criteria](image)

The third part examines the platform, software interfaces, operating system integration, and the kind of data sources. The platform determines the load performance and user scalability. Additionally, the supported programming languages, supplied programming tools and interfaces have an impact to data integration. It is important that data integration is easy and powerful in order to adapt numerous data sources, i.e. relational or multidimensional databases. Within the utility industry, interfaces to common systems like SAP IS-U or GIS are necessary.

---

9 A software evaluation is not the paper’s object.
The final part determines the application functions that are discriminated between graphical representation, report distribution, and the usage of technical methods. It is of major importance for users, to control the layout in order to satisfy their own information need appropriately. Therefore, potential users take emphasizes on aggregation of different data sources, zero row and zero column elimination. Because utility companies have to send several reports to national authorities (like federal net agency), it is important to describe report artifacts with annotations (comments about the measures). Moreover, users prefer to have graphical visualizations to understand complex interrelations. Due to this reason, dashboards or cockpit functions are significant BI applications. In regards of cockpit functions, the ability of navigation within diagrams is mentioned. With regards to a report distribution, BI systems should have extensively printing functionality. Finally, the technical methods are essential for the report quality. As shown in the gravity centers, the sales and the DSO department deal with large data sets. Thus, users take emphasis on data plausibility verification. The following Figure 4 presents the results.

Figure 4: Requirements at BI in context of System Architecture and Report Features

6 CONCLUSION

The contribution of this paper is the deduction of requirements for a BI system in an unbundled utility company. Existing BI studies contain a multitude of criteria, but they do not consider industry specific requirements. It was not published which BI requirements are important and which applications are needed. Due to this reason, a process model based on a case study is presented. It takes emphasizes on process orientation by clarifying the data flow, specifying the input data, data transformation and output data. The process model is structured as follows. Firstly, the necessity of BI within the industry was explained. Since formerly vertical integrated utility companies have to unbundle their DSO to ensure open access to the grid, they are confronted with competition. Due to this reason, the information demand of the management increases to support the decision making process. Moreover, the management starts to think about their business processes, so BI can be an enabler for
harmonizing, collecting, and analyzing data. By implementing a BI system, the IT strategy is affected, because data have to be structured in an unbundling conformable manner. In fact, current studies offer a wide variety of requirements. But not every requirement is relevant for the adoption of BI in the utility industry and the significance of BI requirements is different. In order to deduce utility specific requirements and to scale its impact for a successful adoption, a process model was presented. The main processes are consolidated in a PSM for both utility and DSO based on a case study. In the following step, gravity centers were developed to show the impact in context of data volume and data complexity for each department. Moreover, the utility company was classified by using a maturity model. The goal of these artifacts is to show the meaning of BI on an aggregated level for each department, especially retail and DSO. In the following phase, specific requirements were deduced and weighted in respect to their meaning within the utility sector. Thereby, requirements are divided into four parts: general, administration, architecture, and application functions. According to the utility industry, system stability in context of net metering and energy trading is very important. In context of these two processes, data mining methods become more important. Furthermore, administration tasks play an important role, because of unbundling requirements. A BI system has to supply interfaces to common utility specific systems like SAP IS-U. Especially the DSO integrates a multitude of different data sources that contain both technical and financial data. These data have to be integrated in a single report. Finally, users demand a high layout control, easy navigation, and annotation function.

The presented enumeration of requirements constitutes a guide to a function oriented BI implementation and evaluation of BI software within the utility industry beyond unbundling. Concerning both deducing and weighting requirements, an evaluation of the results is necessary. Although, the developed process model and the deduced requirements are the outcome of a case study, current results cannot be presented. Even if the results of a BI implementation are deliverable, a verification of the artifacts has to be arranged. Such verifications can base upon models that compare the actual and the specified state and formulate the degree of compliance.

References


No Author (2004). Sechzehntes Hauptgutachten der Monopolkommission gemäß § 44 Abs. 1 Satz 1 GWB, 277-278.


FOCUS GROUPS AND CRITICAL SOCIAL IS RESEARCH: HOW THE CHOICE OF METHOD CAN PROMOTE EMANCIPATION OF RESPONDENTS AND RESEARCHERS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0303.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Critical theory, IS Philosphy, Research methodology, Health information systems / medical record systems / care delivery /</td>
</tr>
</tbody>
</table>
FOCUS GROUPS AND CRITICAL SOCIAL IS RESEARCH: HOW THE CHOICE OF METHOD CAN PROMOTE EMANCIPATION OF RESPONDENTS AND RESEARCHERS

Stahl, Bernd, De Montfort University, The Gateway, Leicester LE1 9BH, UK, bstahl@dmu.ac.uk
Chiarini Tremblay, Monica, Florida International University, Decision Sciences and Information Systems, Miami, Florida, United States, monica.tremblay@business.fiu.edu
LeRouge, Cynthia M., Saint Louis University, Decision Sciences/Information Technology Management, Saint Louis, Missouri, United States, lerougec@slu.edu

Abstract

Critical social research in information systems has been gaining prominence for some time and is increasingly viewed as a valid research approach. One problem of the critical tradition is that there is a lack of empirical research. A contributing factor to this gap in the literature is the lack of agreement on what constitutes appropriate methodologies for critical research. The present paper contributes to this debate by outlining the role that focus group research can play in the critical approach. The paper outlines the main characteristics of critical research with an emphasis on its emancipatory faculties. It then goes on to review the focus group method in general and gives an account of two research projects that used focus groups as a method of data collection. It is argued that focus groups can contribute to emancipation of researchers as well as respondents. This argument is built upon the critical theories of the two most prominent theorists currently relied upon in critical social IS research, namely Jürgen Habermas and Michel Foucault. Focus groups can improve communication and move real discourses closer to the Habermas’s ideas speech situation. At the same time, they can contribute to the challenging of prevailing orthodoxy and thereby overcome established regimes of truth in the Foucauldian tradition. The paper ends with a critical reflection of the shortcomings of focus groups as a critical method and of the specific approach chosen in this paper.

Keywords: critical social research, focus groups, research methodology, information systems
1 APPROACH AND MOTIVATION

IS researchers have embraced the critical approach as a viable and adequate alternative to study IT problems (e.g., (Orlikowski and Baroudi 1991; Trauth and Jessup 2000; Richardson and Robinson 2007; Stahl and Brooke 2008). The critical approach allows for a departure of the purely descriptive approaches commonly used in IS. As Richardson and Robinson (2007) indicate: “Critical research in IS serves as a reminder that IS is about the people as well as the business strategy, competitive advantage or systems development regardless of human cost”.

Despite increasing prominence of the critical approach in the field of information systems (IS), there is no agreement on appropriate methodologies for critical studies (McGrath 2005). While it has been observed that questions of methodology are not necessarily central to the critical agenda (Avgerou 2005; Walsham 2005), it would still be desirable to have a clearer understanding of the relationship between certain methods and critical research. The present paper contributes to this debate by offering a discussion of focus groups as a method of data collection and its relationship to critical research.

The potential of the use of focus groups for critical research has been investigated by social researchers (Kitzinger and Barbour 1999; Webb and Kevern 2008). Quality criteria of critical theory research are related to historical situatedness and its ability to overcome ignorance and misapprehensions which can facilitate emancipation. The focus group method can support these criteria as each participant comes in with their own experiential insights, which may change as they learn of the experiences and insights from other participants.

This paper contributes to the literature by discussing how the experience of focus groups as a critical method can be used and built upon in IS research. It argues that emancipation as the central aim of critical research can be furthered using focus groups. Emancipation can be supported for different stakeholder groups, most notably for researchers as well as focus group participants. Developing the argument requires us to give an account of the main concepts, i.e. critical research and focus groups, and to show that these are conceptually compatible. We will then describe and contrast two separate studies to support the argument empirically.

2 CRITICAL SOCIAL RESEARCH

In the field of IS, critical research is often referred to as a paradigm, which offers an alternative to the more established paradigms of positivism and interpretivism (Orlikowski and Baroudi 1991). In order to clearly distinguish our view of critical research, we follow the suggestion of Harvey (1990) and Klein (2008) and add the term "social" in order to render it clear that we follow a particular tradition of thought, which goes beyond the critical considerations of established facts, theories, and data that pervades all research. Critical social research can be defined using a number of different but interlinked characteristics. In this paper we follow Stahl's (2008)view that the most important defining feature of critical social research is its intention to promote emancipation. This critical intention finds its justification in a perception of the world as being imperfect and unjust and the belief that research should address this and aim to improve the status quo. This starting belief sets the critical approach apart from both the positivist and the interpretivist tradition, which are predominantly descriptive.

The concept of emancipation raises a number of complex issues. First, there is its definition. Even following a relatively wide and uncontroversial definition, such as the one suggested by Klein & Huynh (2004 p. 163): "Emancipation means that more people can achieve their potential to a greater degree," raises problems. How can we know what people's potential is and how can we know the best way for people to achieve it? What can the role of research be in promoting emancipation and what is to be done if the views of the research objects on their emancipation does not reflect the researcher's views. How do we avoid critical social research turning into a dictatorship of the intellectual (Stahl 2006)?
Despite these and other open questions, we believe that emancipation is central to understanding the critical view and that it explains most other characteristics of the approach. The interest in emancipation sharpens awareness of impediments and requires sensitivity to non-obvious facts and relationships. Critical social research has a history of concentrating on social structures that alienate humans, which notably include economic structures.

The emancipatory intention has implications for epistemology and research strategy. Critical scholars tend to be sceptical of truth claims that are based on an objectivist ontology. The constructed nature of social reality requires them to understand the process of construction of individual and collective life-worlds. This leads to a dialectical understanding of research, which is often described in direct opposition to positivist approaches. Critical researchers understand that knowledge claims are always contested and aim to understand under which circumstances they become accepted as truth. They also realise that all knowledge claims are based on particular interests and that there is no value-free and objective description of social realities. Knowledge must, therefore, be understood in its historical context. Given the scepticism of truth claims and the starting assumption that the world should be improved, critical social researchers tend to have an ambiguous relationship to empirical data, which can always only say what the world is like but rarely give indications of what it should be. This may to some degree explain the ongoing problem of methodology in critical social research in IS. In addition, the historically contingent nature of critical social research requires the researcher to continually reflect on their activities and the role they play in shaping the outcomes. Reflexivity is therefore a frequently cited aim of critical social research (e.g., Kvasny and Richardson 2006).

Another aspect of the critical approach that has bearing on the role of focus group research is that of theory. A considerable number of theories have been used in critical research. In the field of IS, one can observe that the work of two researchers has gained prominence: Jürgen Habermas and Michel Foucault (Brooke 2002). We will outline brief aspects of their theoretical positions which allow for a wider interpretation of the role of focus group research in the context of critical social research.

Habermas is often considered a leading representative of the Frankfurt School of social theory. Whereas the first generation of Frankfurt scholars were more closely aligned with Marxist thoughts, Habermas developed social critique by establishing the Theory of Communicative Action (TCA) (1981), which posits that there are different ways of coordinating social action and that the most desirable is that of communicative action. This means that interlocutors are accepted as equal and valuable individuals and that the aim of communication is the achievement of consensus on the basis of rational exchange of arguments. Central to this idea is the concept of discourse. In any speech act, the speaker raises a number of validity claims (truth, legitimacy, authenticity). Whenever these are contested, discourses take place where arguments are exchanged to clarify the differences and come to a conclusion. Such discourses presuppose the (counterfactual) ideal speech situation. This is a situation where all speakers are endowed with the same ability to express their views and are willing to accept each others' arguments. Due to the construction of discourses and their presuppositions, the TCA implies certain ethical views, which Habermas has developed in his discourse ethics (1983; 1991). Important aspects of his work that inform our interpretation of focus groups include the discursive nature of social reality, the ideal speech situation as an aim which indicates the way towards emancipation and the ethical nature of communicative interaction.

The second theorist who has gained prominence in critical social IS research is Michel Foucault. Similar to Habermas, he sees the importance of discourses in establishing social reality. Unlike Habermas, however, Foucault is interested in the question of who gets included or excluded from discourses and on what grounds (1971). Where Habermas implies a view of truth that is based on a consensus (albeit often a fictitious one), Foucault holds that there are "regimes of truth", which determine what can count as true in a particular context. In his extensive studies of different social realms, he undertook genealogies of historical discourses with the aim of determining the ways in which legitimate knowledge is created. For our purposes, some of the interesting aspects of his work include the emphasis on bodily discipline in establishing discursive dominance, the whole idea of regimes of truth as well as the importance of power relationships in discourses.
3 FOCUS GROUPS FOR CRITICAL RESEARCH

Focus groups have a long history in market research, social research (Morgan 1988), and more recently in medical research. However, they continue to be under-used in information systems research as a primary method of data collection, even though they can be particularly valuable in exploring and recognizing the socio-technical nature of information systems.

Powell et al. define a focus group as “a group of individuals selected and assembled by researchers to discuss and comment on, from personal experience, the topic that is the subject of the research” (1996 p. 499). The focus group draws upon respondent’s attitudes, feelings, beliefs, experiences, and reactions in a way that is not feasible using other field methods (e.g. one-on-one interviews, surveys, observation). An individual interview provides a one-dimensional perspective, whereas the focus group elicits multiple views through processes that can prompt reaction, synthesis of new information and experiences provided by others, and recall of the forgotten. Focus groups are a special form of group interviewing. General group interviewing has an emphasis on questions and responses between the researcher and participants. Focus groups are distinguished in that they rely on the collaborative construction within the group to inform topics provided by the researcher (Morgan 1997). The researcher prompts things to happen in a focus group in an organized fashion, as opposed to observational methods where the researcher waits for things to happen in a natural way.

Conversing among peers in a comfortable setting with limited moderator prompts may result in more information about organizational issues, knowledge structures, and interacting with users or IT staff than a participant may otherwise provide. Through group discourse, researchers tap into the many different forms of communication that people use in day-to-day interaction, including jokes, teasing, and arguing, (Kitzinger 1995). Gaining access to such variety of communication is useful because people's knowledge and attitudes are not entirely encapsulated in reasoned responses to direct questions, such as those presented in survey or one-on-one interview. Tapping into interpersonal communication is also important because this can highlight (sub)cultural values or group norms. Through analysing the operation of humour, consensus, and dissent and examining different types of narrative used, the researcher can identify shared and common knowledge. Furthermore, when an individual user’s or IT staff’s perspective is limited, focus groups can draw participants from the full set of involved stakeholders. And, because the participants are exposed to a wide range of ideas and perspectives of the information system of interest, these stakeholders can give a multi-perspective view of the technology of interest. The focus group method can highlight individual stakeholders’ reflections and build on these conceptual insights via other participants’ perspectives.

From critical theories such as Habermas's and Foucault's, we learn about the importance of power in any social interaction. Focus groups are no exception to this. Using the critical lens, scholars become more sensitive to power issues, which include obvious ones, such as the choice of focus group topics and selection of respondents, but also less obvious ones as expressed in moderation techniques or seating arrangements. It should not be assumed that groups are, by definition, inhibiting or that focus groups are inappropriate when researching sensitive topics. Group work can actively facilitate empowerment to discuss taboo topics because the less inhibited members of the group break the ice for shyer participants (Kitzinger 1995). Participants can also provide mutual support in expressing feelings that are common to their group, but which they consider to deviate from mainstream culture (or the assumed culture of the researcher).

Criticalists realize that one can never escape power relationships and that this would not even be desirable. The way to address them is to make them explicit and open them up to discussion and revision. The focus group discourse may expose asymmetric power relations in the organization and to allow silenced voices be heard. A focus group provides a means to obtain assessment from a group of stakeholders through collaborative construction of independent attitudes, experiences, and beliefs, rather than consensus or negotiation (Morgan 1997). From Foucault we can learn that power is always
bi-directional and all power leads to reactions, which are often unforeseen. The role of the researcher should, therefore, be discussed as well as other influencing factors, including the power that the focus group members have over the researcher. The critical perspective requires a different interpretation of research findings. Given the contingencies of power relationship and corresponding reality constructions, it is clear that no research, including focus group-based one, can produce a true representation of reality. Rather, all research needs to be understood as a contribution to the discourses that lead to accepted truth claims (Habermas) and that establish regimes of truth (Foucault). Focus groups are arguably better equipped to undertake this role than other means of empirical data collection because of their discursive nature, which allows for an easy recognition of factors of interest to critical research, such as ideology, hegemony, power, perceptions of rationality, etc. In the next section two research studies are used to illustrate how focus groups can be utilized for critical analysis.

4 DESCRIPTION OF TWO RESEARCH STUDIES

In this section we describe two different information systems studies within the healthcare context, which will be used to illustrate how the discursive nature of focus groups contributes to critical social research.

4.1 Emancipating Participants: Exploring Patient’s Concepts of Telemedicine Encounter Quality

The extent to which telemedicine positively impacts the health status of patients and equalizes access to care are social issues purported to be at the core of telemedicine. This is underscored by the motives of our organization and the socio-technical nature of the telemedicine encounter itself. The organizations participating in this study were part of a large government-run telemedicine network. Medical vision conferencing, the form of telemedicine addressed in this study, involves conducting patient encounters (i.e. clinical exams and consultations) using video conferencing equipment and related peripherals. The telemedicine network studied initiated medical video conferencing as a means to provide high-level speciality care to remote patients and as a means to reduce patient wait times for care, rural patient anxiety and time associated with travelling to an urban area for care, and travel reimbursement costs paid to patients. The purpose of this study was to identify the socio-technical attributes that contribute to a successful medical video conferencing encounter workflow and process.

A critical question arose among the research team regarding informants, namely who had the power (using Habermas's and Foucault's concept of power in social interaction) to define utility, efficacy, and process in the telemedicine context. In the telemedicine domain the business philosophy acknowledges that the patient is the icon that underpins each dimension of success and justifies each decision. Yet, this does not necessarily translate into participatory democracy or a public sphere that encourages debate deliberation, agreement, and action (Villa 1992). In a healthcare service encounter, the relationship between patients and medical providers is asymmetrical: the patient subscribes to the professional authority and knowledge of the physician in seeking medical care.

Even though patients may be considered partners in their medical care, patients have some degree of dependence upon medical care providers in the context of a medical encounter. Research indicates a patient’s dependence on a medical care provider (particularly physicians) is based on the physician’s command of an esoteric body of knowledge acquired through training and experience legitimizing the profession, the ability to justify his/her authority and the ability to evoke the client’s trust, confidence, and norm of obedience in following prescriptions for care (Parsons 1975). Despite this power difference, which is of intrinsic interest to the critical perspective, there are functional reasons for exploring patients’ views: their description of symptoms is required for diagnosis, they represent the “consumer” of healthcare, and they shape the competitive market of healthcare. It is, thus, plausible for a healthcare technology study to include data collected directly from patients.

Early forms of data collection in this study and reviews of telemedicine patient satisfaction literature recognize that studies indicate patients are satisfied with telemedicine (Collins, Nicholson et al. 2000),
but provide limited critical analysis from the patient perspective. Critical social research seeks, however, to question shared assumptions and unearth social realities behind such simplistic views. In their review, Collins, Nicholson, and Brown (2000) note some of the limitations of quantitative measurement of patient satisfaction including the propensity for patients to express high levels of satisfaction due in part to their reluctance to say the wrong thing or complain for fear of unfavorable treatment in the future. Direct observation provided relevant data for the general purposes of the study, but did not seem to adequately augment the patient perspective.

Focus groups were chosen to provide the means by which to give the patients a stronger voice in the study. In the spirit of ideal speech, every subject with the competence to speak and act is allowed to take part in the discourse (Habermas 1981). Patient encounters may be sporadic or limited precluding a rich breath of experience and the chance to share experiences with others to understand norms and nuances. We looked to focus groups to provide a means to obtain patient “expertise” through collaborative construction and learning from the related experiences of others (Morgan 1997). In addition, it was thought that the process of interactive discourse in the focus group may provided a collective strength to help empower the patient to overcome reluctance to say the wrong thing or provide an acquiescent response set.

Deliberate attempts were made in the design of the focus groups to enhance patient empowerment and the voice of individual patient participants that might find it more difficult to express their attitudes, desires, and needs. Efforts included the use of aliases by participant, rules to promote an “ideal speech” environment provided in writing before and read to the groups during the session, and the strategic placement of participants around the table.

An attempt to enhance and empower patients was also made through the use of a trained peer moderator the patients might see as someone they would generally have conversation with, as opposed to a research team member or healthcare provider. The assistant commanded a voice only at the end of the focus group to clarify areas of confusion, to facilitate complete information and accurate coding, and to provide a summary of interpretations of focus group comments at the end of the session. This closing discourse between researcher and focus group members was an initial effort towards validating interpretive analysis and some degree of sharing power with participants in the analysis process.

4.2 Emancipating the Researcher: Evaluating Data Quality in Health Policy Decision Making

The second study focused on issues of data quality. Healthcare enterprises can be regarded as ‘data rich’ (Abidi 2001) and the quality of this rich data is a crucial basis of strategic decision. This study proposed several data quality calculations and presentation methods meant to inform the decision maker of three data quality problems in a health planning/aggregated data environment: 1) missing data, 2) small sample sizes and 3) erratic trends in underlying data that could indicate incorrect data.

The focus groups were the final step in this study and were meant as an evaluation of the three data quality presentation methods. “Vignettes” or story lines were used to create fictitious decision scenarios based on current healthcare situations and sample healthcare data. (Table 1 contains some examples for cases of missing data). The strategy taken was to present the data with and without the data quality information in order to detect differences in the collective decision making processes. The moderator presented the experimental vignettes and encouraged the participants to play the role of a healthcare decision-maker. In order to analyze the data, the focus group guided the moderator in exploring the healthcare data. For example, participants were encouraged to ask the moderator to drill down or roll up the OLAP tool data in order to thoroughly understand and compare data for different counties as part of their decision making process.

<table>
<thead>
<tr>
<th>Vignette</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>There is evidence that smoking is responsible for most cancers of the larynx, oral cavity and pharynx, esophagus, and bladder, but possibly for other cancers such as kidney, pancreatic, cervical, stomach, and acute myeloid leukemia.</td>
<td>Is there correlation between smoking and certain types of cancer?</td>
</tr>
</tbody>
</table>
When Hispanics are diagnosed with a certain cancer they’re less likely to receive chemotherapy than non Hispanics.

Is there disparity in care between ethnic groups?

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Example Vignettes</th>
</tr>
</thead>
</table>

The main researcher in this study intended to use the focus groups to empirically demonstrate the utility and efficacy of presenting data quality information to for improved decision making. However, the open-ended discourses generated by the focus group techniques perplexed the researcher. The researcher’s background as computer scientist had guided her research topic choices, initial assumptions and an objectivist approach to research. In the focus group discourses, the researcher noticed non-obvious facts and relationships on how decision support tools were used in decision making, which were contrary to her initial assumptions. This motivated the researcher to retrospectively analyze the focus group transcripts using a critical lens.

5 ANALYSIS USING CRITICAL APPROACH

When focus group dynamics work well, the participants work alongside the researcher, taking the research in new and often unexpected directions (Kitzinger 1995 p. 300). In this section we illuminate some of the unexpected directions the focus groups in the two referenced studies introduced to cause questioning of assumptions. We use Habermas’ concept of ideal speech and Foucault’s concept of regimes of truth to highlight some unexpected directions and the questioning of assumptions.

5.1 Study One: Patient Focus Groups and the Ideal Speech Situation

Though attempts to create a microcosmic public sphere supportive of debate, deliberation, agreement, and action were made and signs of such were evident, there is no guarantee that each and every participant was emancipated in the discourse (Villa 1992). Participants asserted positive views of process and action in the telemedicine encounters as well as incidents of imperfect experiences and possibilities for enhancement that they had not previously revealed by organizational or research surveys or on record as complaints or compliments. As the focus group conversation transpired, shared stories seemed to evoke a greater attention to detail and recall and the ability to more closely examine the process with a critical lens. For example, comments regarding potential issues with the provider interacting with technology included:

“Well, I found that there was a lot of time wasted in adjusting the cameras. Half the time was adjusting the cameras and instructing the nurse to focus in on different parts of my body, rather than to talk to me directly about my symptoms & my problems. I’m sure it was important, but to me that was a negative. In person he wouldn’t have to adjust anything.”

These issues were not suggested through survey responses, even in open-ended survey questions, nor by patient reactions in direct observation of telemedicine encounters.

Empowerment moved from a description of their reality to a more collective inter-subjective vision of what an ideal medical video conferencing encounter should be like. In building this vision, participants would interject prescriptions for process and performance. For example, in discussing the medical assistant responsible for operating cameras in the room with the patient, one patient commented that the assistant should:

“… be familiar with the doctor, have, you know, just don’t have, it has to be someone that and make them like a team. She is that doctor’s representative here and so she has to know the doctor well enough that if he does say something and it can be taken three different ways she usually knows which way he means it.”

Focus group participants also prescribed special training for medical professionals involved in such encounters as demonstrated by the following comment, “they have some video training and to have it critiqued on their performance, because it’s really a performance, they’re on television.”

For many such prescriptions, it seemed ensuing comments rallied support. However, not all statements produced unison. There were differing opinions on the mix of in-person with video exams, though
most groups contained members that supported the first vision in person as a means to start discourse with the doctor as described below:

“There’s a lot you can learn about the doctor by the way he handles you and how close he gets and his demeanour. It's important to the care of the patient and that’s what you get from that first visit in person.”

Some patients also seemed liberated to report on issues of their own shortcomings and awkward moments and stories seemed particularly forthcoming once one focus group member “broke the ice”. For example, some patients commented that they were captivated by the “television” and thus missed some of the doctor’s directives. In follow-up, another patient recounted that they did not realize they did not have to speak very loudly; and when told they did not have to shout, the embarrassment provoked them to cease free-flowing conversation with the doctor. A couple of other patients indicated they overlooked or were not informed that they were not having an in-person exam and shared their reactions as described below:

“I was shocked. I did not know I was having a television exam. My wife was with me and she says who’s talking? I said it was somebody over there. I’d just tell them to give us just a brief orientation, so that they would not be shocked.”

The question for the research team was whether and how to share the “patient voice” expressed in the focus groups beyond intended research outlets. The telemedicine coordinator and a leading doctor of the participating organization with the most telemedicine sites were invited to participate in the coding process to afford a more direct means of “seeing the voice” of the patients than general reports to participating organizations. The practitioners were trained; and practitioner coding was compared to researcher coding. In addition to providing the practitioners with direct access to patient comments, this process provided a means of analytical discourse among the researchers and the practitioners. The convergence found between the coding of the research and practitioner teams provided another level of validation to the study. In the spirit of ideal speech, it also introduced an unexpected discourse into the study. Though practitioners are often the subject and source of data in IS research, they are not often brought into the analysis process. Through their elevated role in this study, the practitioners voiced contextual experience as well as their emotional and intellectual reactions into the design, experiences, and observations associated with the analysis process. The researchers felt this discourse complemented and expanded their capabilities and insight.

In response to patient responses, practitioners involved in the coding and reconciling process were alarmed by some patient comments indicating they had not received a telemedicine encounter orientation to suit their particular concerns and needs for information. Transformation is enlisted in critical research in response to insight and critique. Once management was apprised of the patient orientation issues, the researchers were asked to work with the practitioners to develop patient orientation materials with content reflecting patient orientation needs as indicated by the focus-group discussion. To continue the discourse with patients, a validating survey was sent to patients participating in the focus group by the research team to confirm analytical insights. This procedure of soliciting patient feedback on patient education materials provides further indications of research and the patient voice transforming practice.

5.2 Study Two: Questioning the Orthodoxy: Challenging regimes of truth

The second study investigated issues of data quality in the context of public policy health planning. The intent was to aid human decision makers to better understand the quality of the data they had and in the presence of potentially incomplete and unreliable information from multiple data sources. The underlying premise held by the researcher at the start of the study was that additional data quality information would be valued. Thus, the core issue of the study was what data quality information to present and how best to present this information. Focus groups were used to inform the design and to evaluate if the enhanced decision support tool would improve decision making for public policy health planning issues. The loosely structured format of the focus groups (as opposed to an individual interview where there are structured questions, or a controlled experiment, where all variability is
carefully controlled) had an interesting effect. First, it allowed the conversation to navigate to areas that the researcher had not intended to cover, such as, why data quality problems exist. This new discourse added rich information to the research topic and contradicted some of the initial assumptions held by the researcher. Secondly, as focus groups became more comfortable in sharing their ideas, when they stumbled on to a particular aspect of the topic that excited them or bothered them, their resulting solidarity shifted the power from the moderator to the participants. Finally, because of the unintended discussions, the problems that motivated the researcher to pursue the study appeared in a different light, and forced her to question whether the chosen research focus of presenting data quality information to decision makers was the most pressing to investigate or even of importance.

The major underlying assumption made by the researcher was that decision makers understood that poor data quality was prevalent in many data sources. But in contrast, it appeared that most decision-makers were under the impression that data quality did not cause problems in their organizations. They indicated that they knew which data sources are not reliable, and they just avoid using them. This is evident in the following participant’s comment:

“I think some of it’s probably documented but there are some variables like that every researcher would know like don’t even bother using it. It’s so bad. But you know- it has a reputation for being bad. And you might know that by reading the literature but most likely you’re probably gonna know by informal network.”

In another vignette, participants were asked to determine whether there was a correlation between smoking and certain types of cancer in the presence of significant missing data. The participants indicated it depended on the stakeholder. The information about the missing data, along with the drilling capability of the OLAP tool could be used to either argue either side. This was treated with humour, but it was apparent that data could, and was at times, manipulated for the desired outcome.

Another assumption held by the researcher was that decision-makers would use the provided data quality information for improved decision making. Yet, when presented with the system prototype showing data quality information, the participants collectively began to speculate regarding the reason for the data quality problem, rather the concentrating on using the tool to help them with the test case decision tasks. For example, in the vignette where participants were asked to explore healthcare disparities in cancer screening for Hispanics, the participants speculated on the cause for the missing data, illustrating discourse deviating from the researcher’s initial intention:

“Years ago we started to do press screening in migrant Mexicans - among migrant Mexicans and we tried to give them the free ones and they didn’t feel like they deserved that kind of care. That that was for rich people and wasn’t for them which would you know sort of figure in some of the cultural issues related to this beyond just sort of race.”

The solidarity that formed among the participants had an interesting effect. In one particular focus group that consisted of Ph.D. statisticians at a large hospital centre, participant solidarity actually shifted the power and control from the moderator to the participants. The group disliked the decision task vignettes that were chosen as test cases and refused to make decisions with the data that was presented to them. This was interesting because the examples were based on simplified versions real decision-making scenarios within their context. This particular group resisted any sort of “watering-down” of tasks that were similar to their every day tasks. In retrospect, the participants were incensed that the task design did not consider the situational nuances and unique decision making skills they had acquired by doing this kind of work. They felt that the data quality issues that were being covered minimalized the actual issues they faced. For example, again with the Hispanic disparities case one participant noted:

“My concern is that the census classification of Hispanic allowed for people that aren’t Hispanic, for documentation purposes...Hispanic is anything that doesn’t fall under white or black or that some people may not consider themselves Hispanic if they are of a mixed race.”

In a similar fashion, this particular group nitpicked each task scenario, and the data quality information was ignored. As the moderator tried to re-focus their attention on the data quality issues and the
related tool, they participants basically revolted and said it was just too much to consider, given other information that they also needed to consider in the decision making process. Their overwhelming consensus was that the addition of data quality information to the decision support tool would force them to consider too much information and would actually decrease their decision making efficiency:

“I think that ideally your decisions would be more effective, but I can see you might spend a lot of time if you are digging through that (data quality information), and that time could be wasteful in an of itself”

As the researcher analyzed the conversation and some of the observed resistance to considering the data quality information in the test decision-making scenarios, the complexity of decision making with flexible decision support systems such as OLAP became apparent. Albeit, data quality is an important aspect in decision making, it is just a small part of the myriad of issues in health policy decision making. In fact, inserting more doubt about the fitness of the data used for these tasks may actually exacerbate or corrode the processes that health planners and clinical researchers use to make decisions.

6 REFLECTIONS

Critical research aims to be reflective and question its own assumptions, beliefs, and outcomes. This paper seeks to come to a better understanding of the role that focus groups can play in critical research or, to put it differently, whether and how focus group research can further emancipation.

6.1 Emancipation of the respondents

The stories shared by patients participating in the first telemedicine study eroded elements of ignorance and misapprehensions in sharing what was beneath the surface regarding their medical video conferencing experiences not revealed by other methods of data collection. Positive experiences were not negated, but perceptions of the ideal sometimes shifted. In the telemedicine study, each patient had unique telemedicine exam experience and opinion of success that they brought into the group. The sharing of unique situations and experiences and commonalities evoked a new virtual reality of what it would take to achieve an ideal telemedicine process. Though some individual differences in opinion still existed, a shared common knowledgebase of histories (to varying degrees) transformed the participants. A richer analysis evolved in recognition that even positive situations can merit from a critical lens and multiple organizations applied results from the study to move to a higher level of service and understanding regarding medical video conferencing. Conscious improvement of communication situations in accordance with Habermas' ideal speech situation can improve individuals' understanding of technologies and also lead to a better understanding of different stakeholder groups (e.g. patients and clinicians). Focus groups are uniquely suited for such improvement of speech situations because by their very nature they are close to discourse situations. Making positive use of social dynamics can help researchers enforce emancipatory consequences of qualitative research by using focus groups.

6.2 Emancipation of the researcher

The interaction and often camaraderie that is evident in focus groups guided the researcher to reflect about some of pre-conceived notions about decision making in an important context, health public policy decision-making in our second example. It should be evident that giving information about the quality of the data should empower the decision makers to improve how they approach decisions. Using Foucault's concept of regimes of truth, we experienced the power of focus groups of questioning established beliefs. IS researchers tend to use an unproblematic view of agents' rationality which translates into the belief that high quality data will lead to good decisions. Exploring discourses during the focus group research showed that this orthodoxy is fundamentally flawed. Respondents were not as interested in data quality as standard theories suggest they should or would be. In reality, most decision makers paid limited attention to data quality issues in comparison to alternative realities about decision making that emerged in this context. Personal biases in decision making (such as
questionable racial stereotypes), speculation about reasons for data quality problems, stakeholder issues and limitations in the amount of data that a person can or is willing to process objectively, were in themselves problematic. These findings allowed the researcher not only to come to new insights into data quality issues in healthcare settings but, moreover, to question the very assumptions that the research project was based on. By coming to this new position, the researcher was emancipated from prevailing thoughts and free to explore alternative explanations of social realities.

6.3 Focus Groups as critical method

We believe that this paper gives a convincing account of the positive contribution that focus group research can make to the critical tradition in IS. It can be emancipatory for participants as well as researchers and there are diverse theoretical explanations of the critical faculties of the focus group method. This should not be misunderstood as saying that focus group should be viewed as a "critical method" per se. There are weaknesses to focus group research in general and to our approach in particular that need to be kept in mind.

When using the focus group method for critical research, one needs to be aware that it can never achieve the (transcendental) ideal speech situation. The researcher will always play a central role that has a tremendous influence on outcomes. The researcher sets the topic, established legitimacy of research question, selects participants, influences discourse situations, has control over physical resources, and decides how to use the data. While the process can be made as transparent and democratic as possible, it will always remain skewed towards the perspective of the researcher. An awareness and sensitivity towards critical topics is necessary for the focus group research to display its critical characteristics.

Another problem with our approach is that it is fundamentally functional, i.e. it did not question the technology and its context of usage. It is not obvious whether the socio-economic framework in which our investigations took place is truly geared towards emancipation. Accepting the description of healthcare as a market raises a number of problems. It means that financial considerations are central to healthcare which certainly not all critical scholars would find acceptable. This has to do with questions of distribution and justice. Using a Habermasian perspective, one can say that it represents the colonisation of the life-world which he would not find acceptable. Our approach is thus open to the charge of being fundamentally uncritical because it uses the critical repertoire of ideas to improve a social context and use of technology whose legitimacy per se is not certain.

A further and fundamental problem of all critical social research is the question of successful emancipation. Due to a number of conceptual and epistemological issues, most active critical researchers settle for the idea of "transformative redefinition" (Alvesson and Deetz 2000) as the hallmark of success. The idea is that emancipatory change can be effected by allowing people to reconceptualise their world. The question for us then is whether the focus group approach is truly conducive to such transformative redefinition and whether it is possible to provide evidence of them. Our account of the research suggests that the interaction during the focus groups has given rise to such redefinition but whether they were truly transformative is a difficult question.

One can summarise this paper by saying that we think that there is a good case to be made that focus groups are a research method that goes well with the critical approach. At the same time it raises a number of questions, which require the critical researcher to remain careful. If is of particular importance to underline the fact that the choice of whatever method, including focus groups, will not automatically render a piece of research critical. Criticality requires ongoing attention to issues of emancipation, which may nevertheless remain elusive.

References

Designing and evaluating an interactive video website for organizational learning

<table>
<thead>
<tr>
<th>Journal:</th>
<th>17th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2009-0162.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Organizational learning, Computer-mediated communication (CMC), Web 2.0, Evaluation</td>
</tr>
</tbody>
</table>
DESIGNING AND EVALUATING AN INTERACTIVE VIDEO WEBSITE FOR ORGANIZATIONAL LEARNING

Hrastinski, Stefan, Computer and Systems Science, Department of Information Science, Uppsala University, Box 513, 751 20 Uppsala, Sweden, stefan.hrastinski@dis.uu.se

Monstad, Therese, Media and Communication Science, Department of Information Science, Uppsala University, Box 513, 751 20 Uppsala, Sweden, therese.monstad@dis.uu.se

Abstract

The interest in audio and video technologies has surged as IT infrastructures and network capacities have improved. Surprisingly, there have been rather few studies on such emerging technologies in organizational settings. In this research-in-progress paper, we explore the impact of the use of an interactive video website, comprising videos promoting a company’s core values, on organizational learning. More specifically, we intend to study how the use of a video website affects the awareness of a company’s core values and whether this will also influence the behaviour of the employees. Two web surveys are being designed for the study. The first survey was conducted prior to introducing the video website. In this paper, we present initial results from the first survey. We are currently in the process of designing a follow-up questionnaire in order to assess cognitive and behavioural effects of introducing the video website. As a complement, we will also conduct interviews and observations of how the video website is used.

Keywords: Video website, information technology, organizational learning, cognitive development, behavioural development, internal communication.
1 INTRODUCTION

IT is increasingly used for internal communication. Often, employees drive this development by, for example, communicating with colleagues through instant messaging systems. In other cases, IT is introduced and promoted by the management of an organization. In the study of this paper, the management of a biotechnology and medical company decided to introduce an interactive website with functionality that is similar to YouTube. Figure 1 shows a screenshot of the website. Employees can watch and rate videos, post questions and share opinions regarding the content. We will refer to this technology as a video website. Notably, the interest in audio and video technologies has surged as IT infrastructures and network capacities have improved (Krishna, 2007). Surprisingly, there have been rather few studies on such emerging technologies in organizational settings. Videos that give a background to and describe the core values of the company have been produced. The reason for focusing on the core values was that the management felt that the awareness of the values, and how the values can guide work, needed to be improved. Over the last few years, the founder and management felt that the company stagnated and that the spirit of the enterprise had been lost. By implementing company core values with the overall focus on the entrepreneurial spirit, they believed that the employees’ attitudes and behaviour could improve. Also, they hoped that the employees would become more innovative.

Many scholars have stressed the importance of values and their effect on organizational performance over the years (Baum, Locke & Kirkpatrick, 1998; Collins & Porras, 1996; Ferguson & Milliman, 2008; Lencioni, 2002; Pattakos, 2004). Core values can be defined as describing what is important for a company and they can be used to guide employee behaviour (Ferguson & Milliman, 2008; Lencioni, 2002). A thorough implementation of carefully prepared values can lead to success while meaningless values might be highly destructive (Lencioni, 2002). Our case company has decided that it is time to enhance the employee’s knowledge about the company core values with the purpose to regain the entrepreneurial spirit in the company. In order to do so the core values, “simplicity”, “professionalism” and “innovation”, are being presented in four videos. In the first video, the history of the company is presented and it is emphasized that the core values have been important over the years, ever since the company was established. The founder of the company set up the business with these core values. The remaining three videos each present one core value.

In this research-in-progress paper, we explore the impact of the use of an interactive video website on organizational learning (OL). More specifically, we intend to study how the use of the video website...
affects the awareness of a company’s core values and whether this will also influence the behaviour of the employees. While there is some literature on the relationship between IT on OL (for a review, see Robey, Boudreau & Rose, 2000), the relationship between emerging technologies, such as Web 2.0 technologies, and OL has not been sufficiently explored. We extend previous research that has mainly been focused on text-based knowledge management systems, such as knowledge repositories of best practices, and communication systems, such as e-mail and groupware (Kane & Alavi, 2007; Robey et al., 2000).

While there are a number of competing theories that attempt to explain development, learning and change in organizations, the OL field has consistently emphasized these processes over the years (Argyris & Schön, 1978; Fiol & Lyles, 1985; March, 1991). The basic building block of most theories and models of OL is cognitive and behavioural development (Fiol & Lyles, 1985; Templeton, Lewis & Snyder, 2002). Cognitive development is commonly referred to as learning level (Fiol & Lyles, 1985) and can be described as developing shared understanding among members of an organization (Hedberg, 1981). Behavioural development is commonly referred to as change level (Fiol & Lyles, 1985) and can be described as the change of behaviour and actions of members of an organization (Argyris & Schön, 1978; Daft & Weick, 1984). Similarly, Argyris and Schön (1978) differentiate between espoused theories, i.e., theories we can state verbally, but which do not necessarily affect our behaviour, and theories-in-action, i.e., theories that actually affect our behaviour.

To explore the effects of the interactive video website on OL, we investigate the effect on cognition and behaviour, because these are established concepts and make it possible to extend previous research. This makes it possible to validate previous research in a new setting, contribute to developing a cumulative research tradition, and enables deeper exploration of foundational ideas (Kane & Alavi, 2007). More specifically, we explore how the video website can support or hinder cognition and behaviour, and posit the following questions: How and why does the use of the video website support or hinder an increased awareness of the company’s core values? How and why does the use of the video website support or hinder improved behaviour?

In the section below, previous research on IT, interactive video and OL is reviewed, and it is reflected on possible effects of using the video website. Then, the research setting and methods of the study are discussed. In the fourth section, we present initial results based on the first of two surveys. We conclude by discussing the next steps of our study.

2 INTERACTIVE VIDEO AND ORGANIZATIONAL LEARNING: LEARNING THROUGH TRANSMISSION OR INTERACTIVITY?

In this section, we first review research on IT and organizational learning and then discuss the potential impact of interactive video on learning. The section is concluded with a reflection on to what extent the interactive video website might support transmission versus interactivity.

2.1 Information technology and organizational learning

The relationship between IT and OL is receiving an increasing amount of attention because of the potential to affect organizational outcomes positively. In the literature on IT and OL, there are two streams of research: studies that apply OL concepts for implementing and using IT in organizations and studies focused on the design of IT applications to support OL (Robey et al., 2000). This study contributes to the latter of these streams, by designing and evaluating a video website for OL.

Information supports and stimulates learning and therefore it has been argued that IT can enable and facilitate OL (Janson, Ceccez-Kecmanovic & Zupančič, 2007; Robey et al., 2000). IT can effectively be used to support attributes of OL, such as acquiring, sharing, modifying, interpreting and storing information and knowledge (Templeton et al., 2002). Differences in OL when using IT can be related to characteristics of employees, such as individual learning rates (Kane & Alavi, 2007; March, 1991). For instance, IT may support OL in some situations (e.g., with fast learners) but may be ineffective or even hampering in other situations (e.g., with slow learners). IT has been described as a double-edged
sword, because it has the potential to both help and hinder learning (Kane & Alavi, 2007; Robey et al., 2000). Previous research suggests that successful use of IT as support for OL is dependent on the intertwining of both technical and organizational factors.

The right IT-enabled learning mechanism employed under the right conditions can benefit OL. However, the wrong mechanisms for particular conditions can be detrimental (Kane & Alavi, 2007). Thus, IT and its effect on OL can only be fully understood by examining the ways it is activated in organizational contexts (Leonardi, 2007). Members of organizations often use IT in different ways than managers or system designer intended (Orlikowski, 1996). By adopting the social information processing model (Salancik & Pfeffer, 1978) when studying how organizational members use media, scholars found that members of a work group shared similar attitudes toward IT and use IT in similar ways (Fulk, 1993; Fulk, Steinfield, Schmitz & Power, 1987). Another study by Fulk (1993) found that social influences on attitudes toward IT and the use of IT are consistently stronger if members of a work group are highly attracted to each other than if they feel low attraction to their work group. However, these social constructivist models are not deterministic, which means that they do not presume that social interaction is the sole explanation for acknowledged attitudes and behaviours toward certain IT (Fulk, 1993).

Particular tools, such as information repositories, where information can be stored and retrieved, promote homogeneity and tend to result in improved learning for the short term. Other tools, such as online communities, which connect employees who share common interests, promote heterogeneity and exploration and tend to lead to better long-term results, but are less effective for leveraging knowledge in the short term (Kane & Alavi, 2007). Robey et al. (2000) argue that communication technologies are useful for supporting organization-wide communication, which may support questioning of static assumptions and creation of new knowledge.

### 2.2 Interactive video and learning

As noted in the introduction, the interest in audio and video technologies has surged as IT infrastructures and network capacities have improved. However, there have been remarkably few studies on such emerging technologies in organizational settings. Studies that focus on organizational settings and emerging technologies emphasize the tremendous impact of emerging technologies on the interaction between corporations and their stakeholders, e.g. employees, shareholders and customers (Argenti, 2006; Vielhaber & Waltman, 2008). When reviewing research, we learned that most recent studies on video focused on technical issues, video conferencing or the use of video in educational settings. A number of studies on YouTube and similar technologies are also emerging, although we did not find such studies in organizational settings.

In a literature review on learning from video, Cennamo (1993) identified three key factors that can be expected to influence the mental effort of users. First, characteristics of the media, such as the content of the videos and the questions for discussion, may affect learning. Second, characteristics of the user, such as the perception of the usefulness of interactive video, may affect learning. Third, characteristics of the task, such as the purpose interactive video is used for, may affect learning.

Previous research has indicated that visual information is more memorable and that the combination of audio and visual information can increase comprehension and retention (Baggett, 1984; Kozma, 1991). These findings are supported by a more recent study with 147 psychology students. This study reported that digital video was more effective than text for presenting real-life situations in order to enhance learner satisfaction, comprehension and retention (Choi & Johnson, 2007). In studies of a management course and an English course, it was argued that digital video promoted contextual aspects of learning (South, Gabbitas & Merrill, 2008) and emotional involvement in the learning process (Hakkarainen, Saarelainen & Ruokamo, 2007). South and colleagues (2008) argue that digital video can provide situated cognition, i.e. connecting knowledge to relevant activities, contexts and cultures in which it is used (Brown, Collins & Duguid, 1989). When they compared video with a face-to-face setting, it was concluded that the first setting was characterized by learning through reflection, while the latter setting included more collaboration and conversation. In a study of YouTube, Lange (2008) suggested that social network technologies may need to support publicity or privacy depending
on different individuals’ and groups’ social needs. She argued that technical features could beneficially give users control in deciding whether to create public or private interactions, which may stimulate participation.

Drawing on the above, the use of the interactive video website has the potential to promote homogeneity and short-term learning of the company’s core values. OL theorists have, however, argued that mass communication, such as video messages, or mass meetings where managers communicate core values, are often not sufficient. Top-down communication is then applied and the transmission view on communication is in focus. Members of the organization commonly hear very different things in these situations and therefore the expectations of management are often not met (Schein, 1993). Successful organizational change communication is instead based on the creation of opportunity for the organizational members to be part of the communication and contribute to the organizational framework (Langer & Thorup, 2006; Weick, Sutcliffe & Obstfeld, 2005). Weick (1979) has gone as far as to argue that communication constitutes organizations and that people organize in order to solve equivocal information. A transmission view, i.e. top-down communication with the idea of the organization as a single body and the management voice as the one and only (Langer & Thorup, 2006), would in that sense be fatal for the organization. In order to avoid the transmission view, our case company has encouraged the members of the organization to ask questions and discuss the content of each video. Questions, such as how a particular aspect of the core values can be implemented in daily work, are suggested for discussions on the website. Thus, the interactive features of the website have the potential to, but will not necessarily, support discussion and exploration of ideas.

3 RESEARCH SETTING AND METHOD

The company studied in this paper is a biotechnology and medical device company that primarily develops, manufactures, markets and sells medical products. It was founded in 1987 and has grown rapidly ever since. The company has more than 700 employees in 20 countries, with 480 at the company’s head office and production facility in Sweden. Traditionally, meetings have been organized regularly which most staff has attended. However, as the company has grown, these meetings have turned into information meetings with a transmission view in focus, i.e. management is informing the staff but there is little room for discussions and questions from the staff. The management is therefore searching for new communication forums where questions and feedback from the staff are better supported and encouraged. This has included the use of an intranet and, more recently, the introduction of an interactive video website described above and four video productions. The website and videos were produced by a small company specialized on video, audio and website development. A video website was preferred over streaming video, because of the potential to support discussion and contributions from employees. The videos describe the company core values and include employee stories. The reason for focusing on the core values was that the management felt that the awareness of the values, and how the values can guide work, needed to be improved. With these stories, employees transform tacit knowledge about company core values into more explicit and usable knowledge by giving examples on how they implement company core values into their daily work. The employees are given the opportunity to comment on the videos, ask questions and rate the content. Currently, it is not compulsory to use the website and the discussions are not moderated. It is possible to give anonymous comments, although all participants so far have at least stated their first name. The question is if the interactive video website can support OL and lead to greater integration of company core values into the daily work (Weick et al., 2005).

3.1 Questionnaires

Two web surveys were designed for the study. The first survey was conducted in October 2008 prior to introducing the video website. The second survey is planned to be conducted in August 2009, after the introduction of the website and the four videos. By comparing the results of the surveys, it will be possible to explore the effects of the use of the video website on OL. The development and validation
of measures of OL is an important research contribution in itself as the most critical area of importance has been argued to be the development of methods for measuring OL and the impact of learning on organizations and their performance (Lyles & Easterby-Smith, 2003). A new video will be released every second month. The first questionnaire collected descriptive data describing the respondents and their preferred modes of communication. It also included measures on cognition and behaviour in relation with the company’s core values. The second questionnaire will also include measures on interactivity and employee satisfaction when using the video website.

For cognition, we drew on the concept of cognitive development (Fiol & Lyles, 1985), defined as developing shared understanding among members of an organization (Hedberg, 1981). Shared beliefs play a vital role for enabling the improvement of actions across an organization (Senge, 1990). However, cognitive development does not necessarily reflect behavioural development (Fiol & Lyles, 1985). For example, in the study of this paper, the use of the video website might support an increased awareness of the company’s core values but this does not necessarily mean that the organizational behaviour will change. The measure included six items and achieved a high level of reliability (Cronbach’s alpha = 0.88). For example, one item stated: “I can describe the company’s values for a friend”.

For behaviour, we built on the concept of behavioural development (Fiol & Lyles, 1985), defined as the change of behaviour and actions of members of an organization (Argyris & Schön, 1978; Daft & Weick, 1984). Behavioural development does not necessarily reflect cognitive development (Fiol & Lyles, 1985). For example, in the study of this paper, the use of the video website might stimulate employees to reflect on work practices, but behavioural changes do not necessarily reflect the company core values. We created eight items. The measure achieved a high level of reliability (Cronbach’s alpha = 0.83). For example, one item stated: “I work on the basis of the company’s values”.

In the second questionnaire, we will also include a measure on interactivity. In order to understand the influence of the video website on OL, it is important to understand whether they use it in an interactive way. The use of the video website might support heterogeneity, where employees contribute and share ideas, which has been argued to be one of the key dimensions of OL (Robey et al., 2000). The measure will mainly include items adapted from Webster and Hackley’s (1997) survey instrument. For example, one item is: “I was comfortable when communicating on the web site”.

We will also include a measure on satisfaction. It is important to assess whether employees enjoy using the video website. If employees do not want to use the website it is unlikely that it will support OL. The measure will mainly include items adapted from Webster and Hackley’s (1997) survey instrument. For example, one item is: “In the future, I will try to avoid using the web site” (reverse-coded).

It was estimated that 450 employees were located at the Swedish site when the first questionnaire was distributed. Out of these, 280 employees completed the questionnaire, representing a response rate of 62 percent. Their mean age was 40 years with ages ranging from 24 to 63 years. The mean for duration of employment at the company was 5 years, ranging from 0 to 17 years. Sixty-one percent (n=170) of the respondents were women.

3.2 Interviews and observations

Rather than simply surveying whether OL took place, we also want to explain why the members of the organizations were affected by the video website initiative. Individual interviews will be used in order to include opinions that employees often prefer to express in more private settings (Silverman, 2006). We will continuously observe how the video website is being used. Comments and discussions on the video website are stored and can subsequently be analyzed. We also have access to quantitative data, such as how many times each video has been accessed.
4 INITIAL RESULTS

The management of the company believed that employees were not aware of and affected by the company’s core values, which were underlying reasons for why they decided to produce a website with videos. However, as displayed in Table 1, the cognition measure indicates that employees are aware of the core values (M=5.2). This was especially apparent when asked whether they felt they had good knowledge on the company’s values (M=5.8). As shown in Table 2, the mean for the behaviour measure was slightly lower (M=5.0). Many employees felt they worked on the basis of the company’s values (M=5.4) and that these values could influence the company’s future positively (M=5.5). An interesting finding is that the means were much lower when respondents were asked if it is difficult for employees to work on the basis of the values (M=4.1) and if they think that most employees work on the basis of the company’s values (M=4.3).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have good knowledge on the company’s values.</td>
<td>5.8</td>
<td>0.9</td>
</tr>
<tr>
<td>I can describe the company’s values for a friend.</td>
<td>5.3</td>
<td>1.2</td>
</tr>
<tr>
<td>I can describe the company’s value “simplicity” for a friend.</td>
<td>5.1</td>
<td>1.4</td>
</tr>
<tr>
<td>I can describe the company’s value “professionalism” for a friend.</td>
<td>5.1</td>
<td>1.3</td>
</tr>
<tr>
<td>I can describe the company’s value “innovation” for a friend.</td>
<td>5.2</td>
<td>1.3</td>
</tr>
<tr>
<td>I think that most employees at the company are aware of the company’s values.</td>
<td>5.1</td>
<td>1.3</td>
</tr>
<tr>
<td>Total</td>
<td>5.2</td>
<td>1.3</td>
</tr>
</tbody>
</table>

Table 1. Descriptive data for the cognition measure (n=280).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>The company’s values influence my daily work.</td>
<td>5.2</td>
<td>1.1</td>
</tr>
<tr>
<td>I work on the basis of the company’s values.</td>
<td>5.4</td>
<td>1.1</td>
</tr>
<tr>
<td>I feel that it is difficult for us to work on the basis of the company’s values. (reverse-coded)</td>
<td>4.1</td>
<td>1.5</td>
</tr>
<tr>
<td>I think that the company’s values can influence it’s future positively.</td>
<td>5.5</td>
<td>1.2</td>
</tr>
<tr>
<td>I think that most employees work on the basis of the company’s values.</td>
<td>4.3</td>
<td>1.4</td>
</tr>
<tr>
<td>The company’s value “simplicity” affects my work.</td>
<td>5.2</td>
<td>1.3</td>
</tr>
<tr>
<td>The company’s value “professionalism” affects my work.</td>
<td>5.1</td>
<td>1.4</td>
</tr>
<tr>
<td>The company’s value “innovation” affects my work.</td>
<td>5.1</td>
<td>1.3</td>
</tr>
<tr>
<td>Total</td>
<td>5.0</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Table 2. Descriptive data for the behaviour measure (n=280).

The respondents were also asked how they felt they had gained information about the company’s values (see Table 3). Prior to this study, one web video had been distributed to the staff by sending a link by e-mail. In the video, the CEO discussed the future of the company and the company’s core values. 116 of the respondents (41%) felt they had gained information concerning the company’s core values through web video. The respondents were also asked how they would like to gain information (see Table 4). Most employees preferred other means of communication as compared to web video.
<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass meeting</td>
<td>161</td>
<td>58%</td>
</tr>
<tr>
<td>Intranet</td>
<td>148</td>
<td>53%</td>
</tr>
<tr>
<td>Workshop</td>
<td>145</td>
<td>52%</td>
</tr>
<tr>
<td>Web video</td>
<td>115</td>
<td>41%</td>
</tr>
<tr>
<td>Personal with manager</td>
<td>87</td>
<td>31%</td>
</tr>
<tr>
<td>Newsletter</td>
<td>78</td>
<td>28%</td>
</tr>
<tr>
<td>E-mail</td>
<td>42</td>
<td>15%</td>
</tr>
<tr>
<td>Other</td>
<td>31</td>
<td>11%</td>
</tr>
</tbody>
</table>

Table 3. How employees felt they had gained information about the company’s core values (n=280).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal with manager</td>
<td>2.4</td>
<td>1.6</td>
</tr>
<tr>
<td>Mass meeting</td>
<td>3.0</td>
<td>1.8</td>
</tr>
<tr>
<td>Intranet</td>
<td>3.1</td>
<td>1.6</td>
</tr>
<tr>
<td>Email</td>
<td>3.7</td>
<td>2.0</td>
</tr>
<tr>
<td>Newsletter</td>
<td>4.4</td>
<td>1.6</td>
</tr>
<tr>
<td>Web video</td>
<td>4.5</td>
<td>1.8</td>
</tr>
<tr>
<td>Workshop</td>
<td>5.7</td>
<td>1.9</td>
</tr>
<tr>
<td>Other</td>
<td>7.7</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Table 4. Preferred means of communication (n=280, reverse-coded).

In Table 5, a correlation matrix is presented. It includes descriptive variables, the cognition and behaviour constructs, and the variables video experience and video preference. Video experience refers to whether employees felt they gained information concerning the company’s core values through web video. Video preference refers to the degree employees would like to receive information through web video. A number of strong correlations were identified. Older members of the organization were more likely to prefer web video for gaining information (r=0.19, p>0.01). Employees with a higher level of education were more likely to state that they had gained information concerning the company’s core values through video (r=0.20, p<0.01). In line with the OL literature, there is a strong correlation between cognition and behaviour (r=0.37, p<0.01), i.e., shared understanding among members of an organization is closely related with their behaviour and actions (Senge, 1990). A correlation, although it should be noted that it is rather weak, was also found between whether employees felt they had gained information concerning the company’s core values through web video and whether they felt their behaviour was in line with the core values (r=0.13, p<0.05).

<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>Age</th>
<th>Education</th>
<th>Years empl.</th>
<th>Video exp.</th>
<th>Video pref.</th>
<th>Cognition</th>
<th>Behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.04</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>-.06</td>
<td>.01</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years empl.</td>
<td>.03</td>
<td>.22**</td>
<td>-.12*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video exp.</td>
<td>.07</td>
<td>.11</td>
<td>.20**</td>
<td>.09</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video pref.</td>
<td>-.06</td>
<td>.19**</td>
<td>-.03</td>
<td>-.01</td>
<td>.23**</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognition</td>
<td>-.04</td>
<td>.03</td>
<td>.03</td>
<td>.11</td>
<td>.11</td>
<td>-.15*</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Behaviour</td>
<td>-.01</td>
<td>.11</td>
<td>-.05</td>
<td>.07</td>
<td>.13*</td>
<td>-.04</td>
<td>.37**</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 5. Correlation matrix.

* Correlation is significant at the 0.05 level (p < 0.05).
** Correlation is significant at the 0.01 level (p < 0.01).
5 INITIAL CONCLUSIONS AND FUTURE RESEARCH

The initial analysis of the first questionnaire reveals a number of interesting findings. In general, most employees felt they had good knowledge on the company’s values and felt they worked on the basis of the company’s values. However, when asked if they believe that their colleagues work on the basis of the company’s values, respondents were neutral. Notably, these findings differ considerably from the perceptions of the managers, who believed that the employees had limited understanding of the company’s values. It can also be noted that the employees prefer other means of communication as compared with web video. When we have conducted the second questionnaire, we will assess whether their opinions regarding web video have changed.

Rather unexpectedly, older members of the organization were more likely to prefer web video for gaining information, which contradicts research calling for ways to engage the “YouTube generation” (Duffy, 2007). Thus, the relationship between generations and new media are more complex than is often assumed – web video does not necessarily cater the needs of younger generations in organizational settings just because they are familiar with video websites. We have also learned that education might help explain why some individuals felt they gained information concerning the company’s core values through web video, while others did not. Although it should be noted that all employees at the company have computer access, one explanation might simply be that employees with a higher level of education use computers to a larger extent than those with a lower level of education who are more likely to work in the production line.

The presented results are based on self-reported data, which probably is the main limitation of this research-in-progress paper. However, we believe that these data give an indication of the impact of the use of an interactive video website on OL. As a complement, we will conduct a follow-up survey, observe the use of the video website and conduct interviews. We are currently in the process of designing a follow-up questionnaire in order to assess behavioural and cognitive effects of introducing the interactive video website. The development and validation of measures of OL is an important research contribution in itself as the most critical area of importance has been argued to be the development of methods for measuring OL and the impact of learning on organizations and their performance (Lyles & Easterby-Smith, 2003). As a complement, we will also conduct interviews in order to better explain the results of the surveys. For example, the results of the surveys can indicate whether the cognition and behaviour of the employees were improved in line with the core values, but they cannot explain how and why this occurred. Combining insights from conducting surveys, interviews and observations of the use of the video website will enable cross-validation of findings through triangulation and a richer, fuller description of the effects on OL. In our further research, we intend to explore questions, such as the following: Which are the cognitive and behavioural effects of using the video website? How do other factors, such as demographic factors, interactivity and satisfaction, influence cognitive and behavioural effects of using the video website?

References


MULTI-LEVEL ORGANISATIONAL SENSEMAKING AND LEARNING AS ENABLERS OF INFORMATION SHARING TO MITIGATE IDENTITY AND RELATED CRIMES

Jamieson, Rodger, University of New South Wales, ANZAC Parade, Kensington, NSW 2052, r.jamieson@unsw.edu.au

Land, Lesley, University of New South Wales, ANZAC Parade, Kensington, NSW 2052, l.land@unsw.edu.au

Smith, Stephen, University of New South Wales, ANZAC Parade, Kensington, NSW 2052, Stephen.Smith@commerce.nsw.gov.au

Stephens, Greg, University of New South Wales, ANZAC Parade, Kensington, NSW 2052, g.stephens@unsw.edu.au

Winchester, Donald, University of New South Wales, ANZAC Parade, Kensington, NSW 2052, d.winchester@unsw.edu.au

Abstract

The practice of information sharing as an approach to mitigate identity fraud within, between and from Australian organisations is investigated. Organisations need more flexibility to verify legitimate customers in information system (IS) environments where ‘know-your-customer’ laws or protocols exist for ‘customer-not-present’ or ‘new customer’ transactions. An IS sharing approach is needed, whilst also protecting individual consumers, entities and other stakeholders from identity (identity theft, identity deception, and identity fraud) or related crimes (money laundering, terrorism, trafficking – people, drugs, weapons, or illicit material) and from the breach or abuse of private ‘identity’ information. Privacy, data provisioning laws and protocols need to be adhered to when sharing information whether verbally, coded, encrypted, or interrogating shared data sets through data mining, profiling or other identity crime mitigation techniques. Most information sharing literature to-date focuses within the organisation-level. Our contribution is across organisations as well as within and across, at the sector-level. Asymmetry in information shared between organisations is reduced when the aim is to mitigate a common phenomenon or cause, such as identity crime. However, this sharing relies on trust, integrity, and confidentiality between organisations both within and across industry sectors as well as between the private and public sectors.

Keywords: Identity and Related Crimes, Identity Fraud, Information Sharing, Sensemaking in Organisations, Organisational Learning, Information Systems Security (ISS).

Acknowledgements

The authors wish to acknowledge the assistance and cooperation of participants from the organisations sponsoring this Identity Fraud Linkage Research Project and to AUSTRAC Consortium and the Australian Research Council for their research grant.
INTRODUCTION

Organisations and government agencies since the September 11, 2001, United States (US) terrorist attacks, are often confronted with a ‘know-your-customer’ identity verification requirement for business dealings to mitigate identity fraud and related crimes. However, their preference is for mutual customer trust and integrity, and to not incur search costs and time tradeoffs to prove their customers identity indirectly. Indirect identity verification varies from simple questions where answers should only be known by the other party, to in-depth database cross-checks behind-the-scenes. Dichotomies between privacy of ‘identity’ information and identity (fraud, theft, deception) and related crime (money laundering, terrorism, or trafficking – drugs, people) mitigation have been reported in industry, but they have not yet been fully explained in the academic literature (Main & Robson 2001). To date, identity fraud mitigation has mostly been at the organisational-level (Jamieson & Winchester & Smith 2007), with some exceptions (Federal Deposit Insurance Corporation (FDIC) 2004, Middlemiss & Gupta 2007). Identity fraud “refers to the gaining of money, goods, services or other benefits through the use of a false identity” (Australasian Centre for Policing Research 2006, p. 9). ‘False identity’ is procured by prior identity theft or identity deception acts. Identity theft is where identity details are stolen, while identity deception is when information details or documentation are made up, changed, or lent (or some mix of) for real or fictitious individuals or entities (Jamieson & Stephens & Winchester 2007b). Identity crime is estimated to cost about A$1 billion in Australia (Cuganesan & Lacey 2003, Jamieson & Smith & Stephens & Winchester 2009, see Table 1).

The aim of this paper is to examine the implications of information sharing options available to policymakers in the public and private sectors (in Australia) targeted by identity fraudsters. In this paper, we contribute to the literature and fill an observed gap by examining information sharing (also known as data sharing, data matching and computer matching (in the US) in a database or data set context) laws, protocols, and practice as an identity fraud mitigation approach in Australia (Gordon & Willox 2006, see p.6). We use both an organisational sensemaking and learning lens for this examination in an IS information/knowledge management context. The selection of organisational sensemaking and learning models were based on our aim. Identity crime is an “instance of sensemaking because it involves identity, retrospect, enactment, social contact, ongoing events, cues, and plausibility” (Weick 1995, p. 3). Similarly, the learning lens was chosen because we wanted to understand the identity crime proof of identity (POI) issuers/users learning stage of this new phenomenon. A prerequisite for computerised information sharing across data sets is a common architecture (i.e., servers) that will allow them to communicate in a sharable format. Our analysis revealed that public sector agencies have formally implemented various information sharing strategies (e.g., protocols) to mitigate and detect identity and related crime acts. However, private sector organisations tend to share information in a less formal often bilateral context (verbal) or purchase data outright from third party vendors to help verify an identity. A growing trend is for private industry to reciprocate information sharing with government agencies in order to combat crime for example, fraud in tax or welfare agencies. Information sharing in the government context is (mostly) the computerised comparison of two or more government agency data sets of ‘identity’ and other records which relate to the same individual or entity through a unique identifier. In general, the sharing of organisational information between employees of other organisations on a bilateral or multilateral basis requires mutual trust to avoid asymmetrical knowledge transfer. Symmetrical information sharing or knowledge transfer in this paper would be defined as not holding back identity and related crime perpetrator details or their modus operandi from others in the information sharing ‘group’. There is also a need to create economic incentives to avoid free-rider problems (Gal-Or & Ghose 2005, Gordon & Loeb & Lucyshyn 2003). Gal-Or and Ghose’s (2005, p. 186) work suggested that information technology (IT) security information sharing may be more valuable in some instances than others, for example “when product substitutability is higher, implying that such sharing alliances yield greater benefits in more competitive industries”.

Proceedings ECIS 2009
The next section reviews the literature. Section 3 describes theory and sensemaking and learning models. Section 4 explains our research methodology and Section 5 discusses information sharing results within and across organisations and sectors. Section 6 concludes and suggests future research.

2 LITERATURE REVIEW

Human ‘identity’ information consists of biometric, attributed and biographical attributes. Biometric attribute examples include: fingerprints; signature; hand geometry; and vein recognition. Attributed characteristics include: name; and mothers maiden name. While biographical attributes include: education and vocational history. Collectively, these attribute groups form an individual’s proof of identity (POI) documentation in Australia for identification purposes in opening bank accounts or accessing government welfare entitlements for example. Personal identifying information (PII) is closely related to an individual’s POI because of the knowledge sharing properties of issuance and confirmation. Personal identifying information has been referred to as ‘knowledge-based identification’ (Clarke 1994). Identification is the association of data with a particular individual. Identity authentication is the process whereby evidence of identity is assessed in order to establish a sufficient degree of confidence that data is being associated with the correct individual. In the context of information systems (IS), the purpose of identification is used to link a stream of data with an individual in a database (Clarke 1998). Individual and mass data surveillance can be enhanced by sharing information or data from distinct sources (knowledge management repositories i.e., databases). In addition, a more elaborate data compilation can be interrogated or analysed using data matching (Clarke 1994). In the context of identity validation, cross organisational data matching provides a means of enhancing the degree of confidence an organisation has that the individual they are dealing with is legitimate. This is critical if “other means of confirmation are unavailable, although of course, the mere existence of a matching identity record on an external database is not sufficient to ensure the identity is genuine” (Main & Robson 2001, p. 22).

Identity crime perpetrators have been categorised as organised crime, insiders, sophisticated, and opportunistic perpetrators (Jamieson & Stephens & Winchester 2007a). Perpetrators breach the privacy of individuals and other entities, to steal from or deceive their victims, obtaining their POI documentation and related PII, such as, bank account passwords. Provided with these details, perpetrators target customer bank accounts, retailers for goods and services, or a government agency for benefits thus, committing identity fraud acts (Jamieson & Land & Sarre & Steel & Stephens & Winchester 2008). Organisations need to be constantly investigating methods of improving procedures and systems to stay ahead of perpetrator innovations (Jamieson & Winchester & Smith 2007). The nature and extent to which individuals or organisations are able to share identity information as an approach to mitigating identity fraud perpetrators is based on two basic principles. Firstly, the rule of law and secondly, on the property rights of those holding private information (Gal-Or & Ghose 2005).

Strategies and policies (e.g., information sharing) to mitigate identity and related crimes have usually been modelled at the organisational-level (Cuganesan & Lacey 2003, Jamieson & Winchester & Smith 2007) or government-level (Jamieson & Land & Stephens & Winchester 2008). At the organisational-level, approaches often combine several specific anticipatory, reactionary or remediation elements. While at the national government-level a holistic methodology is often used. This has been the approach in countries such as, the US (e.g., Presidents Identity Theft Task Force), UK (e.g., Home Office), Canada (e.g., Reporting Economic Crime Online (RECOL)), among other countries (Fraud Prevention Expert Group 2007, Jamieson & Land & Sarre & Steel & Stephens & Winchester 2008, Pacific Islands Forum Secretariat 2004). Alternative perspectives on information sharing in IS research have been investigated through social construction, social presence, and task closure theories, while empirically based on small groups within organisations (see Nelson & Cooprider 1996, Miranda & Saunders 2003). A related strand of the literature focuses on information management. This includes: privacy of confidential identity information, and laws, protocols, and standards for protecting this information nationally and internationally (FDIC 2004, Wugmeister & Retzer & Rich 2007).
3 THEORETICAL BACKGROUND

In this paper we investigate information sharing within a knowledge management (Hedlund 1994) context, through a sensemaking in organisations (Weick 1995, Weick & Sutcliffe & Obstfeld 2005), and organisational learning lens (Thomas & Sussman & Henderson 2001). These models (see figure 1) have previously been used in IS research to study banking, manufacturing, education and identity fraud (Cecez-Kecmanovic 2004, Cecez-Kecmanovic & Jerram 2002, Jamieson & Stephens & Winchester 2007b, Janson & Cecez-Kecmanovic & Zupancic 2007). The six identified types of knowledge, corresponding to specific sensemaking levels are shown in Figure 1 with their interactions shown by directional arrows. Cecez-Kecmanovic and Jerram (2002, p.896) explain that the first level ‘individual knowledge’ involves a person’s values, beliefs, assumptions, experiences, and skills, that enable the individual to interpret and make sense of the environment, his/her own actions and the actions of others. The second level, ‘inter-subjective or collective knowledge’ represents shared understanding that emerges through social interaction”. The third level, ‘organisational knowledge’ denotes generic meanings and social structures that emerge in and reproduce within an organisation. The fourth level, ‘knowledge embedded in culture’ assumes a stock of tacit, taken-for-granted convictions, beliefs, assumptions, values, experiences and norms that members of an organisation draw upon in order to make sense of a situation and create meanings at all other levels.

The outer two circles in Figure 1 (left) apply to information sharing across organisations, industries, sectors and jurisdictions – nationally and internationally. The second outer (level 5) circle in the left diagram in Figure 1 named ‘inter-organisational knowledge’ is facilitated throughout organisations and individuals within those organisations (or from sections/departments in) collaboratively sharing information to solve new phenomena e.g., identity fraud (refer Jamieson & Stephens & Winchester 2007b). It brings to an individual organisation, shared knowledge from another or many organisations, and participating individual(s) through their social interaction (levels 1 to 4). While the outermost circle (level 6) ‘intra- and inter-sector knowledge’ encompasses knowledge within and across industry (intra-sector) and sectors (inter-sector). This may span jurisdictions within and across countries (see figure 2, for an Australian example). This will involve synthesising ‘industry culture’, ‘national culture’, professional culture’, and ‘wider cultural contexts’ for sensemaking and learning (see outside level 6 on the left diagram and right diagram in figure 1). As described by Cecez-Kecmanovic and Jerram (2002), and Janson, Cecez-Kecmanovic and Zupancic (2007) the issue is to reach the desired outcome for the phenomena, in this paper namely, to deter, detect, and prevent identity crime via information sharing. The umbrella for this collectiveness could be a committee, association or group.

---

**Figure 1. Sensemaking model of knowledge (left); and learning model in organisations (right)**
The learning model’s (see figure 1, right side diagram) learning loops are explained with examples relevant to sharing information to combat identity fraud in an organisation:

- **Single-loop learning**, “involves adaptive responses: measuring an organisation’s performance, comparing it with its stated goals and taking corrective action to close the gap” (Janson & Cecez-Kecmanovic & Zupancic 2007, pp. 6-7). For example, single-loop learning comprises adjusting a security system to make sure that the system meets privacy standards (legislation, protocols);

- **Double-loop learning** “involves evaluating and changing organisational goals, organisational strategies and mental maps” (Janson & Cecez-Kecmanovic & Zupancic 2007, pp. 6-7). For example, if a system is hacked with the loss of data, a firm is required to rethink its security procedures and develop new mental maps to allow the firm to change its security strategy through information sharing with others (levels 5 or 6 in figure 1); and

- **Triple loop learning** “occurs in response to a realisation that existing mental models and ways of organisational learning no longer suffice” (Janson & Cecez-Kecmanovic & Zupancic 2007, pp. 6-7) e.g., profiling to prevent identity crime (Jamieson & Winchester & Stephens & Smith 2008).

### 4 RESEARCH METHODOLOGY

<table>
<thead>
<tr>
<th>Code</th>
<th>Category</th>
<th>Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bank 1*</td>
<td>1. Head of Fraud – Policy &amp; Strategy</td>
</tr>
<tr>
<td>2</td>
<td>Bank 2</td>
<td>1. Chief Manager Operational Control 2. Fraud Management</td>
</tr>
<tr>
<td>5</td>
<td>Licensing Authority 2*</td>
<td>1. Manager</td>
</tr>
<tr>
<td>6</td>
<td>Telecommunications</td>
<td>1. Fraud Risk</td>
</tr>
<tr>
<td>7</td>
<td>Government Agency 1</td>
<td>1. Compliance, Integrity &amp; Documentation Examination</td>
</tr>
<tr>
<td>9</td>
<td>Government Agency 3</td>
<td>1. Director Internal Audit 2. Internal Issues Manager 3. Programme Delivery 4 &amp; 5 Managers</td>
</tr>
<tr>
<td>10</td>
<td>Government Agency 4</td>
<td>1. Deputy CEO – Corp Services and Regulatory</td>
</tr>
<tr>
<td>11</td>
<td>Government Agency 5</td>
<td>1. Client Account Management</td>
</tr>
<tr>
<td>12</td>
<td>U.S. Criminologist</td>
<td>1. Academic – Professor</td>
</tr>
</tbody>
</table>

*Table 1: Participant Code, Category and Role Key (*Teleconference Participants)*

The research involves a mixed-method approach including: the use of industry-based pilot interview; and cross-sectional participant interviews (quotes are in *italics*, see table 1, for interviewee roles and category). In addition, we use secondary data ‘key word’ database searches for terms such as, information sharing, data matching, computer matching, identity, identity fraud, identity theft, and identity deception to build a better picture of the literature. Recordings were transcribed and transcripts coded using ‘key word themes’ in NVivo 2 (QSR International) qualitative analysis software. Interviewees were provided with transcripts so that they could retract any sensitive information or correct or add details from databases or other sources. Interview duration varied from 45 minutes for teleconferences to approximately 90 minutes for face-to-face interviews. Interviews were conducted (or supervised) by senior experienced academics, industry personnel, and an honours student. Ethical approval was sought and granted by the university ethics committee. A pilot interview was undertaken to test the survey instruments open-ended questions (which are not included here due to space constraints). The four main themes of the survey instrument were: What is Identity Fraud in Your Organisation; Managing Identity Fraud; Identity Fraud Reporting; and Identity Fraud Issues and

Proceedings ECIS 2009
Research. Participants selected were from organisations that issued or used proof of identity documents, personal identifying information or were highly targeted organisations (sectors) by identity fraud perpetrators e.g., Banks, Telecommunication organisations, Licensing or Welfare Agencies.

5 DISCUSSION: INFORMATION SHARING AND LEARNING

5.1 Organisational (Knowledge) Information Sharing

In Australia all financial institutions for bank account opening and government agencies for welfare eligibility are legally bound to identify their customers. Financial institutions are regulated under the Anti-Money Laundering and Counter-Terrorism Financing Act 2006 (similarly, section 326 USA PATRIOT Act 2001). Often protocols are drawn-up for the specific purpose of clarifying the process and types of information that may be exchanged. They are ‘dynamic’ agreements of no legal standing between various parties. Therefore, they are impacted by changes in legislation and guiding principles, and should be amended accordingly. The sharing and exchange of information is not something to be taken lightly, and it is fundamental to the success of any strategy to reduce identity and related crimes.

"Do not under-estimate the intelligence and innovation of the bad guys" (Participant 11). Identity fraud perpetrators have to-date managed to extract a huge economic and social cost of many billions of US dollars from the private and public sectors (Jamieson & Land & Stephens & Winchester 2008). Perpetrators obtain identity details by traditional, mechanical or digital identity replicating devices or via cyber space identity crime methods (Jamieson & Stephens & Winchester 2007a). One way to fight back against identity fraud perpetrators is to match information by sharing proof of identity details presented by the perpetrator to details of the real customer from other public and proprietary sources – especially for remote transactions. Any discrepancies will be: alerted, red flagged, and will require additional evidence of identity or face-to-face presentation. This identity corroboration occurring in the background is very effective and little different to customer service asking ‘non-wallet’ questions (e.g., your mothers’ maiden name) to match with previously supplied identity attributes. When customers supply their identity details, organisations are expected to hold this data under tight security and privacy regimes – often legally defined. There could be serious consequences for getting the identity of the suspected perpetrator wrong, including potential fines, court penalties or reputational loss. Often in the past data breaches have been ‘hushed up’ to protect the reputations of organisations. Now governments are legislating to protect consumers, for example, California's data breach laws.

Besides the common goal to catch identity fraud perpetrators, private organisations’ reasons for information and data sharing with public or other private organisations should be aligned to selfish motives of ‘profits’ (Gal-Or & Ghose 2005, Gordon & Loeb & Lucyshyn 2003). Otherwise these private organisations could be giving away their long-term sustainable competitive advantage (Alavi & Leidner 2001). Competition for these profits may lead to unscrupulous data resellers off-loading ‘dirty data’ (Marx & Reichman 1984). The use of dirty data in data matching processes might reveal a high number of ‘false’ (positives) potential perpetrators being questioned. The aim of identification of individuals through information sharing is to cut off perpetrators’ sources of funding.

The Australian Federal Privacy Act (1988) contains 11 Information Privacy Principles which apply to Commonwealth and Australian Capital Territory (ACT) government agencies. It also has 10 National Privacy Principles which apply to parts of the private sector and all health service providers. The Privacy Commissioner also has regulatory functions under some other laws: the Telecommunications Act 1997 (Commonwealth); National Health Act 1953 (Commonwealth); Data Matching Program (Assistance and Tax) Act 1990 (Commonwealth) - of the Australian Tax Office, Centrelink (a welfare agency) and the Department of Veteran's Affairs and the Crimes Act 1914 (Commonwealth). Our focus is on the Privacy Acts, the Datamatching Program (Assistance and Tax) Act 1990, and the protocols for the conduct of the data-matching program (the statutory data-matching guidelines). In
Australia, some government agencies are exempt from the Australian Privacy Act 1988 (Commonwealth), e.g., the Australian Crime Commission (Australian Crime Commission 2007). From 21 December 2001 the private sector amendments to the Privacy Act 1988 (the ‘Act’) came into force. Schedule 3 of the Act, applies to the private sector. The National Privacy Principles are in addition to existing legislation, guidelines which apply to the private sector. Part IIIA of the Act applies to credit providers and credit reporting agencies. The Credit Reporting Code of Conduct is also a binding law on credit providers and credit reporting agencies – organisations typically targeted by perpetrators. This Code was issued by the Privacy Commissioner (the ‘Commissioner’) under section 18A of the Act in 1991 and is complementary to Part IIIA of the Act. Three other significant areas which are monitored by the Commissioner which affect parts of the private sector are in relation to:
- the collection, storage, use and security of personal tax file numbers by organisations that are authorised or approved to record such information under taxation, assistance agency or superannuation law;
- the disclosure of personal information to law enforcement agencies under Part 13, Division 5 of the Telecommunications Act 1997 (Commonwealth), and
- the handling of personal information under the Anti-Money Laundering and Counter-Terrorism Legislation (2006). Under section 6E (1A) of the Privacy Act 1988 (Commonwealth) certain activities of some small businesses previously exempted are brought within coverage of the Act.

5.2 Inter-Organisational (Knowledge) Information Sharing

The difficulty in current business operations is that organisations, both private and public, are moving away from traditional face-to-face contact with customers to remote cyber space channels – Internet and mobile. “The downside of this situation is that identity fraud perpetrators use it to their advantage as well” (Participant 2:1). This is usually not a problem with genuine customers performing their normal tasks. However, a major problem for organisations is when identity fraud perpetrators use POI and PII that is not their own in the guise of a genuine current or potential customer to extract some illegal benefit or to avoid a cost or loss. It is well established from survey results that there are significant costs to the victim (customer and target entity) to re-establish their identity or situation to a comparable prior position (FDIC 2004). The costs do not only include money and time but also the sociological and psychological impact on them deciding that they can no longer be who they really were when they were born. This goes to the very core of an individual’s ‘identity’. There are many recent instances of ‘identity’ data security breaches where individuals have had their POI details stolen (Main & Robson 2001; Privacy Rights Clearinghouse 2008). Security breaches have occurred more frequently over the past five years with highly sensitive personal data lost or stolen, such as, personal identification numbers and driver’s licence details, or private information like ‘mother’s maiden name’ (Times Online 2007). One cause of this is the high portability of data storage devices e.g., laptops, etc.

One interviewee stated the need to move to alternative methodologies to verify a customer. For example, in less populated areas instances of using a local agent, the solicitor or the policeman or ask the customer to give a range of data about themselves. The details could be read from a driver’s licence, birth certificate, or passport. These are POI documents, with supporting databases that would give the organisation identity verification if the customer were in front of them. At issue now is the need for a way of validating the information provided back to the source of the documents, generally the government agency that issued it, to establish the documents as valid. A valid set of documents still does not tell the validating organisation, in ‘customer-not-present’ situations, that the remotely located person is in fact the person to whom the documents purport them to be. “For example, it could be the son-in-law who’s simply gone into the bedroom and lifted all this data. We need to be able to deal with our customers 24 hours 7 days, no matter what the location is. It’s about getting that veracity underneath the identity that’s the issue, not the documents per se” (Participant 2:1).

Some organisations stated that realistically, the data matching or information sharing tools, the front end tools in use, are not that effective. Examples quoted of front end tools in use, include: the Hunter
(Experian) style system; Decision Point, which is a product through Baycorp (a credit bureau); also a product called Fraud Check via Baycorp for fraud cheque style systems. “I think it’s fair to say, that as we get more sophisticated in the original analysis of facilities (products) we’re preventing more identity related fraud. We’ve now got reasonably sophisticated up front data matching capabilities.” (Participant 2:2). Similarly, Participant 3’s, long-term strategic solution in identity crime involves implementing software now and in the future. “With the Birth Deaths and Marriages systems currently in place, there has been a lot of lobbying with industry representatives, like the Bankers Association and through, the Provider Advisory Group, to push legislative change and get industry (sector) support. (This is) in terms of those state government bodies that are responsible for issuing identification documentation, to try and increase the security around those providers with access to verify the authenticity of those proof of identity documents. So you have to have rules - tell me your mother’s maiden name plus two others (secrets). The detection technology that we have in place for our products - be they electronic banking, cheques, credit cards, or applications – they do pick up a lot of identity fraud, but we just basically want to identify all fraud” (Participant 3:2). However, we need to be aware that it is unlikely technology will solve all the problems of sharing because tacit elements cannot be transmitted (Hislop 2002). Where knowledge is more explicit, sharing would be a lot more amenable with technology.

5.3 Intra- and Inter-Sector (Knowledge) Information Sharing

Participant 1:1, stated that an example of a measure gaining an advantage over perpetrators was “participation in a pilot with Bankers Association and a Licensing Authority Participant 2 and BDM” Often information is passed from the private sector to law agencies as part of a collaborative information sharing. “We do external reporting to XYZ organisation for cheque fraud, to the credit card schemes for credit card fraud, and we’ve also commenced to do some trialling of supplying some fraud data to another government agency (e.g., the Australian Crime Commission see figure 2) and we report fraud data to participant 10” (Participant 2:1). Organisations can also subscribe to commercial organisations that supply identity details to them. Other tools or systems in place by industry organisations check for the identity, verifying phone numbers against electronic White Pages, electoral roll data sets among others. Additional issues regarding data matching and information sharing were “things which are embedded into legacy style controls - checking of signatures on cheques. If the banks still did that, the cost to run that is the cost of control, one of the traditional controls. The cost of asking the wallet and non-wallet questions at the call centre is also a cost to control at a level of unauthorised – it’s also got privacy implications and is a fraud prevention control” (Participant 2:2).

In May 2006 the Australian Federal Government announced an increased focus on identity crime with the establishment of a national Document Verification Service that will allow for ‘blind’ inter-departmental cross-checking of documents (Wentworth Courier 2006). If there has been a hit on a person’s identity, by and large, it will not be captured by the credit reporting bureaus, which is a big problem for banks that use third party verification services. People are having their whole credit history basically altered or adulterated as a result of somebody helping themselves to their identity and perpetrators identity fraud. Government and taxpayers trust government agencies with delivering services, and there would be community revolt if information came out that showed inappropriate people were receiving more payments than they were entitled to. “We actually had a case where we waited until the individual accessed the false bank account and tapped him on the shoulder while he was at the ATM. The Australian Federal Police were there to make the arrest” (Participant 9:5).

Within the tax system identity fraud has been shown to affect the revenue collection, and goods and services tax fraud areas. For instance, the act of people falsely claiming input credits would be considered more fraud than identity fraud. If, on the other hand they are actually pretending to be some other entity in order to make the claim, then it becomes identity fraud. With the Higher Education Contribution Scheme, there is evidence that students are using one tax file number, and acquiring another separate tax file number after completion of their degree, thus, effectively setting
their education debt aside and permitting them to avoid paying down their debt. Other examples are income tax splitting, gaining employment whilst on welfare benefits, and a family tax benefit. This is referred to as ‘double-dipping’, basically an individual claims through the welfare system and through the taxation system; if the identity does not match between the welfare and the taxation systems they avoid detection. With the data matching processes implemented under various pieces of legislation, many identity fraud and fraud perpetrators are being caught. But the data matching is premised on data quality and if the data quality is ‘dirty’ (Marx & Reichman 1984) in the taxation system and in the private health funds, there is a problem. With most government benefits a customer needs a tax file number, so that even a slight change (in proof of identity) reduces the efficiency of the matching process and produces a low level of confidence in any matches detected. We can’t data match on everything. We can’t match data if evidence of wrong doing is not strong enough. In addition, we rely on the processes and documents of others. There is still the Privacy Commissioner who is quick to remind us that it is not mandatory in a lot of instances to quote your tax file number. Although there are some strong inducements like 50% tax rate penalty if a number is not supplied (Participant 11:1).

Immigration authorities around the world have a ‘person alert list’ and a ‘document alert list’. If the proof of identity (POI) document is stolen (identity theft), or altered, or totally fictitious (identity deception), it goes on an ‘alert’ and a risk rating algorithm would pick it up from that information and other sources and try and stop the use of the POI document. For the organisation, the problem is service versus compliance and integrity of the information and POI documents through the vetting process. The community has an expectation that the tax file number, drivers licence, passport are correct (no errors, duplicates, or false entries). Government agencies solve these POI database issues by matching records in the databases in Australia on identity details. Agencies look at the exceptions, for example, ages during matches with death, birth, citizenship and immigration records. In Australia, many citizens are registered in the Centrelink registry, in the tax system, and in the Health Insurance Commission. “The problem is we are now dealing with databases. Some are client-based registers; others have an event or a vehicle type of register. The three axes of data matching tend to be name, address and date of birth. In a perfect world we remove any privacy or that sort of problem; we could all do it by data matching or database security. We are certainly looking at data matching. Essentially, it is a virtual database of the population at any one time. It is not something one agency could do on its own. It is not a single issue, it is not a single agency issue and that is the problem. We couldn’t just swap identity details with any other agencies” (Participant 11:1).

The major issues restricting how entities would implement information sharing are particularly to do with privacy issues. “Which is why it’s important to recognise that the personal information isn’t being passed back to the bank, and that the banks, or whoever, don’t have unfettered access. They can’t have access to the systems themselves, they are just getting that sort of a yes, no answer. And that, of course, limits how much verification you can do. I think that’s the major issue. You could centralise data within government, probably much more easily. But again, the whole privacy issue can’t be cross matched ... all the data matching issues that come up in terms of privacy are obviously something that will have to be looked at” (Participant 10:1). The Australian Crime Commission have an identity fraud register (see figure 2, arrows show information flows), where the major banks and other organisations, including the commonwealth and international agencies, are providing information about identity fraud. This information has actually been compared to other entities databases and there have been a significant number of hits on people on the databases. “The banks can match up, not just whether it’s a fake birth certificate, but whether the person who holds it is dead. Extending that, obviously there are all sorts of issues about whether you could have that sort of access to Medicare information, passport and immigration” (Participant 10:1).

Figure 2 shows the Australian Crime Commission's an Australian Federal agency’s main Australian Criminal Intelligence Database and collaborative information sharing initiatives between regional countries (Pacific Islands Forum Secretariat 2004), and the Australian Commonwealth and state governments Identity Fraud Register. The Identity Fraud Register provides valuable intelligence to law enforcement by keeping a register of identities suspected of being fraudulent. The Australian Law
Enforcement Intelligence Net (ALEIN) “is a secure national extranet used by all Australian police services, the New Zealand Police, state Crime Commissions and a large number of government law enforcement agencies. ALEIN provides a dissemination capability for the Australian law enforcement community and enables cooperative intelligence sharing across jurisdictional boundaries. ALEIN provides real-time, secure intelligence communications between the ACC, police jurisdictions, and other law enforcement agencies across Australia” (Australian Crime Commission, ACC 2004, p. 31).

The police through the Australian Federal Police share information internationally through Interpol.

Figure 2. Australian Crime Commission Information Sharing Regime to Combat Identity Fraud

Information sharing interrelationships shown in Figure 2 also occurs in other government agencies (and they may interact with figure 2 databases). There are inhibitors on this sharing based on what is allowed under legislation. Entities have to work within these parameters. In Australia, driver licensing authorities have access to a national system, which allows authorities to look at licensing information from different states. In a government agency “cross-organisational participation and information sharing agreements do not exist holistically. Participant 8 has access to the Australian Electoral Commission, Participant (11), BDM, and Participant (7) databases but the information flow is one way for privacy reasons”. Similarly outside Australia, information sharing issues may be country specific, depending on what the other country’s privacy legislation and agreements with Australia or another jurisdiction are (see Fraud Prevention Expert Group 2007). In some cases it may actually be easier to share information, but in many other cases it is much more difficult because some countries may be unwilling to provide any information, unless they can see they can get a direct benefit out of it. “The international cooperation branch, that looks specifically at issues to do with countries overseas where Australia has an interest in getting cooperation on certain issues, on sharing information with those countries, and establishing formal systems and structures. It might be with memorandums of understanding or agreements with those countries that we’ll share information on certain issues” (Participant 7:1).

5.4 Organisational Learning

Organisational learning within participant interviewee’s organisations (Table 1, codes 1-11) was at the single-loop learning phase with respect to the identity crime phenomena. This was mainly due to identity crime within an organisation seldom seen as a problem from a whole-of-organisation perspective. It was often seen as a problem for the fraud department. Further interviews with the participant organisations to obtain additional organisational learning information may shed light on whether improvements have occurred in individual organisations. An important observation for the organisational learning model is when a new organisation (association, group, or committee) is set up to share information and learn from many other organisations with a purpose (or intent) to better understand a new phenomenon such as identity and related crimes. This reinforced our choice of lens.
6 CONCLUSIONS AND FURTHER RESEARCH

The sharing of identity attribute information within and across organisations or sectors has real value for those able to do so under current (Australian) privacy and related legislation. Information sharing has been proven to be an effective methodology to detect, deter and prevent identity and related crime perpetrators by many Australian organisations. As more identity fraud cases move through the judicial system the penalties perpetrators receive will act as a further deterrent to others. However, a substantial amount of privacy and related legislation enacted to protect individuals and entities identity attributes (sensitive information) may in some ways actually be helping perpetrators. This may hinder an organisation’s ability to collaborate through information (data) sharing, as the perpetrators hide behind privacy regimes. We applied the extended sensemaking model of knowledge management in organisations (Jamieson & Stephens & Winchester 2007b) and organisational learning model (Janson & Cecez-Kecmanovic & Zupanic 2007) to an intra- and inter-sector context, grounded from Australian based interviews. An example was also shown in Figure 2 highlighting the national and international effect of information sharing to mitigate identity fraud where perpetrators act across jurisdictions. This may create inter-jurisdictional problems or loopholes that perpetrators may exploit.

Interviewee feedback suggests we need to do more information (data) sharing via bilateral or multilateral communication and via linking and interrogating data sets or databases. A second finding is that the data to be matched or shared information needs to be free of errors as possible. Third, we need effective legislation that allows meaningful data sharing, where POI details (presented by the perpetrator) are matched to the bone fide customer details from other public and proprietary sources. Finally, organisational learning was mostly single-loop. Further interviews may detect learning (loop) improvements. These findings from our experts help to support an AC Nielsen poll finding that two-thirds of Australians were willing to sacrifice privacy and civil liberties for protection – especially against terrorists (Dodson 2006). Future research could use these models in different country settings.

References


Proceedings ECIS 2009


